AN INTEGRATIVE SYSTEM APPROACH TO FLASH FLOOD HYDROLOGY

ASHISH MANOJ JASEETHA

AN INTEGRATIVE SYSTEM APPROACH TO FLASH FLOOD HYDROLOGY

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN
(Dr. -Ing.)

von der KIT-Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von
Ashish Manoj Jaseetha M.Tech.
aus Kerala, Indien

Tag der mundlichen Prüfung:
24 April 2026

REFERENT: Prof. Dr. Erwin Zehe
KORREFERENT: Prof. Dr. Axel Bronstert

Karlsruhe 2026

For my father,

whose quiet wisdom and practicality still walks beside me,

guiding my steps long after he has gone.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Systems approach, having its roots in operations research and decision making needs during World War II, provides a structured framework for understanding complex, interacting components within natural and engineered systems. Rather than focusing on isolated elements, it emphasises feedback, nonlinearity, and system-wide responses. In hydrological sciences, system analysis provides a lens for interpreting catchments as integrated systems in which dynamic meteorological forcing drives runoff generation, explicitly considering the landscape controls of the problem.

In this thesis, I characterise and solve the three classes of system problems, namely *Analysis*, *Design* and *Inverse*, with a specific focus on improving hydrological modelling of flash floods. My overarching goal is to improve preparedness for such extreme events, particularly in data-scarce regions. In Chapter 2, I utilise gradient-based simplifications to set up hillslope scale hydrological models and initialise them using a climate reanalysis product to predict convective storm-driven summer flash floods in four headwater catchments over southwest Germany. The results indicate that utilising hydrologic parameter transfer from nearby experimental catchments can eliminate the need for manual calibration when making an initial estimate of flash flood dynamics, thereby making strides toward solving the Prediction in Ungauged Basins (*PUB*) problem.

Chapter 3 opens by addressing the challenge of linking modelling capabilities across scales for flash floods. A coupled hydrologic–hydraulic model is setup at the meso-catchment scale (approximately 200 km^2), using a vectorised depiction of representative hillslopes at the subcatchment level and standard one-dimensional kinematic routing to upscale simulations to the catchment scale. I apply the proposed framework to a historical flash flood in Baden-Württemberg, Germany, enabling the reconstruction of flood dynamics in poorly gauged headwater regions. Based on these results, nature-based flood mitigation measures are explored as alternatives to conventional structural flood-control reservoirs.

In Chapter 4, I leverage recent advances in data-driven learning to develop a regional neural network model that performs precipitation inversion from discharge. The main goal is to examine if discharge data can be used to improve precipitation estimates at the catchment scale for better extreme event characterisation. The trained model

was then used to predict driving storm forcings at out of sample catchments. Chapter 5 integrates the three problems of the thesis and demonstrates, through an exemplary analysis, the importance of linking different approaches for a holistic hydrologic assessment at the catchment scale. The chapter concludes with an outlook section outlining future directions to advance the methodologies proposed in this work.

Overall, this thesis contributes to improved hydrological modelling and enhanced preparedness for extreme events by (a) advancing prediction capabilities in data-scarce regions, (b) bridging spatial scales in flash flood modelling, and (c) inversely correcting precipitation estimates to better reflect catchment responses.

ZUSAMMENFASSUNG

Die Systemansatz, deren Wurzeln in der „Operations Research “ sowie in der Notwendigkeit einer Entscheidungsunterstützung während des Zweiten Weltkriegs liegen, bietet einen strukturierten Rahmen, um komplexe, miteinander interagierende Komponenten in natürlichen und technischen Systemen zu verstehen. Statt isolierte Elemente zu betrachten, stellt sie Rückkopplungen, Nichtlinearitäten und systemweite Reaktionen in den Mittelpunkt. In der Hydrologie ermöglicht die Systemanalyse, Einzugsgebiete als integrierte Systeme zu begreifen, die auf dynamische meteorologische Eingangsgrößen reagieren und Abfluss erzeugen, wobei die räumliche Struktur der Landschaft als wesentlicher Teil des Problems berücksichtigt wird.

In dieser Arbeit charakterisiere und bearbeite ich drei Arten von Systemproblemen – *Analyse*, *Design* und *Inversion* – mit einem besonderen Fokus auf der Verbesserung der hydrologischen Modellierung von Sturzfluten. Das übergeordnete Ziel besteht darin, die Vorsorge auf derartige Extremereignisse zu verbessern, insbesondere in datenarmen Regionen. In Kapitel 2 nutze ich gradientenbasierte Vereinfachungen, um hydrologische Modelle auf der Hangskala aufzusetzen und diese mithilfe eines Klima-Reanalyseprodukts zu initialisieren, um konvektiv getriebene sommerliche Sturzfluten in vier Kopfeinzugsgebieten in Südwestdeutschland vorherzusagen. Die Ergebnisse zeigen, dass durch den Transfer hydrologischer Parameter aus benachbarten experimentellen Einzugsgebieten eine manuelle Kalibrierung für eine erste Abschätzung der Sturzflutdynamik nicht mehr notwendig ist. Damit wird ein Beitrag zur Lösung des Problems der Vorhersage in unbeobachteten Einzugsgebieten (Prediction in Ungauged Basins, *PUB*) geleistet.

Kapitel 3 beginnt mit der Herausforderung, Modellierungsfähigkeiten für Sturzfluten über Skalen hinweg zu verknüpfen. Dazu wird ein gekoppeltes hydrologisch-hydraulisches Modell auf der Mesoskala (ca. 200 km²) aufgebaut, das eine vektorisierte Darstellung repräsentativer Hänge auf Teileinzugsgebietsebene mit einem standardisierten eindimensionalen kinematischen Flood-Routing kombiniert, um Simulationen auf die Einzugsgebietsskala zu übertragen. Das vorgeschlagene Modellkonzept wird auf eine historische Sturzflut in Baden-Württemberg angewandt und ermöglicht die Rekonstruktion der Hochwasserdynamik in schlecht beobachteten Kopfeinzugsgebieten. Aufbauend auf diesen Ergebnissen werden natürliche Hochwasserschutzmaßnahmen als Alternative zu klassischen, baulichen Hochwasserrückhaltebecken untersucht.

In Kapitel 4 nutze ich neuste Fortschritte im datengetriebenen Lernen, um ein regionales neuronales Netzwerk zu entwickeln, das Niederschlag aus Abflussdaten invers bestimmt. Ziel ist zu prüfen, ob Abflussdaten genutzt werden können, um Niederschlagsschätzungen auf Einzugsgebietsskala zu verbessern und Extremereignisse dadurch besser zu charakterisieren. Das trainierte Modell wird anschließend eingesetzt, um die antreibenden Niederschläge in unabhängigen, außerhalb des Trainingsdatensatzes liegenden Einzugsgebieten zu prognostizieren. Kapitel 5 integriert die drei Fragestellungen der Arbeit und verdeutlicht anhand einer exemplarischen Analyse die Bedeutung der Verknüpfung unterschiedlicher Ansätze für eine ganzheitliche hydrologische Bewertung auf Einzugsgebietsskala. Das Kapitel schließt mit einem Ausblick, der zukünftige Schritte zur Weiterentwicklung der in dieser Arbeit vorgeschlagenen Methoden.

Insgesamt trägt diese Arbeit zu einer verbesserten hydrologischen Modellierung und zu einer besseren Vorbereitung auf Extremereignisse bei, indem sie (a) Vorhersagefähigkeiten in datenarmen Regionen weiterentwickelt, (b) räumliche Skalen in der Modellierung von Sturzfluten überbrückt und (c) Niederschlagsschätzungen invers korrigiert, um Einzugsgebietsreaktionen besser abzubilden.

Part I

INTRODUCTION

INTRODUCTION

Hydrology sits at the crossroads of an engineering discipline and earth system sciences (Sivapalan, 2018). As a branch of earth system sciences, it focuses on understanding how water interacts with climate, biosphere, pedosphere, and geology to influence environmental processes. The ideal way is to transfer this knowledge to engineering hydrology, to make informed decisions about water management. (Montanari et al., 2013) and to design infrastructure, such as reservoirs, urban water supply systems, and drainage networks. Tackling hydrological problems becomes quite complicated as the earth system and engineering sides add different layers of complexity (Dooge, 1986), and translating process insights into routine engineering practice often remains painfully slow (Blöschl et al., 2019b).

Why is hydrology unique?

This issue becomes even more critical in a warming climate marked by an increase in extreme events. The sixth assessment report of the Intergovernmental Panel on Climate Change (Arias et al., 2021) highlights that there is growing evidence of human-induced changes leading to intensified extreme precipitation events, which is supported by high confidence. However, there is low confidence regarding changes in regional flood frequency resulting from these increasing precipitation amounts. While flash floods (caused by localised high intensity rainfall exceeding the infiltration capacity of the soil) have increased in most regions (Göppert, 2018; Meyer et al., 2022), droughts have also shown an increasing tendency to occur (Sarhadi et al., 2018). These complexities pose a multitude of challenges for both scientists and practitioners. From an earth system perspective, water, while retaining the memory of its driving meteorological forcings (Zehe et al., 2014), interacts with soils, vegetation, geology, and ecosystems as it navigates through a catchment. These interactions are often non-linear in nature (Bronstert et al., 2023; Zehe et al., 2005) and quite difficult to represent in simplified top-down modelling approaches (Hrachowitz and Clark, 2017).

Extremes make life difficult

The human aspect adds more complexity (Remmers et al., 2025) from the engineering side. Humans heavily modify the water cycle (United States Geological Survey, 2022) as a result of infrastructural and socio-economic changes. Events such as Germany's severe flash floods in 2016 (driven by convective storm clustering - Göppert, 2018; Piper et al., 2016), followed just two years later by an extreme drought in 2018 (Boeing et al., 2022; Rakovec et al., 2022) create significant challenges because they require infrastructure and management sys-

Human perspective

tems that can handle contrasting types of water stress. Improved hydrological modelling approaches can help in this direction by improving our ability to simulate and anticipate such events.

1.1 SYSTEMS APPROACH

Given any natural or man-made system, denoted by the system operator S, which acts on an input $x(t)$ to produce an output $y(t)$ such that

$$y(t) = S[x(t)] \tag{1.1}$$

Three types of problems

Three distinct problems relating $x(t)$, S and $y(t)$ can be distinguished. *System Analysis* or the forward problem, involves finding the output, $y(t)$ given perfect knowledge about the inputs $x(t)$ acting on a well-defined system S. The design problem (*System Design*), on the other hand, is concerned with establishing the system, S assuming some known $x(t)$ acts on the system to produce (either acceptable design or assumed values of) outputs $y(t)$. Finally, inverse problems (*System Inverse*) tries to map the inputs, $x(t)$ based on knowledge of the system, S and the outputs, $y(t)$. Visualising catchments as integrated systems that act as dynamic filters (Zehe et al., 2014) on atmospheric forcing, transforming it into output fluxes such as streamflow, provides a natural basis for applying a systems approach (Miser, 1980) to complex hydrological problems.

1.2 THE HYDROLOGICAL PROBLEM(S)

Hydrological counterparts

The three classes of problems (*Analysis, Design, and Inverse*) have, unsurprisingly, counterparts in the hydrological sciences. The analysis problem is directly related to classical rainfall-runoff modelling, where we seek to model, given that it rained a fixed amount, the level of water observed in our streams and rivers. Design problems can be seen as engineering problems of constructing flow channels and flood defence structures to ensure that, given a fixed amount of rain, the safe carrying capacity of our streams isn't exceeded and the cost to humans (both economic and social) is minimised. While the inverse problem is less commonly discussed in hydrology (compared to contemporary progresses made in image processing and robotics - Vahrenkamp et al., 2013), it is sometimes vital to stress that often (at least not *everything everywhere all at once*) we do not know what 'fixed' amount it actually rained and it makes sense to try to understand this by observing the water in the streams along with our previously established knowledge of hydrological systems.

Catchments are crucial for understanding hydrology, as they are

often defined as the natural boundaries (system) within which the different processes interact to influence the response. The law of mass conservation (Lavoisier, 1789) nicely encapsulates the essence of the three system problems in hydrology by providing the fundamental constraint linking inputs (I), storage (S), and outputs (O) for a catchment. The catchment water balance can be described as:

$$\frac{dS}{dt} = I - O \tag{1.2}$$

$$\frac{dS}{dt} = P - ET_a - Q \tag{1.3}$$

where P is the spatially averaged precipitation from the atmosphere that reaches the surface of the catchment, Q is the total runoff (including both surface and groundwater flow) that discharges at the catchment outlet, ET_a is the spatially averaged actual evapotranspiration loss from the catchment system, and S denotes the integral catchment storage, representing the total water stored in the system.

Catchment as a hydrological system

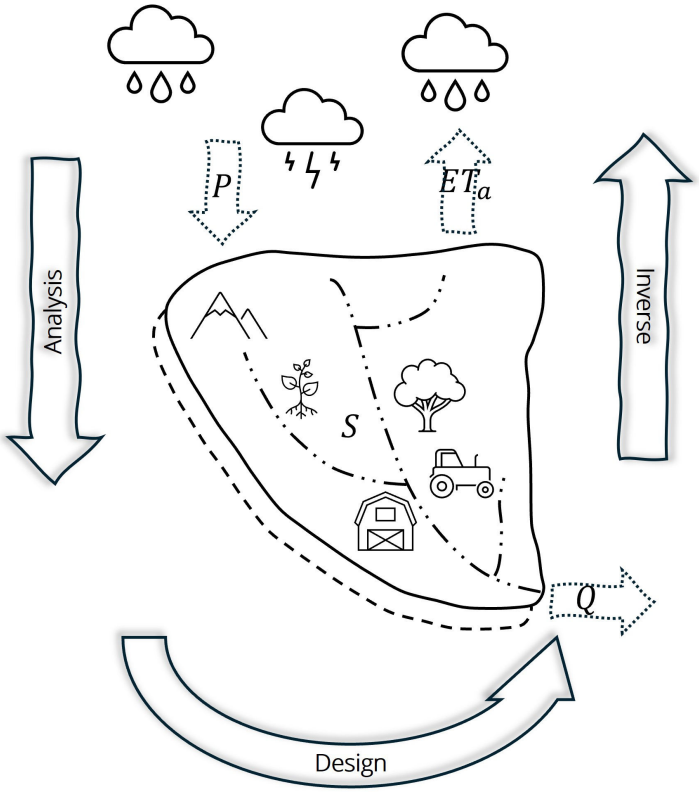

Figure 1.1: Conceptual visualisation of an integrated hydrological system, showing the three classes of problems

1.3 MOTIVATION

The main motivation of my thesis is to characterise and address these three fundamental problems in hydrology, focusing specifically on extreme events, particularly flash floods caused by high rainfall intensities. While a multitude of definitions exist for flash floods (Borga et al., 2008; Marchi et al., 2010; Meyer et al., 2022), in this work, they are defined as floods caused by short-duration, intense (convective) rainfall during summer. These events often occur in rather poorly (hydrologically) observed regions, and have much in common with the Predictions in Ungauged Basins (PUB) problem. The PUB initiative (Hrachowitz et al., 2013; Sivapalan et al., 2003) aimed to improve hydrological predictions in ungauged basins by reducing uncertainty in existing models and developing new models that better represent the space–time variability of hydrological processes.

Analysis

As I briefly mentioned in the introduction, the earth system and engineering discipline sides of hydrology often pose confounding challenges, particularly in the case of extremes. While the catchment system can be approximated by simple, linear top-down approaches (Hrachowitz and Clark, 2017) for analysis during most of the hydrological year, these approximations severely mischaracterise the system behaviour during extremes. An example of such a misrepresentation is when confronted with high rainfall intensities during convective storm events, the soil storage of the catchment system sometimes activates certain preferential flow pathways (Bronstert et al., 2023; Villinger et al., 2022). Not accounting for them would introduce a bias into modelling the response during the event.

Design

On the engineering side, designing a system for an extreme flood such that the resulting water levels and discharges are within the prescribed safety limits is an optimisation problem (Roy et al., 2008) subject to constraints on safety, cost, and environmental impact. However, as shown in experiences around the world, flood protection measures have sometimes been ineffective during such extreme events and in some rare cases even compounded the flood hazard (Suresh and Hossain, 2025). While structural flood defence measures have long been the norm, recent advances in so-called Nature-based Solutions (NbS; Guillaume et al., 2025) that provide environmental benefits irrespective of extremes have started gaining prominence, particularly in agricultural and other vulnerable settings.

It is important to note that while we assume perfect knowledge of the system, observational uncertainties (Nearing et al., 2016) plague all the terms in the catchment water balance. This is usually taken care of in the analysis problem by predicting an ensemble or reporting

the spread instead of a single value prediction (Klotz et al., 2022), and in design by using factors of safety (Terzaghi et al., 1996). The inverse problems offer another rather appealing, unconventional way forward for hydrology (Kirchner, 2009). Assuming that a perfect relationship can be learned about the rainfall-runoff relationship as the forward problem, this also means that such a relationship should also be invertible in the backward direction. Thus, we can invert the *effect* to get back to the *cause*. While such an inverse problem usually has reduced practical benefits, in the case of flash floods, this mapping could help to correct precipitation biases, as more often than not the *effect* (higher discharge rates in the stream) is observed more clearly than the *cause* (localised high intensity rainfall rates).

Inverse

1.4 STRUCTURE OF THESIS

1.4.1 *Chapter Organisation*

In the following chapters, I present three studies that directly address the three problem classifications: *analysis*, *design*, and *inverse*, as they relate to flash floods. In Chapter 2, I tackled challenges related to predicting flash floods in rather data-scarce settings. By integrating a spatially distributed, process-based model with climate reanalysis products for model initialisation and using transfer learning from experimental catchments for parameterisation, we successfully tackle the Prediction in Ungauged Basin (PUB) problem in four poorly gauged headwater catchments. Chapter 3 evaluates the effectiveness of current flood defence standards in a mesoscale catchment under high-intensity rainfall rates (subsequently triggering flash floods) that have become the norm in a warming climate due to anthropogenic climate change. I also propose distributed nature-based solutions (NbS) as a viable alternative to mitigate flash floods and demonstrate that these environmental measures can attenuate peak storm runoff and provide downstream benefits throughout the catchment, even when placed in only headwater catchments. In Chapter 4, an inverse model is set up using machine learning methods by tapping into large sample hydrological datasets spanning over a thousand catchments across Europe, to infer rainfall from runoff (thus inverting the more commonly used rainfall-runoff relationship). The learned relationship is then used to do *'a posteriori'* correction of catchment scale extreme event precipitation estimates given by reanalysis products. Finally, I present an exemplary analysis that links all three problems and offer a few concluding remarks in Chapter 5.

1.4.2 *The Venue*

Kraichgau

The Kraichgau region (Bundesamt für Naturschutz (BfN), 2026; Wikipedia contributors, 2025) in Baden-Württemberg, Germany, is located between the cities of Heidelberg, Karlsruhe, and Heilbronn. It is a hilly landscape characterised by a dense pattern of mostly rural settlements, many of which date back to the Middle Ages, embedded within a mosaic of agricultural land and forest. The presence of fertile loess soil (deposited during the ice age as silt) and its relative milder climate makes it one of the breadbaskets of southern Germany. Alongside cereals, the landscape supports widespread orchards and viticulture (the famous *Badische-Weinstrasse*, also routes through Kraichgau), as well as the cultivation of potatoes, sugar beet, and (historically important) tobacco.

At the same time, the area has been repeatedly affected by severe floods over the past three decades (Disse and Engel, 2001; Villinger et al., 2022; Zehe et al., 2005). In particular, the winter flood in late December 1993 and the subsequent summer flash flood in 1994 highlighted the region's high flood hazard and exposure. The dispersed settlement pattern, with populations located within and between cultivated fields and forested hillslopes, further increases the vulnerability to pluvial and fluvial flooding. The region was again affected by the convective storm clustering in the summer of 2016 in Germany, and faced widespread flash floods (Bronstert et al., 2018; Göppert, 2018), which even led to the overtopping of flood defence reservoirs in the region.

Weiherbach project

The region was also home to the multidisciplinary Weiherbach project (1989-1997; Plate and Zehe, 2008) with the aim of developing a physically based numerical model capable of describing all relevant water and solute fluxes within and out of a small rural catchment (the Weiherbach catchment). During the course of the project, extensive data (Delbrück, 1997; Schäfer, 1999; Schierholz et al., 2000) regarding the hydrological, agricultural and soil properties of the research area were collected.

Elsenz Schwarzbach

During the 2016 summer floods in southern Germany, headwater areas in the Elsenz Schwarzbach catchment in Kraichgau faced intense storms leading to flash floods and widespread destruction. Unfortunately, there was limited data available for quantitative evaluation of the flood response, as these regions were ungauged except for water level recordings in the flood reservoirs downstream. In Chapter 2, I chose four such headwater catchments (up to 6 km^2) severely affected by the 2016 floods in the Elsenz Schwarzbach, Kraichgau and modelled the storm response while tackling issues related to

data scarcity. The entire mesoscale catchment (196 km^2) of the Elsenz Schwarzbach was selected in Chapter 3 to reconstruct the impacts of the 1994 summer flood on population centres and to examine the adequacy of the current flood design standards. For inverting the classical rainfall-runoff relationship using machine learning models in Chapter 4, we make use of recent advances in large sample hydrology (Kratzert et al., 2023), and try to infer the inverse relationship using a training dataset consisting of 1800 European catchments, ranging from 2 to 150,000 km^2 in size. After training, an out-of-sample test is conducted across four catchments, including the Elsenz Schwarzbach.

1.4.3 *The Model(s)*

All models are simplified representations of a system's reality. We build models by expressing our understanding of the system in the form of mathematical equations, giving (sometimes undue) importance to the parts and processes that we think are more important. As quoted by the British statistician George Box *'All models are wrong, but some are useful. The practical question is how wrong do they have to be to not be useful'*. Applying the right models for the right problems can ensure that the solutions obtained are both efficient and accurate.

To model or not to

Hydrological models can be broadly classified (Fatichi et al., 2016; Hrachowitz and Clark, 2017) as giving either a macroscopic (top-down) or microscopic (bottom-up) system perspective for catchments. Models following the bottom-up philosophy (also more commonly referred to as physically based models) rely on an explicit representation of each process in a catchment. These models are often computationally expensive and parameter intensive but can provide detailed information on the various state variables and fluxes driving the flow in a system. On the other end of the spectrum, top-down models (also called conceptual or bucket type models) often rely on a macroscopic representation of the system processes and driving gradients. Data-driven models which rely on learning from the data explicitly have also gained prominence over the last few years (Kratzert et al., 2018). Long Short-Term Memory (LSTM; Hochreiter, 1998) networks is a type of data-driven model that has been recently widely used for hydrological modelling.

Choosing the *right* model for the *right* problem often also involves some subjectivity on the part of the modeller. In Chapter 2, I use the spatially distributed, process based (bottom-up) model CATFLOW (Zehe et al., 2001) at the four headwater catchments as an explicit representation of the catchment system is required to account for the intricacies of Hortonian overland flow (due to the high intensity associated with convective storm) and the onset of preferential flow.

Horses for courses

I then also test the representative hillslope (Loritz et al., 2017) concept which tries to find the right balance between the explicitness required for such process based models with simplicity considerations stemming from parsimony considerations. The CATFLOW model and representative hillslopes was again made use of in Chapter 3 to model flash floods over the larger mesoscale catchment. This time linking the individual representative hillslopes with a standard 1D kinematic wave routing (Fenton, 2019) helps to account for the preferential flow processes even at the larger scale. For the inverse problem (Chapter 4), motivated by their success in rainfall-runoff predictions at larger scales (Kratzert et al., 2018) and the universal approximator theorem (which states a neural network based model can approximate any continuous function to any desired degree of accuracy, provided certain conditions are met; Hornik et al., 1989), I use the standard LSTM model structure as commonly used in runoff modelling (Acuña Espinoza et al., 2024) with the change that instead of using predictors including precipitation to simulate discharge, I use future discharge and other predictors to model catchment average precipitation sums inversely.

1.5 OBJECTIVES

The overarching goal of this thesis is to improve both the modelling of flash floods and the preparedness for such events by adopting a hydrological systems approach. The following research questions are proposed to guide this work:

- **Analysis -** To what extent can the initialisation using climate model simulations and parameter transfer enhance the reliability of modelled responses in data-scarce headwater catchments during flash floods while accounting for preferential flow?
- **Design -** Can modelling simplifications developed for headwater catchments be scaled to the mesoscale using standard hydraulic routing schemes to assess the design adequacy of existing flood mitigation measures and to propose nature-based solutions for addressing increasing flood risk at these scales?
- **Inverse -** Does the discharge measured at the catchment outlet retains sufficient memory of the underlying hydrological processes to allow reliable inversion for estimating the driving precipitation forcings of extreme events?

Part II

TOWARD FLASH FLOOD MODELING USING GRADIENT RESOLVING REPRESENTATIVE HILLSLOPES

This study is published in the scientific journal *Water Resource Research* (WRR). The remainder of part II is a reprint of:

Manoj J, A., Loritz, R., Villinger, F., Mälicke, M., Koopaeidar, M., Göppert, H., and Zehe, E. (2024). "Toward flash flood modeling using gradient resolving representative hillslopes." Water Resources Research, 60, e2023WR036420. https://doi.org/10.1029/2023WR036420.

2

TOWARD FLASH FLOOD MODELING USING GRADIENT RESOLVING REPRESENTATIVE HILLSLOPES

ABSTRACT

It is increasingly acknowledged that the acceleration of the global water cycle, largely driven by anthropogenic climate change, has a disproportionate impact on sub-daily and small-scale hydrological extreme events such as flash floods. These events occur thereby at local scales within minutes to hours, typically in response to high-intensity rainfall events associated with convective storms. In the present work, we show that by employing physically based representative hillslope models that resolve the main gradients controlling overland flow hydrology and hydraulics, we can get reliable simulations of flash flood response in small data-scarce catchments. To this end, we use climate reanalysis products and transfer soil parameters previously obtained for hydrological predictions in an experimental catchment in the same landscape. The inverted mass balance of flood reservoirs downstream is employed for model evaluation in these nearly ungauged basins. We show that our approach using representative hillslopes and climate data sets can provide reasonable uncalibrated estimates of the overland runoff response (flood magnitude, storm volume, and event runoff coefficients) in three of the four catchments considered. Given that flash floods typically occur at scales of a few $km^2$and in ungauged places, our results have implications for operational flash flood forecasting and open new avenues for using gradient resolving physically based models for the design of small and medium flood retention basins around the world.

2.1 INTRODUCTION

As early as 2008, the Organisation for Economic Co-operation and Development (OECD) highlighted climate change and hydrometeorological extremes as among the most pressing challenges facing humanity. Flood events, a key component of these extremes, manifest at varying spatial and temporal scales, each driven by distinct meteorological conditions. Flash floods, for instance, occur on local scales within a span of minutes to hours. These events are triggered by high-intensity rainfall from convective storms, often resulting in strong infiltration excess and significant Hortonian overland flow (Bronstert et al., 2018; Marchi et al., 2010; Marchi et al., 2016; Meyer et al., 2022; Ruiz-Villanueva et al., 2012). While these floods pose immediate risks, such as loss of human life, their consequences extend to long-term impacts like soil erosion, sediment transport, and subsequent deterioration of water quality and soil fertility, particularly in agricultural settings. On the other end of the spectrum are large-scale riverine floods, which occur due to synoptic scale low-pressure systems characterized by widespread and sustained precipitation. Unlike flash floods, these events are governed by capacity-controlled runoff formation processes like saturation excess, known as Dunne overland flow (Dunne, 1978), and subsurface storm flow. Additionally, flood routing in channel networks and snowmelt contributions, play crucial roles (Blöschl et al., 2007). This stands in contrast to the Hortonian overland flow (Horton, 1933) typically observed in flash floods driven by convective storms and affected by preferential flow processes (Bronstert et al., 2023).

Flood forecasting and risk management have to cope with both types of flood events, and both are naturally highly sensitive to climate change (IPCC, 2021). The largest observed floods in many European rivers have occurred in the last three decades, which count among the most flood-rich periods in the past 500 years (Blöschl et al., 2020). The same holds true for local flash floods. For instance, 22 flash floods in southwest Germany occurring in the past 20 years, had estimated design return periods exceeding 500 years (Göppert, 2018). While return periods of that magnitude are uncertain, they indicate nevertheless that these are exceptional events. In a broader context, the finding is in line with the recent accumulation of flash floods in Europe (Meyer et al., 2022). This likely reflects the already ongoing acceleration of the hydrological cycle, with expected increasing frequencies of more intense convective rainstorms (Bürger et al., 2019; Mueller and Pfister, 2011) and flash floods due to Clausius–Clapeyron scaling (Pall et al., 2007). This is alarming, as the flash flood series in the summer of 2016 alone caused about €2.5 bn of damage in Germany (Munich Re, 2016). All this recent evidence calls for improving

the current standards in (a) flood predictions and (b) methods for deriving hydrological extreme values for design. Flash floods are, unfortunately, rarely observed with conventional rain and discharge measurement networks (Borga et al., 2008), which implies that the sample for model learning and testing is small. While being rare, flash floods typically impact small catchments or even specific hillslopes, which are largely unobserved and have never experienced such events in the near observational past. We thus argue that the problem of flash flood predictions strongly relates to the classical "predictions in ungauged basins – PUB problem" (Hrachowitz et al., 2013; Sivapalan et al., 2003).

Despite their success, conceptual hydrological models have clear limitations in the context of local flash floods and Hortonian overland flow formation (Fatichi et al., 2016; Hrachowitz and Clark, 2017). Infiltration excess or Hortonian overland flow occurs when rainfall intensity exceeds the soil infiltration rate and is quite difficult to capture and characterize. This is because infiltration is a highly discontinuous and spatially variable process. Previous studies have reported a possible elevated infiltration capacity due to the emergence of macropore flow (Bronstert et al., 2023), which allows for rapid, purely gravity-driven infiltration, bypassing the slow soil matrix. An increasing macropore density can thus drastically reduce Hortonian overland flow generation (Zehe and Blöschl, 2004). Rain splash may, the other way around, lead to aggregate breakdown and siltation of the surface and significantly reduce infiltration capacity during high-intensity rainfall. This may drastically enhance Hortonian overland flow production, especially on agricultural fields during stages of low canopy coverage (Niehoff et al., 2002; Villinger et al., 2022).

Here, we hypothesize that gradient-resolving, physically based hydrological models (Fatichi et al., 2016; Paniconi and Putti, 2015) are better suited to address the challenges of flash flood predictions at small scales. By solving coupled partial differential equations that represent infiltration, soil moisture dynamics, overland flow and streamflow hydraulics, and evaporation, such models offer better opportunities for capturing spatially distributed production and concentration of overland flow (Pérez et al., 2011; Steinbrich et al., 2016; Zehe et al., 2001).

The primary challenges in employing physically based models for flash flood predictions in largely unobserved areas lie in their demand for extensive input data and their computational expense. To address the latter, we favor the use of the recently proposed representative hillslope concept by Loritz et al. (2017), which has not yet been tested for uncalibrated predictions and extreme flood events. To overcome

the data requirement challenge, we propose to leverage existing information from well-studied experimental catchments within the same hydrological landscape. Specifically, we suggest utilizing these well-instrumented catchments as 'donors' to transfer data on soil hydraulic properties and surface roughness to the poorly instrumented target catchments (donor models) for making a priori (uncalibrated) predictions. This could then give way for a posteriori (manual calibration) exercise once more relevant event information becomes available. It's also important to understand the significance of such a parameter transfer in comparison to estimating soil hydraulic functions using pedotransfer functions (Szabó et al., 2021; Zhang and Schaap, 2017) and readily available soil texture data.

Our main aim is thus to evaluate the potential of physically based hydrological models for predicting flash floods triggered by convective rainstorms in data-scarce regions. In this context, we put special emphasis on interactions between overland flow velocities, controlled by surface roughness, and Hortonian overland flow generation, controlled by the initial soil water content and infiltration characteristics, including macropores and surface sealing. To address the related challenges, we thus explore whether:

1. It is feasible to transfer model parameters from the well monitored experimental catchment to the data-scarce catchments for a priori uncalibrated flash flood predictions in response to convective storms, and how does this compare to simulations with parameters derived from soil pedotransfer functions?

2. Climate reanalysis data can be used to initialize process-based hydrological models in these data-scarce regions, and how much uncertainty is reduced using such a dynamic spin up compared to a random estimate of the antecedent soil conditions?

3. Reservoirs in headwater catchments (< 5 km^2) and their inverted mass balance offer largely untapped potential to study flash floods, complementary to a rather low number of operational stream gauges that are available at this scale worldwide?

2.2 VENUE AND MODEL

2.2.1 *Study Area*

The four target headwater catchments "W22," "W32," "W39" and "W44" belong to the Elsenz-Schwarzbach catchment in the State of Baden Württemberg, Southern Germany (Fig. 2.1 & Fig. A.2 in Appendix A). The catchments are located within the eastern "Kraichgau," west of Bad Rappenau and around 50 km from the nearest cities - Heidelberg and Karlsruhe. Due to a series of catastrophic flooding

episodes in 1993–1994, a comprehensive flood protection concept for the entire region was envisaged, which led to the development of local flood retention basins throughout the catchment area. The size of the catchments varies from 1 to 6 km^2; they all drain into the Krebsbach, which joins into the Schwarzbach near Waibstadt (the nearest gauging station—Eschelbronn Schwarzbach, being more than 12 km from our study area). The Elsenz-Schwarzbach finally merges into the Neckar, one of the Rhine's largest tributaries. From Figure 1, it is clear that even though the catchments are primarily agricultural in nature (major crops being—cereals, maize, sugar beets and potatoes), they are situated upstream of the population centers of the region. While these catchments are in the same hydrological landscape as the previously monitored Weiherbach experimental headwater shed (Schierholz et al., 2000), they are, despite the available water level gauges in the reservoirs, completely unmonitored with respect to rainfall, streamflow, soil moisture and soil hydraulic properties.

Figure 2.1: Overview of the location of the four headwater catchments considered in the study. Also, shown in the figure are the downstream flood control reservoirs, which afford protection to the towns in the region. The overlay layer depicts a Sentinel-2 (Drusch et al., 2012) multispectral true color composite image for the region during May–June 2016. Figure A.2 in Appendix A shows the stream network of Krebsbach and Schwarzbach and also the total accumulated precipitation during the event.

During the end of May to early June 2016, several strong convective rainfall events clustered in Germany because of persistent atmospheric conditions (Bronstert et al., 2018; Meyer et al., 2022; Piper et al., 2016). Rain totals exceeding 100 mm were reported in a day (in some cases

even within 2 hours), triggering flash floods in many small catchments over Southern Germany. The impacts in the Elsenz–Schwarzbach were also severe, with several of the flood control reservoirs being overtopped. To investigate the feasibility of our approach (Figure 2.2) and of the CAFLOW model to simulate such events, we focus our attention on the severe event of 08 June 2016 in the region (Appendix A.2). Since no streamflow gauges are available for the four headwater catchments, we use the water level measurements in the flood control reservoirs to estimate the runoff response based on the reconstructed reservoir inflow (W22, W32, W39, and W44). The storm runoff response is calculated by inverting the reservoir mass balance with knowledge of the reservoir geometry and stage–outflow relationship (Appendix A.3). Related uncertainties are accounted for by using a relative percentage error value (5%) in the stage level measurements.

2.2.2 *CATFLOW in a Nutshell*

The quest for accurately identifying and modeling the governing processes of the water balance and reactive pesticide transport in rural catchments motivated both the setup of the Weiherbach experimental catchment in the 90s (Plate and Zehe, 2008; Zehe et al., 2001) and the development of the process-based model CATFLOW (Zehe et al., 2001). The model relies on the subdivision of a catchment into several 2D hillslopes and an interconnected drainage (optional) and river network. However, each hillslope is modeled separately and hence this provides the opportunity to run each hillslope individually without the associated stream network. Hillslopes are discretized along a two-dimensional cross section using terrain-following curvilinear orthogonal coordinates. Soil water dynamics within the hillslopes are characterized using the potential-based form of the 2D Darcy–Richards equation, solved by a mass-conservative Picard solver using adaptive time stepping (Celia et al., 1990). Soil hydraulic properties can be parameterized according to van Genuchten (1980) and Mualem (1976), Tang and Skaggs (1977) or the recently proposed PDI model (Peters et al., 2021).

Overland flow is simulated using the diffusion wave approximation of the Saint-Venant equation and explicit upstreaming, in combination with the Gauckler–Manning–Strickler formula. CATFLOW can account for partially saturated and fully saturated subsurface flow as it solves the Darcy–Richards equation in potential-based form. The study of Loritz et al. (2017) explains how this, together with no-flow boundary conditions in the lower parts of the hillslope, can be used to create a riparian zone with a local active groundwater body to simulate baseflow. Evaporation and transpiration are usually simulated using a SVAT (Surface Vegetation–Atmosphere Transfer) module based on

the Penman–Monteith equation, accounting for annual cycles of plant phenology, albedo, and roughness using tabulated data. Stomatal conductance is characterized after Jarvis (1976), or via the inversion of sap flow data (Loritz et al., 2022).

The model can represent macropores either as stochastically generated spatially connected pathways of very low flow resistance (similar to fractures) or via an effective macroporosity factor, which scales the ratio between infiltration into the macropore and the matrix domain (Loritz et al., 2017; Wienhöfer and Zehe, 2014). For the simplified macroporosity factor approach, as detailed in Zehe et al. (2001), he bulk hydraulic conductivity k_s^B is linearly increased by a relative scaling factor f_m (if the relative saturation (S) at a grid point exceeds a certain threshold (S_o)) as follows:

$$k_s^B = \begin{cases} k_s + k_s f_m \dfrac{(S - S_o)}{(1 - S_o)}, & \text{if } S_o \leq S < 1, \\ k_s, & \text{otherwise.} \end{cases} \tag{2.1}$$

2.3 METHODOLOGY

Setting up a process-based model of any hydrological system mainly requires two types of information (Remson et al., 1971). The first concerns the fundamental laws governing the dynamics of system state variables and fluxes and related process parametrizations (e.g., preferential macropore flow). Our selection of the appropriate model implicitly selects these conditions for the given problem.

The second involves the data representing the "landscape" in the equation set. A proper identification of these properties is crucial for reliable model performance, and they can be divided into (a) system geometry (topographic and river network properties), (b) system parameters (soil and vegetation characteristics), and (c) initial and boundary conditions (antecedent soil moisture, rainfall and meteorological forcings). The current section details the steps required for setting up the model in this respect. We first explain the concept of the representative hillslope and its derivation from digital topographical data, then elaborate on the transfer of soil and land use parameters from the Weiherbach, along with the derivation of parameters via pedotransfer functions. Finally, we explain the spin up of the model using ERA5-Land and the radar-based precipitation product used during the event simulation.

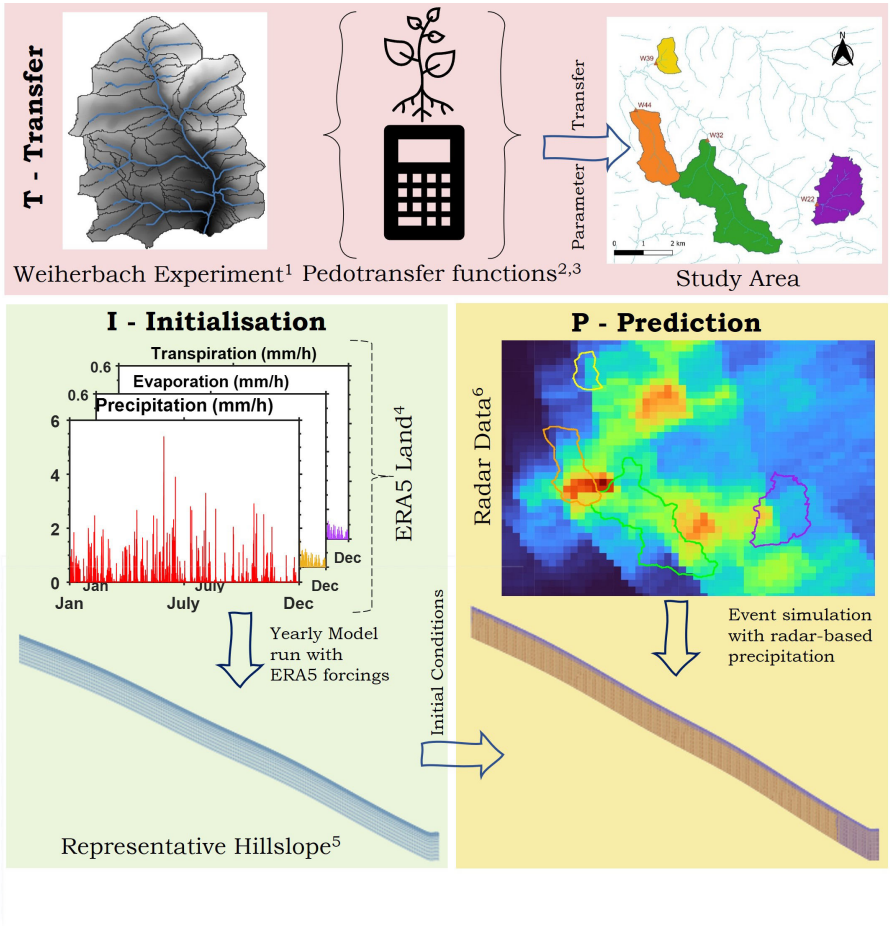

Figure 2.2: Illustration of our methodological approach–Transfer (T), Initialization (I), and Prediction (P). In Transfer (T), we transfer our knowledge of hillslope properties and soil parameters from the Weiherbach to our study area in Krebsbach. The Initialization (I) phase involves deriving the representative hillslope (detailed in Fig. 2.3) for the catchments and using the ERA5-Land forcings (hourly and $0.1^\circ \times 0.1^\circ$) to run the hillslope model for an entire year. In the prediction phase (P), the same model is run with the fine-resolution radar forcing and initial conditions from Initialization (I) for predicting the flash flood discharge (1—Zehe et al. (2001), 2—Zhang and Schaap (2017), 3—Szabó et al. (2021), 4—Muñoz-Sabater et al. (2021), 5—Figure 2.3 & Loritz et al. (2017), 6—Kachelmannwetter (2023)).

2.3.1 *The Representative Hillslope Concept*

Physically based hydrological models are renowned for their substantial computational demands, often impeding their broader application (Paniconi and Putti, 2015). As a result, catchment hydrology research has pivoted toward simplifying these models, ensuring they retain their physical underpinnings (Güntner and Bronstert, 2004). Notable models that exemplify this approach include the hillslope storage

Boussinesq model by Troch et al. (2003) and the representative elementary watershed model proposed by Reggiani and Schellekens (2003). In this study, we adopt a gradient- based simplification termed "representative hillslopes," as introduced by Loritz et al. (2017). Their work demonstrated that the water balance and streamflow generation in the Colpach catchment (19 km^2) could be accurately simulated using a single representative hillslope, negating the need for an associated river network.

Here we provide a concise explanation of why this approach works. The concept behind a representative hillslope is that both surface and subsurface water fluxes are propelled by differences in potential energy (Loritz et al., 2017; Zehe et al., 2013). These differences emerge from rainfall distribution over varied topography. In the context of intense convective rainstorms, our focus narrows to the energy balance of overland flow. Here, the driving potential energy difference hinges on the relative elevation between a location and its corresponding flow outlet. It is crucial to recognize that only a minute portion of this potential energy is converted into overland flow kinetic energy, with the majority being dissipated, primarily influenced by factors such as Manning's roughness (Schroers et al., 2022).

Preserving this energy dynamic on average implies that the topography of the representative hillslope should be structured to maintain average topographic gradients along the flow path to the nearest drainage point. A viable method involves segmenting geopotential energy by proximity to the river and averaging within each segment. Specifically, we consider the distribution of flow profile lines shown in Figure 2.3 for catchment W22. For any distance class (also shown in Fig. A.1: Appendix A), the total flow potential is the sum of the potentials of all the cells within the class, which is proportional to the relative elevation difference of the cells. For the catena profile, we require a representative value for this class so that the total energy remains on average conserved. We use the Linear Average Representative Slope Profile concept from Francke et al. (2008) for the same. The method involves a weighting factor based on the relative occurrence of each cell in a flow path (characterized by the flow accumulation values). Therefore, the mean elevation value (h_i) for a class at distance i is

$$h_i = \frac{\sum_{j=1}^{n} h_j \sqrt{f_j}}{\sum_{j=1}^{n} \sqrt{f_j}}, \tag{2.2}$$

where h_j and f_j are the relative elevation and flow accumulation values for each cell in the class at distance i, and $j = 1, \ldots, n$, with n denoting the total number of cells in the class. The representative value for any other attribute (e.g., width) can be calculated analogously.

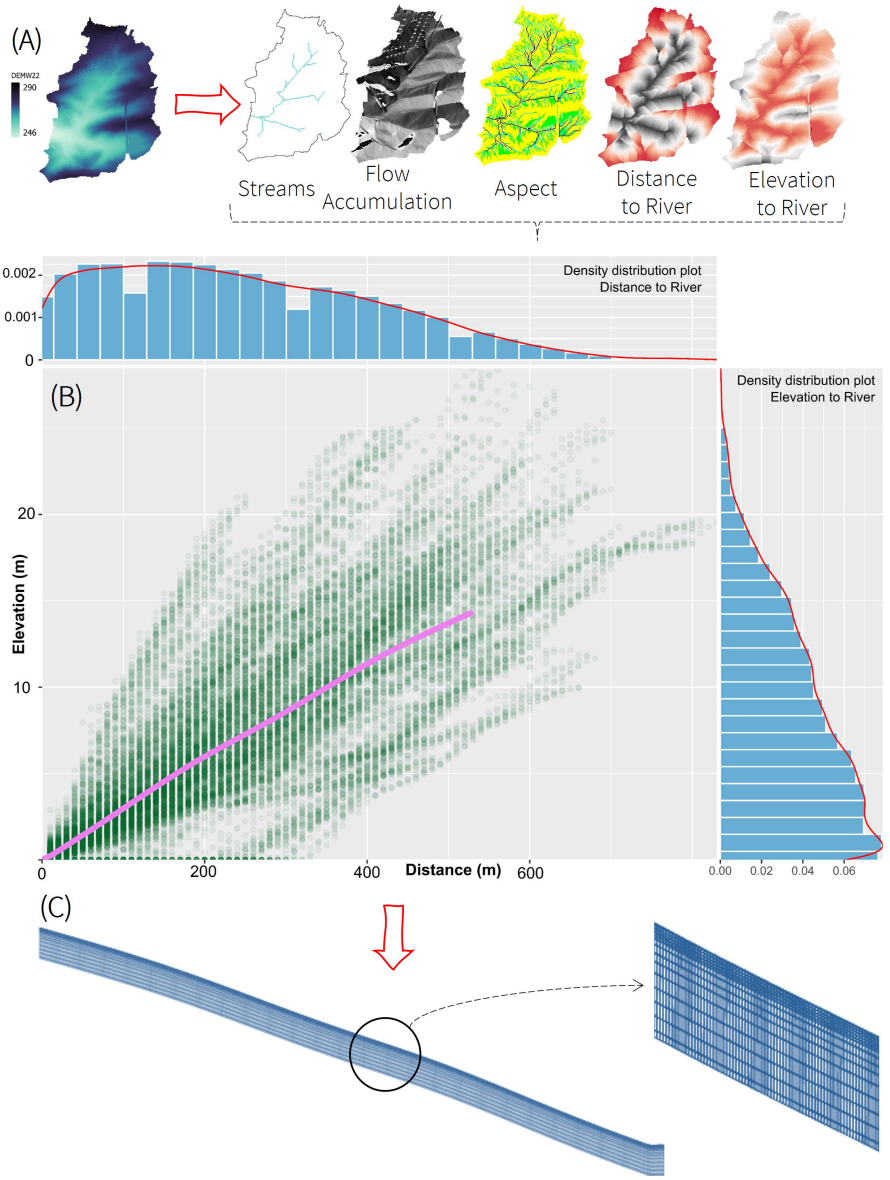

Figure 2.3: Workflow diagram illustrating the major steps involved in deriving the representative hillslope catena for the catchment W22. Derivation of raster maps (streams, flow accumulation, aspect, distance, elevation to river from the filled digital elevation model (a). Selection and binning of every distance and corresponding elevation to the nearest river segment(b). Calculation of mean distance using flow accumulation weights (see also Appendix A). Final derived representative hillslope (pink overlay line in panel (b)) in panel (c). The panel shows the grid mesh used to model the single representative hillslope. As detailed in Section 2.3.2, we opt for more finer grid spacing toward the surface as shown in the inlay figure.

2.3.2 *System Geometry (Deriving the Hillslope Profile)*

The representative hillslope topography is derived for each of the four catchments (Fig. 2.1) as illustrated for the catchment W22 in Figure 2.3. First, the digital elevation model (DEM) is pre processed to fill all depressions and sinks. We then derive the flow accumulation, aspect, and stream rasters from this filled DEM. The distance to the river and elevation to river rasters (which indicates the relative horizontal and vertical distance from a cell to the nearest river segment, respectively) are then extracted.

The distance from the nearest river segment and the corresponding relative elevation difference is plotted for all the cells within the catchment of interest (Fig. 2.3b). In Figure 2.3b, each green dot denotes a 10×10 m^2 cell in the catchment. The representative hillslope catena is then derived based on the methodology explained in Section 2.3.1. The potential energy conservation along the direction of the flow profile by means of the weighted mean elevation values is validated using four different distance classes (100, 200, 300 & 400 m: Fig. A.1 in Appendix A). The representative hillslope profile obtained for W22 is shown as a pink overlay in Figure 2.3b. The catena length is chosen intuitively based on the relative elevation and distance from stream distribution plots. In essence, the catena length depends on the distributions of distances and elevations to the river raster, as shown in Figure 2.3. We choose a suitable value for the catena length such that all major classes of elevation and distance to river cells are included (e.g., 530 for catchment W22 as shown in Fig. 2.3). Based on the elevations in each distance class and the chosen catena length, the representative hillslope (pink overlay line) geometry for the target catchment is calculated.

The entire catchment is then represented by this single hillslope, the corresponding patterns of soils (Table 2.1) and landuse in CATFLOW (Fig. 2.3), when simulating the catchment water balance and flash flood runoff. To this end, the hillslope W22 was discretized into 531, 1 m wide horizontal segments and 15 vertical layers. The total hillslope depth was set to 2 m, based on the transfer of knowledge from Weiherbach (see Section 2.3.3). The vertical grid resolution varied from 0.05 m near the surface to 0.25 m toward the bottom node (Fig. 2.3c). For ease of numerical simulation, we choose a uniform width (area of catchment/representative hillslope length) for all hillslope elements. Boundary conditions were set to the atmospheric boundary at the top, no flow boundary at the left margin. At the lower boundary a gravitational flow condition was established.

Table 2.1: Soil hydraulic parameters after van Genuchten (1980) and Mualem (1976) used to set up the soil properties in the CATFLOW model. The first two parameter sets were obtained from undisturbed soil cores in the Weiherbach/donor catchment. The remaining parameters were derived from two soil pedotransfer functions, Rosetta3 and EUPTFv2, using soil textural data (shown in Table A.3 - Appendix A)

Soil	k_s (m/s)	θ_s (%)	θ_r (%)	α (m^{-1})	n
Calcaric Regosol on Loess (Zehe et al., 2001)	3.1e-7	0.44	0.06	0.40	2.06
Coluvisol on Loess (Zehe et al., 2001)	1.7e-6	0.40	0.04	1.90	1.25
Rosetta3_LGRB	1.66e-6	0.45	0.08	0.43	1.54
EUPTFv2_LGRB	7.12e-8	0.44	0.006	0.35	1.21
Rosetta3_LUCAS	2.44e-6	0.44	0.08	0.38	1.56
EUPTFv2_LUCAS	1e-7	0.44	0.004	0.43	1.21

2.3.3 *Transfer of System Parameters From Weiherbach*

Since our present study area in the Elsenz-Schwarzbach consists mainly of agricultural loess catchments, which share similar geological and pedological characteristics and the same major crops with the Weiherbach/donor catchment (Zehe et al., 2001), we attempted to transfer the soil and land use parameters (particularly surface roughness) from the previous field experiments (Fig. 2.2). We transfer, furthermore, the finding that a successful simulation of the water balance and stream flow generation could be achieved, assuming that all hillslopes in the area share the same relative soil catena, (Zehe et al., 2001), consisting of Calcaric Regosol on Loess FAO/UNESCO, 1988; *Pararendzina*) in the upper 90% and Coluvisol on Loess FAO/UNESCO, 1988; *Kolluvium*) in the lower 10% of the hillslope.

Hence, we assigned the same relative catena to the representative hillslopes of the target catchments. In the donor model version, we also transferred the corresponding soil hydraulic functions for these soils aftervan Genuchten (1980) and Mualem (1976), which were determined by Schäfer (1999) and Delbrück (1997) for these soils in the Weiherbach using both field and laboratory experiments (Table 2.1 and also Table 3 in Zehe et al., 2001).

To investigate the value of this *a priori* information for predictions in our poorly gauged regions, we alternatively used pedotransfer functions to estimate the hydraulic parameters required for CATFLOW model from soil textural data. We employed two well known pedotransfer function - Rosetta3 (based on the UNSODA database: Zhang and Schaap, 2017) and EUPTFv2 (based primarily on European samples: Szabó et al., 2021). These functions require the relative proportions of sand, silt, and clay in the soil over our study

region (Table A.3 in Appendix A) as inputs. The corresponding van Genuchten Mualem parameters are also shown in Table 2.1. (For e.g. Rosetta3_LGRB refers to the parameter set obtained using Rosetta3 model and LGRB Textural data). Please note that the corresponding model versions used a single soil for the entire representative hillslope.

Estimates of surface roughness after the Manning Strickler coefficient, K_{st}, for different crop types and maturity stages were obtained from more than 60 rainfall simulation experiments conducted in the Weiherbach to characterize frictional losses for shallow overland flow (Scherer et al., 2012) (Throughout the remainder of this work, we use both Manning's roughness and Strickler values interchangeably to refer to the roughness coefficient (K_{st}) in the Gauckler Manning Strickler formula. Interested readers are referred to Hager (2015) for a historical anecdote). Previous studies (Lumbroso and Gaume, 2012) have shown that the traditional estimates of the Manning's coefficient obtained for riverine floods do not adequately represent flash flood conditions. Specifically, due to overbank flow during such extreme events, changes in the associated roughness properties are invariable. Hence, we use an ensemble approach for the surface roughness. Instead of running the model for a single value of Manning's roughness, we run the simulations for the range of roughness values within the informed experiments in the Weiherbach (6–12 $m^{1/3}/s$) and report the mean and spread of the ensemble predictions. This is of particular interest as it sheds light on the sensitivity of flash flood timing, peak and volume to changes in overland flow velocities.

As stated, CATFLOW also includes an advanced evapotranspiration subroutine, which enables time continuous simulations for a model spin up. However, the use of this module requires detailed information about the relative fraction of each crop, which is not available for the summer of 2016, as well as ground based data on radiation, wind speed, air humidity and temperature, which are neither at hand for our study area nor for most regions in the world. Hence, we decided to circumvent this challenge because the model offers the option to use evaporation data as input as well. We thus ran the model using globally available climate data sets for the model spin up. Specifically, we drove the hillslope models with the climate reanalysis product ERA5 Land (Muñoz-Sabater et al., 2021), using precipitation and evapotranspiration during the event simulation and for model spin up as detailed in the next section.

2.3.4 *Initial and Boundary Conditions*

The problem of inferring the initial conditions is a key challenge in event based modeling (Beauchamp et al., 2013; Zeimetz et al.,

2018), because simulation results are highly sensitive to changes in antecedent wetness (Zehe et al., 2005; Zehe and Blöschl, 2004). The challenge is usually not estimating the "actual" soil moisture state but establishing an initial state that is coherent with the land atmosphere interactions and the model physics (Koster et al., 2009). In essence, we seek an initialization identical to the dynamics being captured by our model.

In the present work, we use the ERA5 Land hourly precipitation and evapotranspiration reanalysis data for such a dynamic initialization of our representative hillslope model (Fig. 2.2) using a spin up period of 1 year. The model was run using the mean catchment values of forcing data from ERA5 Land until the event of interest (08.06.2016 00:00 UTC); the corresponding soil moisture patterns were saved and then used as initial conditions for a-priori event simulations with a radar based precipitation estimate (temporal resolution of 5 min).

During the event of 08 June 2016 (Appendix A.2), there were no operational rainfall gauges in the target catchments. The nearest gauge operated by the Baden Württemberg State Institute for the Environment, Survey and Nature Conservation (Landesanstalt für Umwelt, Messungen und Naturschutz Baden Württemberg, LUBW) lay toward the southeast of catchment W22 in Bad Rappenau Bonfeld (LUBW Station ID 76730: Fig. A.2 in Appendix A.2). The gauge recorded a total precipitation sum of around 28 mm on 08 June. The German Weather Services (Deutscher Wetterdienst, DWD) operates a nearby gauge in Waibstadt (DWD Station ID 13674), west of catchment W44. The DWD gauge reported a total daily precipitation of 11 mm. Considering the mismatch between the two gauges (Fig. A.4 in Appendix A), and the need for a finer spatiotemporal estimate of the convective storm activity, we opt for a radar product (temporal resolution 5 min, spatial resolution 221 m $\times$ 221 m) provided by Kachelmannwetter (Kachelmannwetter, 2023) as the forcing boundary condition for the model. The Kachelmann product merges DWD radar data with official and private rain gauges, including clutter correction, and so on. This product is also used by the federal flood forecasting center in the state of Baden Württemberg. Appendix A.2 depicts radar images of the storm on 08.06.2016 over our study region. Overall, it can be seen that the storm activity is captured quite well by the fine resolution radar product. The direction of the storm explains the smaller magnitude of total precipitation reported by the DWD gauge compared to the LUBW gauge (which seems to be nearer to the storm center: Fig. A.2 and Fig. A.4 in Appendix A.2). Note that the total sum registered by the DWD gauge corresponds well to the accumulated precipitation from ERA5 Land, while the radar product indicates almost 50 mm of rainfall for catchment W22.

2.4 RESULTS

In the following section, we first show the initialization using ERA5-Land and evaluate the simulated hillslope scale soil moisture compared to globally available soil moisture products (Section 2.4.1). We then detail the a priori event based flash flood simulations with the donor model versions using radar based precipitation forcing in Section 2.4.2. In Section 2.4.3, we present an *a posteriori* effort to improve the model in two out of the four catchments by including spatially variable precipitation forcings. Lastly (Section 2.4.4), we explore the sensitivity of the simulated flash floods to varying infiltration capacities due to macropore/surface sealing, soil hydraulic properties inferred from pedo-transfer functions and variations in antecedent soil water content compared to the dynamic initialization.

2.4.1 *Model Initialization With ERA5 Land*

Figure 4 shows the top soil (0–5 cm) water content simulated with the representative hillslope that was forced by ERA5 Land precipitation and evapotranspiration for catchment W22. Note that the variations of Manning Strickler, K_{st} $(\mathrm{m}^{1/3}\,\mathrm{s}^{-1})$ lead to a variation in soil water content during the summer period. To characterize the coherence of these soil moisture simulations with the gridded ERA5 reanalysis product, we calculated the Kling Gupta Efficiency (KGE; Gupta et al., 2009) between the averaged CATFLOW top layer soil moisture ensemble with the spatially averaged ERA5 Land surface soil moisture (0–7 cm) (Fig. 2.4 and Table 2.2)

Table 2.2: Goodness of fit measures between the modeled soil moisture values of different CATFLOW runs (varying Manning–Strickler coefficient K_{st}) and ERA5 Land surface soil moisture for catchment W22.

K_{st} $(\mathrm{m}^{1/3}/\mathrm{s})$	KGE	r	*Gamma*	*Beta*
6	0.651	0.798	0.812	0.786
7	0.654	0.797	0.821	0.784
8	0.662	0.800	0.836	0.782
9	0.691	0.822	1.050	0.752
10	0.690	0.822	1.051	0.751
11	0.671	0.800	0.861	0.779
12	0.699	0.824	1.019	0.756

While this revealed high KGE values, CATFLOW simulations were consistently drier than the ERA5 Land reanalysis product and the yearly CATFLOW runs (Fig. 2.4). This mismatch likely reflects the different soil parameterizations, vertical grid depths and scale disparities in the two models. It is important to note that we do not expect a perfect fit between the two modeled soil moisture products, our

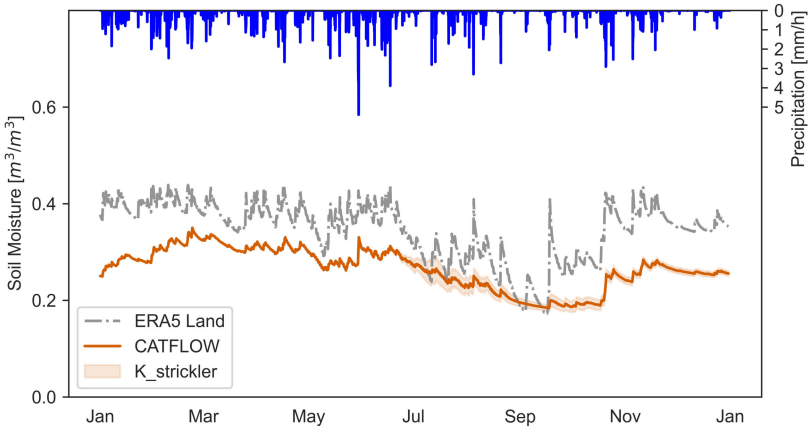

Figure 2.4: Time series of ERA5 Land surface soil moisture (0–7 cm) averaged over the entire catchment (gray). The precipitation data from ERA5 Land is shown in blue, while the soil moisture simulations with CATFLOW (0–5 cm) are depicted in red, representing the ensemble mean. The shaded regions illustrate the uncertainties ($\pm$ the standard deviation) corresponding to different values of the Strickler coefficient (K_{st} = 6–12).

interest is in capturing the overall local dynamics in soil moisture changes, as explained below.

In order to gain a more comprehensive understanding of the impact of bias on the overall Kling–Gupta efficiency (KGE) calculations, we calculated the three main components of the modified KGE (Kling et al., 2012). These included the Pearson correlation (r), the bias ratio (*beta*), and the variability ratio (*gamma*), as outlined in Table 2.2. As expected, we obtained high pearson correlation values (around 0.80) for all the different runs (varying K_{st} values). The high correlation shows that our approach reproduces the yearly dynamics of soil moisture changes in the region well (using the coarse resolution globally available ERA5 Land data as a benchmark). The values of *beta* and *gamma* indicate the overall bias and variability of the modeled values compared to the ERA5 Land data.

As we used ERA5 Land forcing variables (precipitation and evapotranspiration) to run the CATFLOW model, and then again ERA5 Land soil moisture states to evaluate the model performance, it remains to be seen whether the correlation is not only inherited from the forcing product. While a comparison with direct soil moisture observations would be preferable, unfortunately, these are not available for the catchments considered in this work, which are essentially ungauged except for the water levels in the reservoir. Due to the absence of in-situ soil observations, we use the following reanalysis and remote

sensing soil moisture products (Fig. A.8 in Appendix A) for comparison: (a) GLEAM (Global Land Evaporation Amsterdam Model; (Miralles et al., 2011)), (b) ERA5 Land (Muñoz-Sabater et al., 2021), (c) GLDAS (NASA Global Land Data Assimilation System, GLDAS 2.2 GRACE DA; (Li et al., 2019)), (d) MERRA (Modern Era Retrospective analysis for Research and Applications, version 2, tavg1_2d_lnd_Nx; (Global Modeling and Assimilation Office, 2015)), and (e) SMAP (Soil Moisture Active Passive L Band radiometer Level 3 product; (Chan et al., 2018)). The pairwise correlation for the different products for catchment W22 is shown in Figure 2.5.

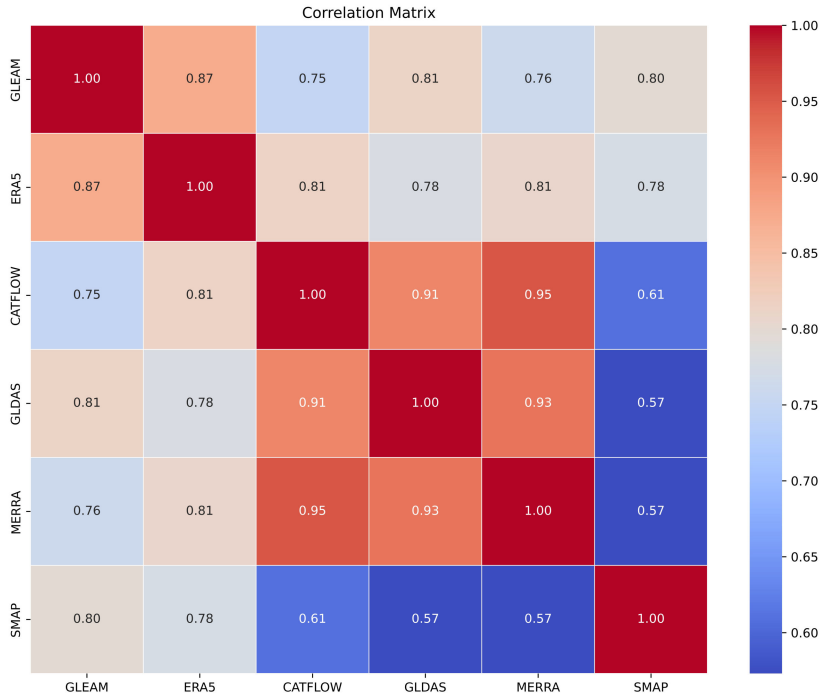

Figure 2.5: Correlation matrix plot illustrating the pairwise correlations between the different soil moisture products.

We stress here that we don't expect a direct match of the simulated time series in CATFLOW and other products due to varying spatial resolution, vertical grid sizes and hydraulic parameters. However, this is not necessary for initialization purposes. What we are aiming for is a strong temporal rank correlation, which means that the non-exceedance probabilities match well. The high correlation values obtained between CATFLOW and the different products (0.75 – GLEAM, 0.81 – ERA5, 0.91 – GLDAS, 0.95 – MERRA and 0.61 – SMAP) point to the feasibility of our approach to robustly estimate initial conditions using such reanalysis products in the absence of in situ observations. The multimodal comparison exercise also indicates that the choice of the reanalysis product (ERA5 Land) is not important. This is a valuable takeaway for setting up the model in a forecast mode. The initialisation could be done with a forecast product (compared

to reanalysis products, these are available at shorter lead times), and then radar forcing forecasts could be employed to provide event-based predictions.

2.4.2 *A Priori Flash Flood Modeling Using Representative Hillslopes*

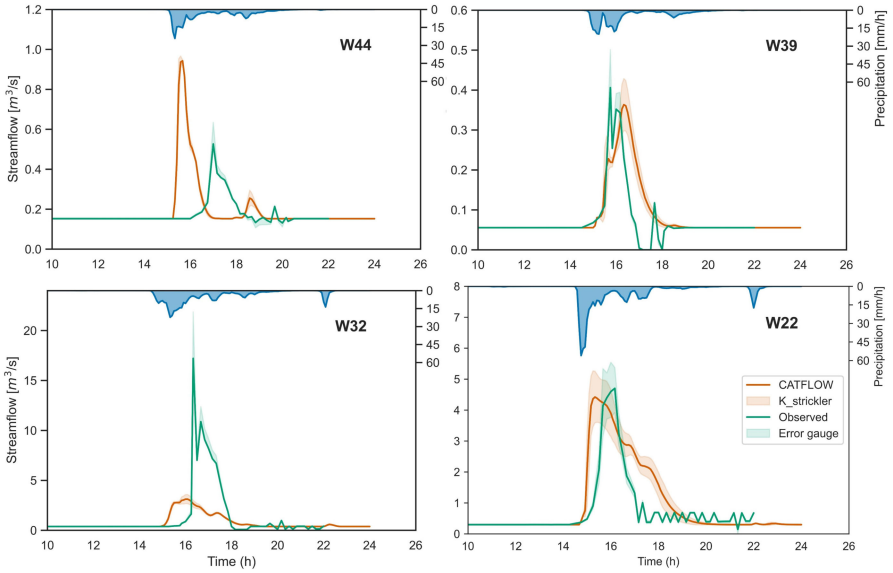

Figure 2.6: Rainfall—Runoff hydrographs for the flash flood event on 08.06.2016 at the four headwater catchments (W22, W32, W39, and W44). Green curve indicates the reconstructed inflow to the flood defense reservoir (Appendix A.3), assuming measurement uncertainties of 5 %. Red curve indicates the mean values (±SD) of the predicted flood discharge by the CATFLOW model ensemble (varying Strickler coefficient values). All simulation times are in UTC time zone.

The "donor" representative hillslope models were then used to simulate the runoff response for the convective storm event on 08.06.2016 in the four catchments in the study area using the dynamical initial conditions obtained from the yearly scale runs using ERA5 Land. This approach of initializing the models using the reanalysis data sets helps in avoiding a random guess of the initial states, which leads to considerable uncertainty, as detailed in Section 2.4.4. Figure 6 displays the simulated catchment response modeled using a spatially uniform precipitation input—the spatially averaged radar precipitation over each catchment. The simulated runoff in the four catchments is evaluated against the reconstructed inflow hydrograph obtained from the reservoir mass balance (Appendix A.3), assuming relative measurement error measures for peak flow, volume, and time to peak, as given in Table 2.3.

The donor model captures the steep ascent of the rising limb of

the flood hydrograph, albeit with a time lag, and matches the magnitude of peak discharge values in at least two out of the four catchments (W22 and W39). Note that the uncertainties in the simulated response due to variable surface roughness are almost identical to the possible observational errors in the gauge level measurements (5%) that propagated into the estimated storm hydrograph.

The peak flow errors (Table 2.3) in W22 (6%) and W39 (11%) are remarkably small, given the fact that these are uncalibrated predictions and the high uncertainties involved in local flash flood predictions (Bronstert et al., 2018; Marchi et al., 2010). It is also interesting to note that the simulations slightly underestimate the peak flow but overestimate the flow volume for both the catchments. Note that the simulated peak is delayed in W39, while it occurs earlier in W22.

In the two remaining catchments, the performance is clearly worse. In catchment W32, the model severely underestimates the storm response, while in W44, it slightly overpredicts the discharge values. To better understand the underlying reasons for this mismatch, we closely examined the storm pattern compared to the relative shape, LULC and orientation of the catchments.

Table 2.3: Characteristics of simulated and reconstructed storm hydrographs. Note - The error values are calculated between the mean values of the ensemble CATFLOW predictions and the inverted flood hydrograph for each catchment. Area of each catchment is indicated in brackets.

Flood characteristics	W22 (2.91 km^2)		W32 (5.6 km^2)		W39 (0.73 km^2)		W44 (2.44 km^2)	
	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim
Storm Precipitation (mm)	49.0	–	35.0	–	26.0	–	24.0	–
Peak Discharge, Q (m^3/s)	4.70	4.42	17.21	3.12	0.41	0.36	0.53	0.94
Time of Peak, t (s)	58200	55200	58800	57900	56700	58800	61200	56400
Flood Volume, V (m^3)	45637	57978	72209	49868	5189	5959	13466	14537
Flood Volume, V (mm)	15.7	19.9	12.9	8.9	7.1	8.2	5.5	5.9
Runoff Coefficient, R	0.32	0.41	0.37	0.25	0.27	0.31	0.23	0.25
Percentage Error in Peak Discharge, P_Q (%)	–	6.0	–	82.0	–	11.0	–	-79.0
Error in Time to Peak, P_t (s)	–	3000	–	900	–	-2100	–	4800
Percentage Error in Flood Volume, P_V (%)	–	-27.0	–	31.0	–	-15.0	–	-8.0

2.4.3 *Role of LULC, Distributed Rainfall Forcing and a Posteriori Model Updating*

From Figure 1, we can observe that the catchments W32 and W44 appear to be more elongated and fan-shaped in contrast to the broader shaped catchments W22 and W39. A look at the storm's direction (Figure A.3 in Appendix A.2) suggests that our initial assumption of uniform precipitation across the representative hillslopes might be too simple for these elongated catchments (W32 and W44).

On the other hand, the sharp discharge increase in W32 of around 15 m^3/s within 15 min seems astonishingly high, given the overall precipitation input and the responses in other catchments. One possible explanation could be an obstruction in the flow path perhaps due to debris like wood or sediment from the agricultural upstream areas of W32, which, as indicated by Figure A.2 in Appendix A.2, was closer to the storm center. This might have inadvertently created a temporary retention area, which then burst after a certain point, mimicking the effects of a dam break and resulting in a sudden inflow to the reservoir. Magnitude amplification due to such debris flows and driftwood blockages during flash floods have been reported in regions around the world (Chen et al., 2021; Schalko et al., 2018; Spreitzer et al., 2019).

The total event runoff coefficients calculated for each catchment (Table 2.3) also shows that while the approach slightly overestimates the response in all the other three catchments (W22, W39, and W44), it clearly underrepresents the runoff response by around 12% in Catchment W32. This apparently stronger runoff production could be explained by the presence of a larger fraction of impervious sealed built-up surface in W32. From Figure 2.1, it is seen that the small town of Haselbach lies within the catchment area, this contrasts with the other three catchments which are mostly only of agricultural or forest type. Note that this imposes limitations on the parameter transfer from the agricultural rural Weiherbach catchment. Another interesting point is that there is a well defined distribution of agricultural and forested areas along the stream profile in W32: crops at the upstream plateaus and forest along the tributaries or near the outlet. These regions could hence behave like sub catchments having distinct concentration times. However, it is worthwhile to note that out of all the four catchments the timing of the peak is most accurately captured in W32, which also has relevant implications for flood warning systems.

To investigate whether a distributed forcing input could help in better characterization of response in these elongated catchments, we re-ran the simulations for catchments W32 and W44 using a variable

precipitation along the representative hillslopes. Intuitively, we divided the catchment as having two different precipitation forcings over the upslope and downslope regions, to better reflect the storm pattern over the region (Appendix A.2). While this didn't lead to major changes for catchment W32 (apart from a minor increase in the peak flood), this lead to a clear improvement for W44 (Figure 2.7). The a posteriori predicted discharge values for catchment W44 match the observation remarkably better, as relative peak errors reduce from around 80% to just 2% (Table 2.4). The relative volume error decreases from 8% to 2% while the time to peak error remains nearly constant.

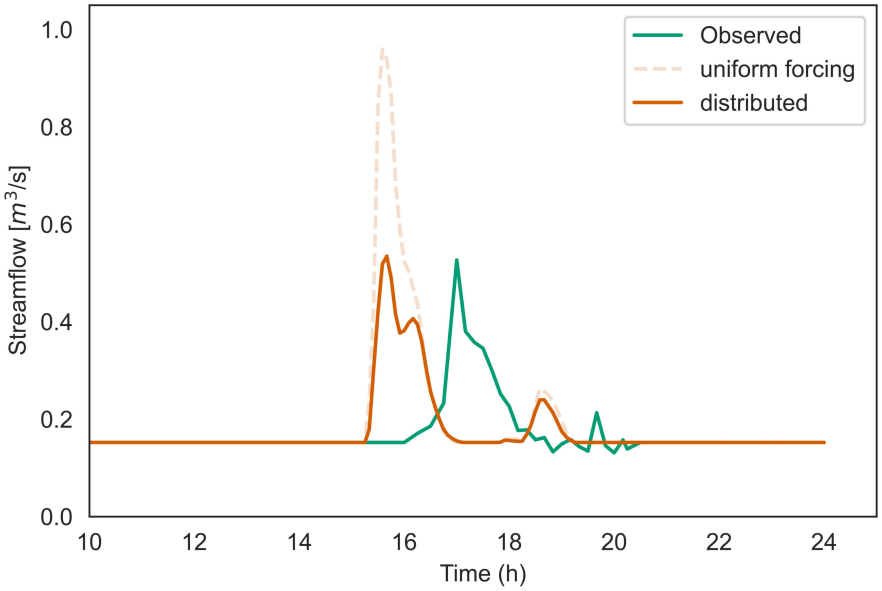

Figure 2.7: Flood hydrographs for catchment W44. Green curve indicates the reconstructed reservoir inflow, dotted red curve stands for model run using uniform precipitation forcing for the entire representative hillslope, solid red line denotes the model run with distributed forcing. All simulation times are in UTC time zone.

Table 2.4: Goodness of fit measures for the representative hillslope modeled discharges (K_{st} = 9) with the reservoir stage inverted streamflow measures for distributed forcings over catchment W44

Flood characteristics	W44 (2.44 km^2)		
	Observed	Uniform forcing	Distributed forcing
Peak Discharge, Q (m^3/s)	0.53	0.96	0.53
Time of Peak, t (s)	64800	56100	56400
Flood Volume, V (m^3)	13465.9	14567.4	13714.54
Runoff Coefficient, R	0.23	0.25	0.25
Percentage Error in Peak Discharge, P_Q (%)	–	-82.2	-1.5
Error in Time to Peak, P_t (s)	–	53.1	50
Percentage Error in Flood Volume, P_V (%)	–	-8.2	-1.8

2.4.4 *Sensitivity to Preferential Infiltration, Estimated Soil Hydraulic Functions and Initial Conditions*

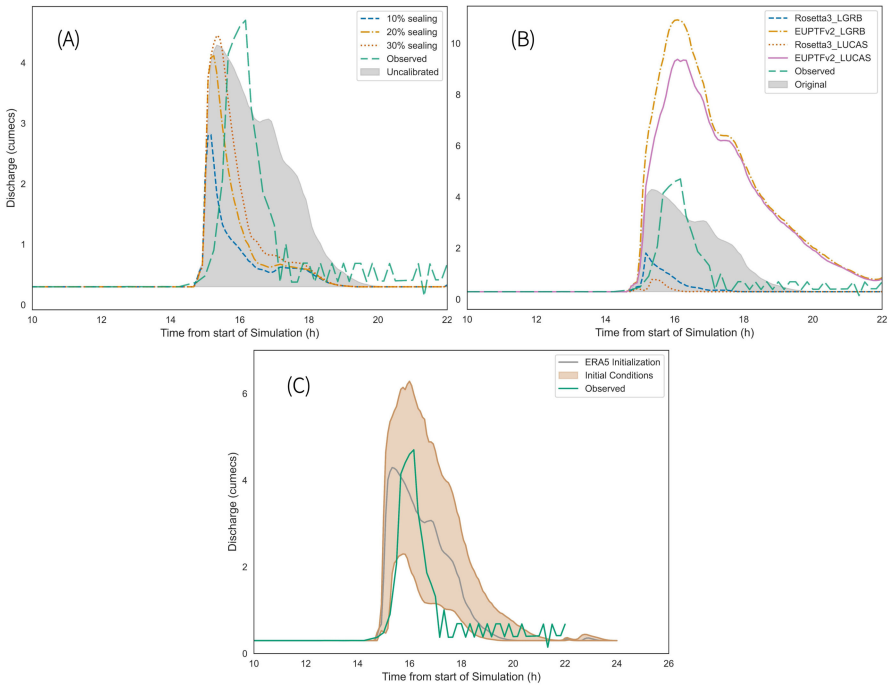

Figure 2.8: Sensitivity of the simulated flash floods (Catchment W22) to varying infiltration capacities due to macropore/surface sealing (a), soil hydraulic properties inferred from pedo-transfer functions (b) and variations in antecedent soil water content compared to the dynamic initialization (c).

In order to test the sensitivity of the simulated flood discharge to preferential flow processes in the catchment, we employed the simplified macroporosity approach described in Section 2.2.2 (Equation 2.1) for Catchment W22 (having a significant amount of agricultural coverage). Essentially, we simulated the increased infiltration due to macropore flow and reduced infiltration due to surface sealing by relative scaling of the bulk hydraulic conductivity along the representative hillslope element. We applied a scaling factor of 10, which is consistent with the results of previous studies (Niehoff et al., 2002). We then ran simulations changing the longitudinal fraction (10%–30%) of the hillslope where surface sealing and a reduction in infiltration occurred. The multiplicative scaling factor for saturated hydraulic conductivity was then applied to the remaining region (90%–70%). As a result, the top parts of the hillslope experience higher macropore flow, while the foot of the hillslope experiences stronger ponding and less infiltration.

The corresponding sensitivity is shown in Figure 2.8a. By increasing the extent of sealed surface from 10% to 30% we get a clearly better

matching to the observed hydrograph for catchment W22. The storm volume errors are considerably reduced (around 10%) compared to the uncalibrated simulation, which did not consider preferential flow. This is in line with the observation by Niehoff et al. (2002) for the Lein Catchment (also part of the Kraichgau region in South West Germany) of an overestimation of direct surface runoff when preferential flow processes are neglected during convective events. Villinger et al. (2022) tested the same idea when revisiting the flash flood of 1996 in the Weiherbach. Although the significance of such preferential flow processes is widely acknowledged in literature (Bronstert et al., 2023; Zehe et al., 2001), they are not yet incorporated in conventional flood forecasting methods.

The flood simulations for catchment W22 have been illustrated in Figure 2.8b using the derived parameters from pedotransfer functions (Table 2.1 & Table A.3). The results show that while the model can indeed be set up using such derived parameters, the relative error values for a priori uncalibrated predictions are significantly large. The simulations derived from Rosetta3 drastically underestimate the runoff generation, while those from EUPTFv2 strongly overestimate the storm response. This deviation is in line with the difference in the saturated hydraulic conductivity estimated by the two pedotransfer models (Table 2.1) and speaks clearly for Hortonian overland flow.

To understand the sensitivity of the flood simulations to variations in antecedent soil water content, we designed a virtual experiment in which the initial soil conditions were taken from a random perturbed sample (30%–90% relative saturation). These experimental runs show a considerable spread both in flood and volume values compared to the ERA5 Land Initialization (Fig. 2.8c). The scatterplot presented in Figure A.9 also shows a wide range of errors, highlighting the benefits of utilizing ERA5 Reanalysis for making a priori model predictions in regions with limited data availability.

2.5 DISCUSSION

In line with our main hypothesis, we demonstrated the feasibility of the physically based CATFLOW model for a priori flash flood predictions in data-scarce, small (< 6 km^2) headwater catchments in response to convective storm events. We used the representative hillslope concept and transferred soil hydraulic and roughness parameters using the well observed Weiherbach as donor catchment. Due to the absence of a coherent observational network in the area, we employed initial conditions from a spin up driven by climate reanalysis data. The key to evaluating the event-based simulations was to inversely estimate the storm inflow using water level data and reservoir geome-

tries. This comparison allowed us to quantify the relative simulation error values.

2.5.1 *Toward Short Term Predictability in Ungauged Basins*

The "Predictions in Ungauged Basin" problem (IAHS PUB Initiative: Hrachowitz et al., 2013) and the prediction of local flash floods has much in common, as the latter, unfortunately, usually occur within poorly instrumented areas. The PUB initiative attempted to solve the problem of hydrologic predictions in ungauged basins using the concept of regionalization, that is, to undertake a transfer of hydrological understanding from gauged to ungauged environments. Spatial proximity is, in this context, one of the most widely used and simple regionalization techniques.

While parameter transfer attempts are not uncommon, using specially trained regionalization functions for conceptual models (Hundecha and Bárdossy, 2004), we have shown that this works decently well with parameters that are directly measured without the need for posterior parameter tuning. We were able to demonstrate that parameters and entire model structures of a physically based model can be transferred (and in a sense generalized) from well observed "donor" catchments to largely unobserved regions. This is an important finding, particularly for setting up such process-based models in an uncalibrated setting. And it suggests that experimental catchments provide information that is useful beyond the single catchment case. In consequence, the relevance of hydrological observatories like the Weiherbach (Zehe et al., 2001), the HOAL (Blöschl et al., 2016) or the Attert experimental basin (Pfister and Kirchner, 2017) continues undebated.

It is important to have information about behavioral hillslope structures and soil hydraulic parameters before transferring them. This information can be obtained from experimental watersheds within the same hydrological landscape. However, such data may not be available in all areas, so we explored simulations (Fig. 2.8b) using soil pedotransfer functions based on standard soil maps. We compared these simulations to the original donor simulations and found significant errors in the pedotransfer simulations. This is not surprising as pedotransfer functions are based on a wide range of soil samples and may not be suitable for localized predictions without posterior parameter calibration. However, the fact that the model can indeed be set up using globally available soil texture data along with open source pedotransfer functions, implies that our approach is generally applicable globally. However, a posteriori calibration exercises may likely be needed to ensure that the model predicts Hortonian overland flow response in such new regions. However, this might not necessar-

ily require a large set of events, as previous studies have shown that physically based models can be trained on a few, informative events (Villinger et al., 2022; Zehe et al., 2005).

2.5.2 *Tackling Data Scarcity for Initialization*

One way to overcome the challenge of estimating initial conditions for event simulations is by using climate reanalysis products such as ERA5 Land. The importance of such antecedent soil moisture conditions in constraining the flood response cannot be overemphasized (Manoj J et al., 2022, 2023), as has been shown for many catchments across Europe (Berghuijs et al., 2019; Blöschl et al., 2019a; Marchi et al., 2010) and also here in Section 2.4.4. Global climate models have delivered commendable outcomes when it comes to capturing climate and weather extremes on regional scales. However, they remain largely unusable for estimating the impacts of smaller scale hydrological events (IPCC, 2021; Poschlod, 2022).

Our method of utilizing the ERA5 Land as a spin up for the CATFLOW model is beneficial in initialization over data-sparse, small headwater catchments. The strong correlation to the other soil moisture products indicate, that the choice of the product is not necessarily crucial. It is crucial to perform a dynamic spin up because a random variation in initial soil saturation caused a large uncertainty in the a-priori runoff simulations. Using the reanalysis product for deriving the event initial conditions significantly reduces the large uncertainty in the spread of peak flood and volume, as depicted in Figure 2.8c.

2.5.3 *Managing Observational Uncertainties*

Flash floods usually come as (bad) surprises, often impacting regions when and where we least expect them (Borga et al., 2008). Hence, strategies that provide robust warnings are essential. However, since they are also quite rare in nature, there lacks a coherent motivational starting point to invest time and resources into them (Montz and Gruntfest, 2002).

Marchi et al. (2010) analyzed around 25 major flash floods over Europe and showed that proper observational records didn't exist for more than half of the investigated events. During such intense flash floods, direct current meter measurements are often not possible due to safety and technical considerations. Furthermore, these events usually occur in remote ungauged regions with limited accessibility (Borga et al., 2008). It is important to stress here that even in cases with flow measurement gauges, prediction of discharge values during such convective events usually involves lot of uncertainties due to

faulty devices, dynamical riverbed changes and floating debris in the stream (Lumbroso and Gaume, 2012).

As is common in such poorly gauged catchments (Bronstert et al., 2018), we didn't have a streamflow gauge to compare our model performance. Hence, we made use of the reservoir geometry and downstream flood retention reservoirs to obtain a crude estimate of the storm characteristics. This strategy should create a win-win situation because local water resource managers are natural end users of such a warning system, and we tremendously increase the sample of historical test cases and complement the small sample that is available from the few gauging stations that observe catchments < 10 km^2. To investigate the plausibility of the reservoir inverted

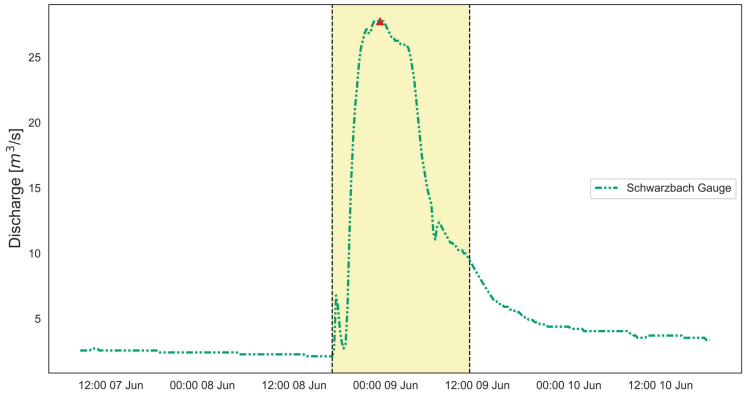

Figure 2.9: Discharge record at the nearest LUBW gauging station to the Krebsbach (Eschelbronn Schwarzbach –Gauge ID 461) denoting the instantaneous streamflow (temporal resolution—15 min) values during the event of 08–09 June 2016 (All times in MEZ).

estimates for the four headwater catchments of the Krebsbach, we compare those to streamflow record from the nearest downstream gauging station on the Schwarzbach River (Eschelbronn Schwarzbach, Figure A.2 in—Appendix A.2—LUBW Gauge ID 461). Figure 2.9 indicates that the streamflow increased quite abruptly from around 2 m^3/s in the evening of 8th June to 27 m^3/s (increase of around 1,250%) at midnight. The abrupt increase occurred within 6 hr, followed by the recession by the end of the next day, pointing out the flashy nature of the flood. It is also interesting to note that the peak flow of 27 m^3/s (Figure 2.9) matches quite well with the sums of individual flood values (W22—4.7 m^3/s, W32—17.2 m^3/s, W39—0.4 m^3/s, and W44—0.5 m^3/s) obtained for the four headwater catchments by inverting the flood reservoir mass balance. This indicates that the individual flood values obtained are well plausible.

While it is true that the LUBW stream gauge in Eschelbronn drains a much larger area (around 200 km^2), the precipitation values recorded by the DWD precipitation gauge (11 mm; Figure A.4) underpin, in line with the reports on the flash flood summer in 2016 (Göppert, 2018), that these floods were caused by localised convective events. Moreover, the timing of the peak at the gauging station is also generally in line with the observed timings in the four catchments (after accounting for the time required for the flood wave to reach the gauge along the river network from Krebsbach to Schwarzbach (Figure A.2) and possible time conversions between UTC and MEZ).

2.6 LIMITATIONS AND OUTLOOK

The perils of applying physically based models at large scales, while the governing equations were developed for much smaller spatial scales, are reviewed in literature (Hrachowitz and Clark, 2017). Such distributed, process-based approaches are also criticized for their complexity and larger data requirement compared to mathematically simpler conceptual models. While conceptual models are without doubt helpful, they are not so intuitive to be combined with a perceptional model of a catchment, as argued by Loritz et al. (2017). Top down (conceptual) models do not perform well in regions and scenarios which deviate from their well-calibrated range of conditions (Fatichi et al., 2016; Hrachowitz and Clark, 2017), and they have challenges in representing Hortonian overland flow.

We believe that, under the threat of a non-stationary climate (Milly et al., 2008) and unprecedented flow regime changes (Pérez Ciria et al., 2019), strategies which involve a convergence of both modeling philosophies are called for. The representative hillslope approach for flash flood modeling is a venture in this regard. And the related use of a gradient resolving, physically based model corroborated well known and provided surprising sensitivities to overland flow. The former is the importance of macropores and infiltration characteristics; the latter is the sensitivity of surface roughness to flood volumes. These findings suggest that the mass and momentum balance of overland flow during flash floods are strongly coupled and thus need coupled modeling, which, however, is not the case in standard flood forecast models.

However, limitations remain that need to be properly understood and accounted for. The 2D effective representative hillslope used to represent the catchments implies the assumption of symmetry where the runoff production is controlled by hillslope parallel and vertical fluxes and their driving gradients (Loritz et al., 2017). The derived effective catena profile depicts our best guess based on the

available topographical data (DEM). Furthermore, we used only one transferred (typical) soil catena for modeling. Any uncertainties and errors in the terrain representation will invariably propagate to our model geometry. Another point is the sensitivity to different DEM resolutions, raster filling and flow direction algorithms (Loritz et al., 2019). The hillslope scale again emerges as an interesting sub-unit within a catchment (*virtual laboratories*: Fatichi et al., 2016) for testing the impact of such uncertainties on runoff response.

So far, the direct flood simulations were essentially event based with no separate baseflow component (a constant baseflow was considered from start till end). Moreover, in our case, we do not attempt to fit the model response to water level curve obtained from the reservoir level measurements. Our main aim is to a priori simulate the catchment response during high intensity events in a simple, parsimonious manner. Since the overarching focus of this study was to test if the representative hillslope models and the parameter transfer could work in getting predictions right in rather poorly gauged catchments, we opted for an experimental design in which the approach was tested by spatial sub-sampling rather than temporal. The approach should indeed be extended to investigate other storm events in individual catchments. This provides for an interesting research avenue worth exploring in future works, particularly on the value of dynamic initialization in different years.

We also endeavored to consider the uncertainties in our modeled response (by varying the surface roughness) and the observational benchmark (relative error in gauge measurements). It is indeed true that the choice of process based model implies that we deal with a much larger number of system parameters and boundary conditions, compared to conceptual models. The strength is that these parameters are interpretable and, as shown in this study, transferable.

The forcings and soil moisture simulated by any land surface model (ERA5 Land, in our case) is model-dependent and direct transfer of one model product to another can lead to inconsistencies due to deviations in formulations (Koster et al., 2009). Attempting such a switch of forcing from a coarse gridded reanalysis product (ERA5 Land) to a fine resolution radar precipitation product would usually entail a re-engineering of the model and associated variables. However, we show here that a process based, spatially distributed model can capture the dynamics due to their mechanistic description of the flow system (conserving both the energy and mass balances). Moreover, as we show in Section 2.4.1, we are more interested in the temporal soil moisture variability rather than the absolute values predicted by the models. Hence, we expect very less model bias due to the choice of

the reanalysis product.

One argument frequently put against the use of process-based models in flash flood modeling and forecast strategies is their higher computational times. In the current attempt, we reiterate that by employing a representative approach which spatially averages along the main driving gradient of flow, we can preserve the total flow potential of the catchment without significant computational effort For reference, the spin up phase for the entire year had run time of less than 10 min while the event simulation for each catchment took around 2 min, in a normal Windows PC with 32 GB RAM only.

2.7 CONCLUSIONS

Based on the provided evidence, we conclude that physically based modeling in data-scarce, small headwaters using representative hillslopes, supplemented by climate reanalysis products, appears to be a viable pathway for achieving dependable rainfall-runoff simulations during high intensity storm events. By ensuring that these representative hillslopes align with the principles of energy conservation, we strike a balance between the intricacies required by physically based models and the desired simplicity rooted in parsimony considerations. Integrating with global climate reanalysis products effectively addresses the persistent challenges of data availability, a crucial aspect when modeling extreme events in data-limited regions globally. Our endeavor to model and understand flood dynamics in these specific regions, despite the data limitations, gave crucial insights which presents a step forward in mitigation and preparation for such extreme events. The findings indicate that the modeled hydrograph aligns well with the observed flood curve, derived from reservoir gauge level measurements, in three of the four studied catchments. We also looked at the role played by localized soil surface conditions and preferential flow during convective storms which are quite important for improving our understanding of small scale but high impact flash floods. While the approach demonstrated limitations in one of the region's larger catchments, further exploration and research, as outlined in the subsequent text, could provide more insights into modeling elongated catchments, especially those with urban developments.

Part III

LEVERAGING REPRESENTATIVE HILLSLOPES FOR ENHANCED FLOOD RISK MANAGEMENT IN MESOSCALE CATCHMENTS

This study is currently under review in the scientific journal *Hydrological Processes* (HP). The remainder of Part III is a reprint of:

Manoj J, A., Villinger, F., Wienhoefer, J., Loritz, R., and Zehe, E. (Under review). "Leveraging representative hillslopes for enhanced flood risk management in mesoscale catchments." Hydrological Processes. Preprint available at: https://doi.org/10.22541/au.176797104.47760019/v1

3

LEVERAGING REPRESENTATIVE HILLSLOPES FOR ENHANCED FLOOD RISK MANAGEMENT IN MESOSCALE CATCHMENTS

ABSTRACT

Understanding and preparing for extreme events in a warming climate is fraught with significant challenges, particularly when modelling (the ever increasing) flash floods in small- to mesoscale catchments. While top-down modelling approaches that attempt to describe fluxes directly at the scale of the system perform rather well for riverine floods caused by saturation-excess runoff generation, models that rely on bottom-up paradigms have been successful in handling intensity-controlled runoff generation, along with associated preferential flow processes. In this work, we develop a meso-catchment (196 km^2) scale, spatially distributed process-based model based on the gradient conserving approximation of representative hillslopes. The approach helps us balance the required complexity for modelling such coupled flows with the practicalities of scaling our simulation to the mesoscale catchment.

The model is used to investigate a severe flood that happened in 1994 in southwest Germany. After validation, the model is used to reconstruct flood magnitudes in the poorly gauged yet severely affected headwater regions of the catchment. The results reiterate the importance of accounting for differing gradients and landuse patterns across headwater regions and their role in runoff generation. To further investigate the implications of these findings for flood hazard assessment and to explore the reliability of design flood values in this area, we conducted additional simulations for a range of precipitation return periods, demonstrating that errors are often most pronounced at smaller spatial scales, largely due to data scarcity. Finally, given the growing interest in nature-based solutions (NbSs) as decentralized, process-oriented alternatives to conventional structural flood protection, we implemented representative NbSs at the hillslope and headwater scales and assessed their downstream effects at the catchment scale

Our work has broad implications for modelling and the appropriate characterisation of overland flow storm responses at such crucial intermediate scales, given the increasing frequency of convective extremes.

3.1 INTRODUCTION

Floods are usually defined as the inundation of an otherwise dry area or a significant streamflow event extending well beyond the regular river banks (Mishra et al., 2022). They usually garner significant public and governmental attention when a severe disruption to normal life occurs. Due to the increasing population and economic assets in flood-prone areas, exposure to floods is expected to grow by a factor of three by the end of 2050 (Jongman et al., 2012). Floods can result from various atmospheric processes (Merz et al., 2021), which are then modulated by local catchment and river network dynamics. Anthropogenic climate change and other human-induced modifications (Pérez Ciria et al., 2019) to the global water cycle, which may include land use changes, construction of settlements, and channel modifications, also leads to the intensification of floods.

Preventing floods in a warming climate is becoming increasingly challenging, in part explained by the Clausius-Clapeyron (Pall et al., 2007) relation, which states that the moisture-holding capacity of the atmosphere increases as air temperature rises. However, effective pre and post flood management can significantly mitigate the human and economic costs associated with such disasters. Global scale analyses (Kreibich et al., 2022) have shown that flood management measures (both structural and non-structural) have led to a pronounced reduction in flood-related casualties and damages. Flood risk management can be further improved by adopting better modelling strategies (Brunner et al., 2021) that explicitly account for the rarity of such events.

Various types of hydrological models have traditionally been employed for flood modeling (Hrachowitz and Clark, 2017). Simple, lumped conceptual hydrological models, such as the unit hydrograph, are often favored for simulating extreme flood events and designing flood protection measures. These approaches have demonstrated commendable performance in mesoscale catchments for riverine floods. However, limitations remain in applying these methods to diverse scenarios. Recent experiences with increasing flash floods (Meyer et al., 2022) have raised questions about the validity of the assumptions inherent in these idealized models. The tragic events of Ahr valley (Mohr et al., 2023) and Braunsbach (Bronstert et al., 2018) are stark reminders of the disastrous impacts caused by such extreme events. Unlike floods controlled by storage (saturation excess - Dunne (1978)) or driven by channel networks (Merz et al., 2021), flash floods typically result from Hortonian overland flow (Horton, 1933), which is usually driven by convective storm activity (Meyer et al., 2022) in small headwater catchments and influenced by localised preferential

flow processes, such as macropore flow.

It is important to emphasize that, even in the case of flash floods, the flow patterns and subsequent attenuation of flood waves in river channels (Fenton, 2019) are more thoroughly observed and characterised than the localised processes (Bronstert et al., 2023) occurring at the hillslope scale. These localised processes may include, but are not limited to, antecedent conditions, reinfiltration, surface sealing, and the activation of preferential macropore flow. These nonlinear processes interact to create significant and evolving hazards, whose impacts cannot be accurately predicted by generalised linear models (Kirkby and Cerdà, 2021). Conversely, as shown by recent advances in so-called nature-based solutions (NbSs - Guillaume et al., 2025, Richet et al., 2017, Rosier et al., 2024), localised, site-specific water retention measures could attenuate the flood at the cause scale (e.g., hedgerows at agricultural fields) rather than seeking measures to tackle the flood at the effect scale (e.g., flood reservoirs near the catchment outlet).

Spatially distributed, process-based modelling approaches (Fatichi et al., 2016) have much to offer to tackle the challenges associated with such intensity-controlled events. These models rely on correctly representing the gradients driving the flow, hence requiring far fewer events for model learning. Transfer learning implies that parameter values from catchments sharing similar characteristics could be used throughout the entire region. They also enable the virtual implementation of nature-based solutions at the hillslope scale, allowing for intricate details such as the spatial arrangement of these measures with respect to the dominant flow paths within the catchment.

One potential drawback of using such models is their high demand for data and computational complexity. This consideration becomes increasingly important in design, where rapid and cost-effective alternatives are usually preferred. Innovative catchment simplifications that maintain the system's total flow potential could effectively balance computational complexity with necessary process representation. One such approximation is the concept of representative hillslopes (Loritz et al., 2017) which tries to model only a fictitious representative hillslope profile derived from the topology of all existing hillslopes within a catchment. While the concept has shown promising results in smaller experimental catchments (Loritz et al., 2017; Villinger et al., 2022)) and recently to predict flash floods in poorly gauged headwater catchments (Manoj J et al., 2024). It remains to be seen how such simplifications impact process representation and model performance at the larger mesoscale.

Flood risk management should ideally also take care of 'surprises'

(Mishra et al., 2022). A growing concern is the failure of flood retention structures designed for a T-year return period when faced with a storm that has a return period shorter than T. Idealised design workflows often overlook responses to short-duration events that mainly trigger intensity-controlled (Hortonian) runoff generation. This issue is exacerbated in some cases due to the lack of records for past destructive flooding episodes, especially those that occurred in localised (often poorly gauged) headwater catchments.

At the hillslope scale, non-linearities and complexities due to changing landuse patterns significantly influence the shape of the flood hydrograph. However, flood management practices often focus on much larger scales, typically due to social and administrative constraints. Changes in land use, especially soil sealing, compaction, and the removal of hedgerows, have contributed to increased water loss through runoff in post-World War II Europe (Auerswald et al., 2025). On the other hand, if implemented properly, such changes have the potential to mitigate some of the adverse effects of global warming on terrestrial environments. Unfortunately, they are neither adequately represented in hydrological models nor considered a viable pathway to achieving sustainability goals. With the following challenges in mind, in this study, we explore:

1. Can representative hillslopes coupled with flow routing support physically consistent, transfer-learning-based prediction of convective flash floods at the mesoscale?

2. Whether floods can be reconstructed using the model to evaluate the impact of past flooding and to assess the adequacy of existing flood design values within the catchment?

3. Is it possible to provide protection (in comparison to classical flood reservoirs) against such flash floods at the larger mesoscale by using localised nature-based solutions on the hillslope scale?

3.2 VENUE AND MODELLING PHILOSOPHY

3.2.1 *Study Area*

The Elsenz Schwarzbach catchment (located in the federal state of Baden-Württemberg, Germany) was chosen as the study area (Fig. 3.1), as it was severely impacted during the catastrophic 1993–1994 (Disse and Engel, 2001; Villinger et al., 2022) flood series in Germany. The German federal state of Baden-Württemberg lies towards the country's southwest, sharing land borders with France and Switzerland. The two major climatic regimes according to the Köppen-Geiger classification (Beck et al., 2018) are temperate oceanic climate (Cfb)

and humid and warm continental climate (Dfb). The state is primarily drained by the subcatchments of the Rhine, with the rest by the Danube catchments (Ho and Ehret, 2025).

This medium sized mesoscale catchment (196 km^2; Fig. 3.1), while being largely hydrologically unobserved (except for streamflow) shares similar geological and pedological characteristic as the heavily experimented (but smaller: 3.5 km^2) Weiherbach catchment (Figure B.1 in Appendix B.1: Zehe et al., 2001). Both the catchments lie in the 'Kraichgau' region, mainly consisting of agricultural catchments with Loess soils. The Elsenz Schwarzbach drains into the Neckar river at Neckargemünd. The river Neckar then joins the Rhine river at Mannheim.

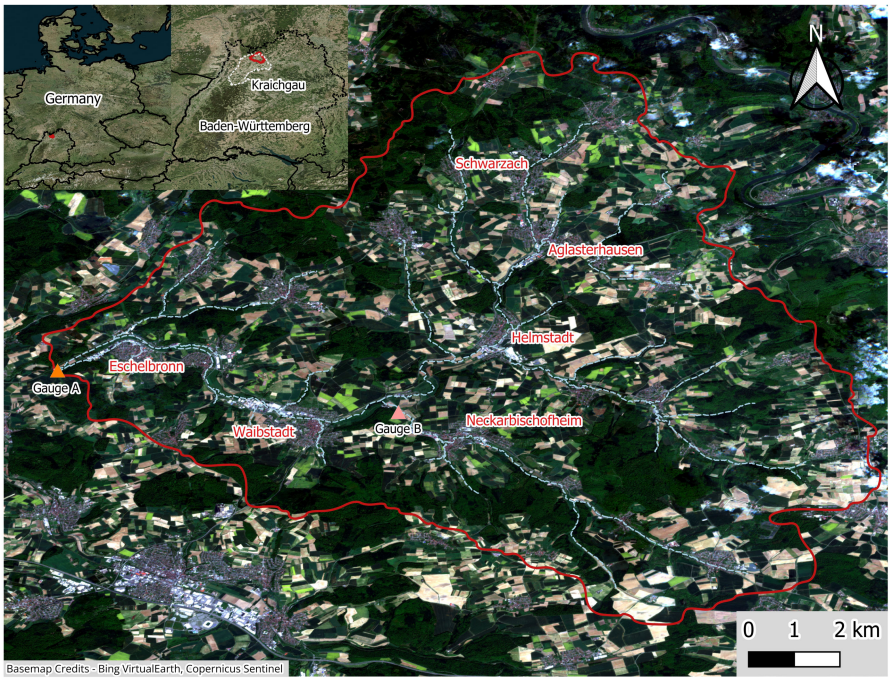

Figure 3.1: Remote sensing image showing the location of different population centres (Fig. B.6in Appendix B.1) within the mainly agricultural and forested Elsenz Schwarzbach catchment. Also shown in the inlet is the location of the catchment within Baden-Württemberg, Germany.

During the convective storm-driven summer flood (27 June) of 1994 (Appendix B.2), an estimated destruction (Figure B.8 in Appendix B.2) of 250 to 300 million Deutsche Mark was reported (Henkel, 1994) in the ten municipalities within the Elsenz Schwarzbach catchment (Appendix B.2). Sadly, three human fatalities were also reported for the event. The most severely affected (Vogt, 2024) where the towns of Eschelbronn, Helmstadt, Waibstadt, and Neckarbischofsheim (Figure 3.1 and B.6 in Appendix B.1). The temporal preconditioning of the event by the winter floods of 1993–1994 also led to a compounding of

impacts (Ruiter et al., 2020) in the region. Due to these catastrophic flooding episodes in 1993–1994, a comprehensive flood protection concept (Zweckverband Hochwasserschutz Elsenz-Schwarzbach, 2016) was developed for the entire region, leading to the construction of local flood retention basins throughout the catchment area. Around twenty flood reservoirs are now operational within the catchment with a combined design flood storage volume of 885,000 m^3. The region again faced a severe convective storm clustering in June 2016, leading to localised flash floods (Manoj J et al., 2024) overtopping the flood reservoirs in some regions (Figure B.10 in Appendix B.2).

3.2.2 *Representative Hillslopes*

Balancing the intricacies required for process based models with the required simplicity stemming from parsimony considerations (Hrachowitz and Clark, 2017) has long been considered one of the holy grails of hydrology. While top-down approaches (such as the unit hydrograph) offer readily applicable solutions, they are often criticised for their simplifying assumptions. Models resting on bottom-up paradigms (Fatichi et al., 2016), on the other hand, demand a larger parameter space and higher computational complexity.

The concept of representative hillslopes (Loritz et al., 2017) is an attempt to bridge the gap between these two worlds. Building upon related works (Cochrane and Flanagan, 2003; Francke et al., 2008) on deriving representative hillslope catena for catchments, the approach addresses the question of what is a meaningful approximation for the large number of dominant flow paths in a hydrological system, and up to what scale do such approximations hold good. By creating and modelling a *fictitious hillslope* (which doesn't really exist in the system but is rather an derivation of the averaged distribution of potential energy of all the existing hillslopes along the averaged distance to river), the approach shows commendable results (Loritz et al., 2017; Manoj J et al., 2024; Villinger et al., 2022) in reproducing the rainfall-runoff behaviour in headwater catchments up to 20 km^2. Here we expand this concept to the scale of the entire Elsenz Schwarzbach catchment – the idea is to connect representative hillslope models for the sub-catchments with river network and a hydraulic routing scheme

3.2.3 *Modeling approach*

3.2.3.1 Hydrological model

Beginning from around 1990, the Weiherbach catchment (Figure B.1 in Appendix B.1) in the Kraichgau region was the focus of detailed

experiments on the interplay between infiltration, runoff generation and related transport processes of solutes and sediments with a focus on explanatory, physically based modelling (Bronstert et al., 2023; Zehe et al., 2001). The spatially distributed, process based hydrological model CATFLOW (Zehe et al., 2001) was developed as part of this detailed field investigations. The basic modelling unit is a 2D hillslope. Water and solute dynamics of each hillslope element are simulated on a terrain following curvilinear grid. Therefore, the fundamental equations for mass and water transport must be transformed from their typical Cartesian form into curvilinear coordinates.

Assuming that $u(\eta, \xi)$ and $v(\eta, \xi)$ are the mapping functions of the curvilinear coordinate directions η and ξ into the cartesian coordinates x and z such that

$$x = u(\eta, \xi) \qquad z = v(\eta, \xi) \tag{3.1}$$

Since the physical line element (distance, $ds^2 = dx^2 + dy^2 = d\eta^2 + d\xi^2$) has to be invariant under the transformation and by using chain rule of differentiation successively:

$$\frac{\partial}{\partial x} = \frac{1}{f^\xi}\frac{\partial}{\partial \xi} \qquad \frac{\partial}{\partial y} = \frac{1}{f^\eta}\frac{\partial}{\partial \eta} \tag{3.2}$$

where

$$f^\xi = \left(\frac{\partial u}{\partial \xi}\right)^2 + \left(\frac{\partial v}{\partial \xi}\right)^2 \qquad f^\eta = \left(\frac{\partial u}{\partial \eta}\right)^2 + \left(\frac{\partial v}{\partial \eta}\right)^2 \tag{3.3}$$

are the metric coefficients that have to be evaluated locally at each grid point. The use of curvilinear coordinates simplifies modeling complex hillslope soil surfaces, allowing for more accurate representation of surface runoff and erosion processes. Soil water dynamics within the hillslopes are then characterized using the 2D Darcy–Richards equation in these transformed curvilinear coordinates.

$$\begin{aligned}\frac{\partial \theta}{\partial t} = &\frac{1}{f^\xi}\frac{\partial}{\partial \xi}\left\{K(\theta)\left[k^{\xi\xi}\frac{1}{f^\xi}\frac{\partial}{\partial \xi}(z-\psi) + k^{\xi\eta}\frac{1}{f^\eta}\frac{\partial}{\partial \eta}(z-\psi)\right]\right\} \\ &+ \frac{1}{f^\eta}\frac{\partial}{\partial \eta}\left\{K(\theta)\left[k^{\eta\xi}\frac{1}{f^\xi}\frac{\partial}{\partial \xi}(z-\psi) + k^{\eta\eta}\frac{1}{f^\eta}\frac{\partial}{\partial \eta}(z-\psi)\right]\right\} - \omega\end{aligned} \tag{3.4}$$

The coefficients $k^{\xi\xi}, k^{\eta\eta}, k^{\xi\eta}$ account for the anisotropy of the hydraulic conductivity and ω represents the source and sink terms. The soil hydraulic conductivity $K(\theta)$ as a function of the relative saturation ,S and the soil water retention function (the soil water content, θ as a function of the matrix potential ψ) can be parameterized according to van Genuchten (1980) and Mualem (1976), Tang and Skaggs (1977) or the recently proposed PDI model (Peters et al., 2021).

Overland flow heights are simulated using the diffusion wave approximation to the Saint-Venant equation and explicit upstreaming, in combination with the Gauckler-Manning-Strickler formula. (Hager, 2015).

The model can represent preferential macropore flow by using an effective macroporosity factor. This factor (Loritz et al., 2017; Wienhöfer and Zehe, 2014) scales the ratio of infiltration into the macropore domain compared to the matrix domain. For the effective macroporosity factor approach, as detailed in Zehe et al. (2001), the bulk hydraulic conductivity k_s^B, is linearly increased by a relative scaling factor, f_m (if relative saturation (S) at a grid point exceeds a certain specified threshold (S_o) but only till complete saturation) as follows:

$$k_s^B = \begin{cases} k_s + k_s f_m \dfrac{(S - S_o)}{(1 - S_o)}, & \text{if } S_o \leq S < 1, \\ k_s & \text{otherwise.} \end{cases} \tag{3.5}$$

Thus, emulating the preferential flow paths (of increased saturated hydraulic conductivity and consequently larger flow velocities) through which water could flow to the deeper layers in the hillslope.

3.2.3.2 River routing model

The river routing model, which connects the subbasins through the main channel, uses a one-dimensional kinematic wave approximation of the Saint-Venant equations (Fenton, 2019; Troy et al., 2025). The law of conservation of mass is expressed as:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \tag{3.6}$$

Where Q is streamflow, x is distance along the river channel, A is the cross-sectional area of the river channel, t is time, and q is the lateral flow from the representative hillslopes into the river channel. Under the kinematic wave approximation, local and convective acceleration as well as pressure gradient terms, are neglected in the full dynamic momentum equation, and the friction slope (s_o) is assumed equal to the bed slope (s_f). The momentum equation then reduces to an algebraic relation between streamflow, Q and the flow area, A. The Gauckler-Manning-Strickler formula is then used for the same.

$$Q = K_{st} \sqrt{(s_o)} \left(\frac{A}{P} \right)^{2/3} A \tag{3.7}$$

Where P is the wetted perimeter, A is the wetted cross section, and K_{st} is the strickler roughness coefficient (which is also more commonly

expressed as inverse of the Manning roughness coefficient, n). Eqn (3.7) is then substituted into Eq. (3.6), resulting in one equation with one unknown. This is then solved using a finite difference scheme that enforces a Courant condition on the time step, ensuring numerical stability and monotonicity of the solution, thereby suppressing numerical dispersion.

3.3 METHODOLOGY

3.3.1 *Setting up the model*

3.3.1.1 System Geometry

The representative hillslope profile lines were extracted individually for all the subbasins (124 in total: Fig. 3.2A) of the Elsenz Schwarzbach in parallel using the workflow outlined in Appendix B.3. Figure 3.2 depicts the main methodological steps involved in the extraction process of one such representative hillslope profile (Fig. 3.2B). Firstly, the digital elevation model is preprocessed to fill all depressions and sinks. This is followed by the extraction of other rasters (Fig. 3.2C). The relative elevation of each cell above the nearest stream segment and the corresponding distance to the stream segment define the flow profile lines (Fig. 3.2B) in each subbasin. Flow profile lines can be defined as the paths water takes as it flows from one cell to the next, starting from a cell with no water inflow at the hillcrest and terminating at a river cell. Our aim is to derive a single representative hillslope for the subbasin rather than modelling all these flow profile lines individually. This is done by considering the total potential energy of all the hillslopes (Manoj J et al., 2024).

Preserving this total energy implies that topography of the new representative hillslope should be chosen to maintain average topographic gradients along the flow path. We achieve this by binning the geopotential energy based on proximity to the river and then performing a weighted average (Francke et al., 2008) using the flow accumulation values) within each segment. Consider all the cells (Fig. 3.2B) at a relative distance of x m from the nearest stream cell, we require an estimate of relative elevation, $\bar{h}(x)$ for our representative hillslope, by evaluating the elevation (h) of all cells, at the same distance. This elevation is obtained by multiplying the elevation of each cell by their corresponding flow accumulation values (f), which denotes the number of flow profile lines that pass through a cell and hence its relative importance. The process is then repeated for all the binning distances.

$$\bar{h}(x) = \frac{\sum_{i=1}^{\text{All cells at } x} h_i^x \sqrt{f_i^x}}{\sum_{i=1}^{\text{All cells at } x} \sqrt{f_i^x}} \tag{3.8}$$

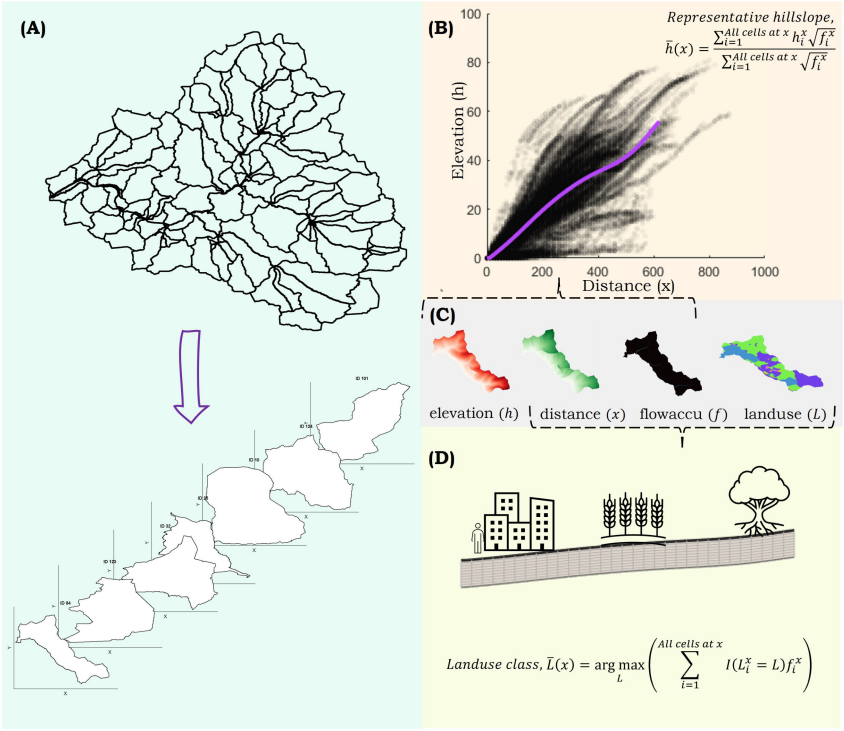

Figure 3.2: Schematic overview of the representative hillslope derivation. The catchment is first divided into different subbasins (Fig. 3.2A), the representative hillslope profile for each subbasin (Fig. 3.2B) is derived from the elevation, distance and flow accumulation raster files (Fig. 3.2C). A similar approach is utilised to get the landuse distribution along each representative hillslope (Fig. 3.2D). The process is then repeated for all the subbasins in the catchment (Fig. 3.2A) in parallel using the reproducible research workflow detailed in Appendix B.3.

Interested readers are referred to Loritz et al. (2017) and Manoj J et al. (2024) for a validation of the total energy conservation using different distance classes and a more detailed explanation of the methodology. The same process is then repeated for all the other subbasins (Figure B.2 & B.11).

We converted the stream network (Fig B.3) of the entire catchment into a system of interconnected channel nodes, connecting each representative hillslope at the location of the corresponding sub-basin outlet. The 1D kinematic wave routing was established to model the propagation of the flood wave along the channel network. Due to the absence of detailed field data, the Gauckler-Manning-Strickler roughness (Throughout the remainder of this work, we use both Manning's roughness and Strickler values interchangeably to refer to the roughness coefficient (K_{st}) in the Gauckler-Manning-Strickler formula. Interested readers are referred to Hager (2015) for a historical anecdote) coefficients for stream routing were assigned based on literature values (United States Geological Survey, 1989; Ven Te Chow,

1959). Lower-order streams tend to be narrower, more vegetated, or irregular, resulting in greater resistance to flow than the well-defined, larger, higher-order channels (Heldmyer et al., 2022). Consequently, streams with a higher Strahler stream order (Strahler, 1957) were assigned correspondingly higher Strickler values, as indicated in Table 3.1.

Table 3.1: Gauckler-Manning- Strickler Coefficient values for different stream orders

Stream Order	Gauckler-Manning- Strickler Coefficient (K_{st} – $m^{1/3}\,s^{-1}$)
1	15
2	20
3	30
4	40

3.3.1.2 Soil and landuse mapping

The mesoscale catchment shares similar geological and pedological characteristics with the heavily experimented Weiherbach catchment (Zehe et al., 2001). Hence, we transferred the relative hillslope soil distribution profile (Loess and Coluvisols) and corresponding hydraulic functions (which include the van Genuchten parameters: See Table 4 in Zehe et al., 2001) for these soils from the Weiherbach catchment.

Accurate land-use scenarios are a prerequisite for assessing the impact of landuse gradients on runoff generation at the catchment scale. We use the Dynamic World landuse dataset (Brown et al., 2022) for mapping the dominant landuse classes in our catchment. The classification taxonomy involves nine dominant classes. The relevant classes in our study area (Fig. B.4) and the corresponding Manning roughness values chosen (again informed from the Weiherbach field experiments: Zehe et al., 2001) are shown in Table 3.2.

Table 3.2: Gauckler-Manning- Strickler Coefficient values for different landuse classes

Landuse ID	Class	Gauckler-Manning- Strickler Coefficient (K_{st} – $m^{1/3}\,s^{-1}$)
1	Trees	8
2	Grass	10
4	Crops	9
5	Shrub	10
6	Urban	30
7	Bare Earth	12

For each subbasin (and corresponding representative hillslope), we now require the relative distribution of the different landuse classes along the gradient. This is done analogously to our extraction of the representative hillslope profiles. We were interested in getting a mean value for elevation there; here, the focus shifts to getting the dominant landuse class for each distance (Fig 3.2D). This is attempted by looking at the relative flow accumulation values for each cell with a specified landuse class. Instead of a weighted average, we choose the class with the largest areal share.

To determine the representative land-use class $\bar{L}(x)$ at a relative distance of x m from the nearest stream cell, we consider all raster cells located at that same distance in the different hillslope profiles (Fig 3.2B). For this group of cells, we evaluate the land-use category by weighting each class (L) according to the flow accumulation of the cells that belong to it. Thus, giving greater influence to cells through which many flow profiles pass. We then sum the weighted contributions for each landuse class separately. Finally, choosing the class with the highest value for this weighted sum as the representative landuse class element at this distance. Mathematically (Eqn 3.9), we seek a value L that maximises the weighted sum of the indicator function $I(L_i = L)$, which equals 1 when a cell belongs to the class L and 0 otherwise; multiplied by its corresponding flow-accumulation value f_i.

$$\bar{L}(x) = \arg\max_{L} \left(\sum_{i=1}^{\text{All cells at } x} I(L_i^x = L)\, f_i^x \right) \tag{3.9}$$

3.3.1.3 Initial and Boundary conditions

Previous works (Liu et al., 2025)) have demonstrated that initialising a process-based model with climate reanalysis data can help address the challenges of data availability in data-scarce regions. For our simulation of the historical flood that occurred in June 1994, we opted to use ERA5 Land Reanalysis (Muñoz-Sabater et al., 2021) to initialise the model. The initial conditions for each model run were derived (after normalisation) from the hourly soil moisture band values, specifically: "*volumetric_soil_water_layer_1 - 4*", from the ERA5 Land reanalysis product.

For the historical summer flood of 1994 (Appendix - B.2), gauges with finer resolution precipitation estimates are unfortunately not available within the catchment boundary of the Elsenz Schwarzbach, while the German Weather Services provides daily accumulated sums (DWD, 2024; Rauthe et al., 2013) in the HYRAS gridded product (our study area received an average precipitation of 90 mm) this is not enough for estimating the dynamics of flash flood generation due to intensity

excess Hortonian flow. To overcome this limitation of forcing data, we make use of the quality checked higher resolution (6 minutes) precipitation data recorded at the gauging station (Zehe et al., 2001) in Weiherbach experimental catchment (total sum – 83 mm). Since the Weiherbach catchment also had a strong runoff reaction during the event and due to its spatial proximity (around 20 km), we assume that the storm intensity distribution is similar. We derive daily sums individually for each subbasin (representative hillslope) in our catchment and then assume the same distribution (rainfall hyetograph) of storm as recorded in the Weiherbach gauge. The derived series is then used to simulate each representative hillslope separately. To validate our approach, we also conducted simulations of a more recent flood event (April 2023) using the fine-resolution radar precipitation estimates (DWD, 2021) from the German Weather Services.

3.3.2 *Evaluating design adequacy*

In Germany, flood reservoirs are typically designed in accordance with the DIN 19700 standard (LUBW, 2007). Dams are classified into four categories based on their height and capacity: very small, small, medium, and large reservoirs. The design must account for a flood event with a return period of 50 - 10,000 years, depending on the category of the reservoir. The German Weather Services and the State Institute of Environment, Baden-Württemberg (LUBW) provide location specific standardised estimates for extreme precipitation sums corresponding to different return periods and storm durations (DWD, 2023; Junghänel et al., 2017; LUBW, 2023) which are then used for design purposes. We ran our validated model for five different scenarios (Table 3.3) to simulate design floods for various return periods (assuming same storm duration as the 1994 flood) along the Elsenz Schwarzbach.

Table 3.3: Design precipitation sums for different return period floods according to DWD (2023) and LUBW (2023)

Return period (years)	Precipitation (mm)
100	55.73
200	61.5
500	69
1000	74.5
10000	93

3.3.3 *Implementing nature based solutions (NbS)*

Historically, vegetative features such as hedgerows were integral to agricultural landscapes, marking the boundaries of individual

landowners' plots. This network of vegetation provided significant benefits to the catchment's surface hydrology (Rosier et al., 2024). By creating impediments to overland flow, hedgerows slowed down water movement and enhanced soil infiltration. However, due to a lack of previous scientific evidence (Rosier et al., 2023) supporting their effectiveness and to changes in land ownership patterns (most notably the shift from smaller to larger consolidated fields), these elements have largely disappeared from many agricultural regions in Europe. After the 1994 floods in the Elsenz Schwarzbach, the consequences of disappearing hedgerows (Fig. B.9 in Appendix B.2) once again became a topic of discussion, as evidenced by historical newspaper archives (Henkel, 1994) from that period. On the other hand, contemporary forest management practices often involve the use of heavy machinery, which has seen an increase in wheel loads (Auerswald et al., 2025) over the past century. This has resulted in increased subsoil compaction (Brus and Akker, 2018; Schneider and Don, 2019), further diminishing the infiltration capacity of the soil.

To model the impact of nature-based solutions on the destructive floods of 1994, we selected scenarios for different landuse classes (fig. 3.2D and Table 3.2) in our catchment. The first involves establishing a row of hedgerows (Rosier et al., 2023, 2024) within the agricultural fields (landuse class – crops: Table 3.2) of the catchment area. The second assumes the implementation of practices aimed at minimizing soil compaction (Auerswald et al., 2025), such as using designated skid trails for heavy machinery in forested areas (landuse class – trees: Table 3.2). To assign the locations of the hedgerows, we first identified all continuous stretches of landuse type crops on the representative hillslope element. We then placed a hedgerow land-use class in the middle of each of these stretches, ensuring the hedgerow length was approximately 10% of the total length of the cropped area. The practices aimed at minimising soil compaction were applied in all portions of the representative hillslope with landuse type trees.

Table 3.4: Summary of the main NbS measures considered in our work

Implementation	Parameter	Changes
Hedgerows in the middle of the area designated as crops in the representative hillslope	Gauckler-Manning- Strickler Coefficient (K_{st})	Reduced by 50% compared to landuse – crops
	Saturated hydraulic conductivity of the first 40 cm of soil (k_{sat})	Increased by an order of 2 compared to landuse – crops
	Saturated water content of the first 40 cm of soil (θ_{sat})	Increased by 20% compared to landuse – crops
Practices aimed at minimising soil compaction for forest regions spanning the representative hillslope	Saturated hydraulic conductivity of the first 40 cm of soil (k_{sat})	Increased by an order of 1 compared to landuse – trees
	Saturated water content of the first 40 cm of soil (θ_{sat})	Increased by 10% compared to landuse – trees

While substantial research has been conducted to understand how these nature-based solutions influence key parameters in hydrological modeling, uncertainties persist, as their effects are highly location-specific. In this study, we utilise recommendations based on previous research (Guillaume et al., 2025; Richet et al., 2017) about the Manning-Strickler roughness (K_{st}), saturated hydraulic conductivity (k_{sat}) and saturated water content of soils (θ_{sat}). This focus is primarily due to the study regions located in France (Richet et al., 2017) and Belgium (Guillaume et al., 2025), included agricultural loess soils similar to the Elsenz Schwarzbach in Germany. The parameter values for the new hedgerow landuse class were provided and modified for the tree landuse class as outlined in Table 3.4.

3.4 RESULTS

3.4.1 *Model building and testing*

We simulated representative hillslopes using the CATFLOW hydrological model for all subbasins of the Elsenz Schwarzbach during the 1994 summer event and then used the 1D channel routing to estimate the flood at the outlet (Gauge A).

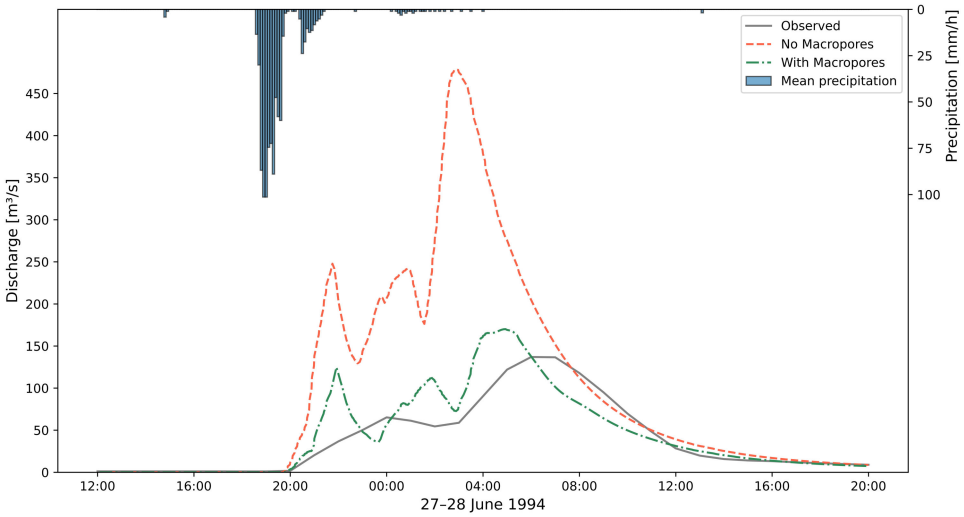

Figure 3.3: Modelled hydrographs (smoothed 1-hourly windows) compared to the observed outlet time series. The green curve represents simulations with macropores, while the orange curve shows model run without considering any macropore flow.

Figure 3.3 shows the overestimation of the observed flood dynamics by our initial modelled set-up. These overestimations align with findings from other studies (Beven and Germann, 1982; Bronstert et al., 2023; Niehoff et al., 2002; Weiler, 2005), which reports that in soils with very low soil hydraulic conductivity (as in our case with the

agricultural loess soils), when confronted with high intensity rainfall during convective storms, a large part of infiltration may pass through so-called soil macropores into deeper soil layers, thus not contributing to overland runoff.

To simulate the effect of macropore flow during the high intensity storm of 1994, we used the simplified macroporosity approach (Eqn. 3.5). This essentially replicates the increased infiltration due to macropore flow by relative scaling of the bulk hydraulic conductivity along the representative hillslope element. A scaling factor of 10 is chosen, consistent with findings from other studies in the regions (Manoj J et al., 2024; Niehoff et al., 2002). We assume increased macropore flow along 90% of the hillslope catena. The hillslopes are again simulated, considering the macropore effect. As expected, the new flood simulated flood hydrograph with macropores is closer to the observed curve (with a Nash Sutcliffe Efficiency (Nash and Sutcliffe, 1970) value of 0.58). The total flood volume (Table 3.5) aligns closely with the observed flood volume (relative error of -12 %). We also observed that the dynamics of the flood are well captured. The model accurately captures the rise of the flood wave at around 20:00 CET and 04:00 (+ 1 day) CET. While there are some overestimations and sharper recessions (possibly due to our choice of 1D hydraulic routing), the model captures the flood peak with a relative error of -24%. The subsequent return to normal flow levels is the same in both runs at around 12:00 CET (+1 day). It is interesting to see that the flood response still clearly reflects the temporal pattern of the rainfall event, even though the latter is an educated guess based on the observational record from the nearby Weiherbach catchment on the same day.

Table 3.5: Flood event characteristics at the outlet gauging station near Eschelbronn (Gauge A)

Flood Metrics	27.06.1994		28.04.2023	
	Flood Volume (1000 m^3)	Flood Peak (m^3/s)	Flood Volume (1000 m^3)	Flood Peak (m^3/s)
Observed	4683.74	136.97	1146.35	24.31
Simulated	5276.64	170.06	1232.75	38.75
Relative Error (%)	-12.6	-24.2	-0.1	-59.4

To again validate the performance of our model on another event, we simulated a flood (with far less intensities compared to the 1994 flood) event that occurred in April 2023. It is important to note that reservoirs introduce greater uncertainty into simulations of flood peaks in more recent periods, as they are now heavily regulated (Zweckverband Hochwasserschutz Elsenz-Schwarzbach, 2016) and the present set-up does not account for reservoir operations at this stage. Our model

run again showed low errors in the total flood volume (Table 3.5). However, the higher modelled flood peak led to greater relative errors in peak discharge.

3.4.2 *Reconstruction of the 1994 flood*

The model set up is then utilised to investigate the flooding within the Elsenz Schwazbach catchment during June 1994. While this historical flood lead to widespread damage (Fig. B.8) within the catchment, as recorded in grey literature (Henkel, 1994; Vogt, 2024), not much information is know about the relative storm volumes and peak values that hit the different population centres within the catchment. The gauge near Neckarbischofsheim (Gauge B in Fig. 3.1) records only flood-level markings, and because the observed level was higher than the established rating curve, it is not possible to derive the flood volume for this event from observations alone.

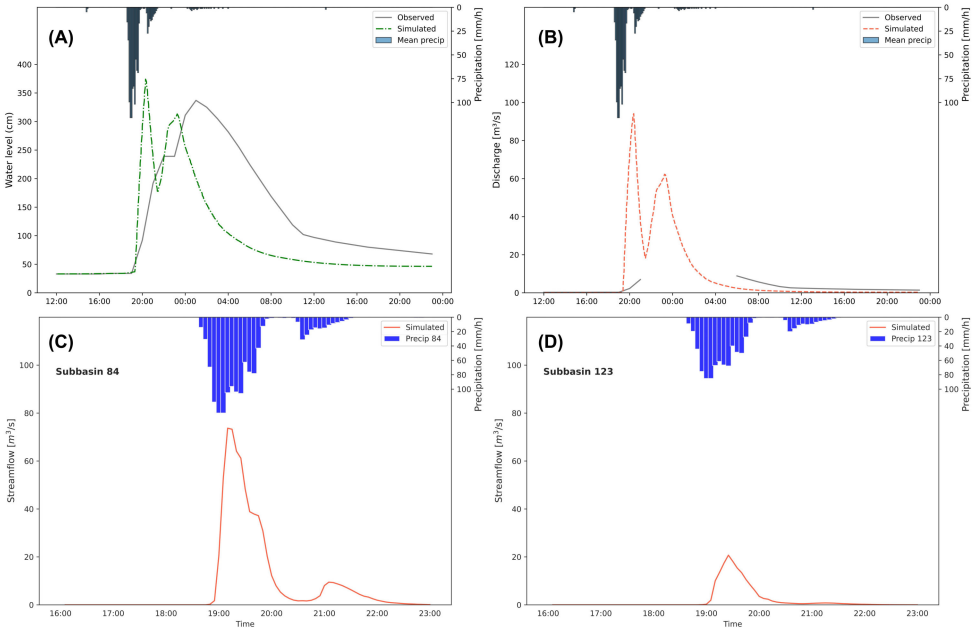

Figure 3.4: Comparison between simulated (orange) and observed (grey) water levels at the Neckarbischofsheim gauge (A); reconstructed flood hydrograph at Neckarbischofsheim (B); simulated flood hydrograph for hillslope 84 during the June 1994 event (C); and simulated flood hydrograph for hillslope 123 during the June 1994 event (D).

We compared the water level observations at Gauge B to the channel water levels modeled in our setup (Fig. B.3) at the nearest node. Figure 3.4A shows the simulated wave heights alongside the observed water levels. Although the model does not replicate the observed curve exactly, it satisfactorily captures the overall timing, magnitude, and recession behaviour. It is also important to consider the uncertainties

in the observations, especially for such short-duration, high-impact events. While the observational records indicate that the peak occurred approximately 8 hours after the flood began, experiences with rapid responses (up to 2-3 hours) during such events in areas of this scale (catchment area up to Gauge B = 35 km²) question the accuracy of the measurements during the flood.

The double peak nature of the flood hydrograph (Fig. 3.4B) warrants a closer look at the smaller headwater catchments that drain to this city. We focused on two headwater subbasins (Fig. B.5 in Appendix B.1) that subsequently drain into the city of Neckarbischofsheim.

Subbasin 84 encompasses part of the city, which is situated at the foot of the basin. The impervious area in such downstream parts leads to low infiltration and, conversely, faster flow. This compounds the flood hazard for the city. As seen from the flood hydrograph (Fig. 3.4C) for the subbasin, there is an abrupt rise in the rising limb, possibly due to the urban built up area downstream of the subbasin. The location of the agricultural plots upstream (Fig. B.5) in subbasin 84 also leads to more sediment and debris flow to the city, as seen from archived footage of the aftermath of the flood in Neckarbischofsheim.

Subbasin 123 is located (Fig. B.5) just upstream of the town of Obergimpern, the basin is predominantly agricultural in nature with large open fields. The large open fields without any hedgerows or rills again exacerbate the flooding by more sediment flow into the stream. The slower rising limb (Fig. 3.4D), when compared to hillslope 84 (Fig. 3.4C), suggests gradual overland flow generation with erosion, also more mixing time for the suspended sediments in the channel. Analysing the individual responses reveals that the earlier and higher observed flood peak at Gauge B (Fig. 3.4B) predominantly originates from the headstream regions (Subbasin 84), which drain directly into it. In contrast, the smaller second peak is associated with other catchments that flow into the city along the river network. The shorter travel duration between the two peaks indicates a scenario of temporal preconditioning and hazard cascading (Zscheischler et al., 2020), where the initial flood has already created a hazardous situation that is then worsened by the subsequent flood, leaving very little time for recovery or rehabilitation.

It is curious to note that the flood reservoir downstream (designed for a 100 year flood) of subbasin 123 was again overtopped (Fig. B.10) in the summer flood series of 2016 (Reservoir W22 in Manoj J et al., 2024), raising interesting questions about the adequacy of existing

design measures in response to increasing flash floods in a warming climate.

3.4.3 *Unraveling the disconnect in design floods*

Using our model setup we ran simulations for the entire Elsenz Schwarzbach catchment to predict the various return period floods from the standard design extreme precipitation estimates (Table 3.3) provided by the German Weather Services (Junghänel et al., 2017). The initial conditions and storm duration were assumed to be the same as those of the 1994 summer flood. The flood values predicted by the model were then compared to the official statistical T-year return period floods currently established for planning and design practices for flood protection in the German state of Baden-Württemberg.

Table 3.6: Comparison of the simulated and standard design floods over the two gauging stations

Return period (years)	Eschelbronn (Gauge A)			Neckarbischofsheim (Gauge B)		
	Simulated	Design flood	Error (%)	Simulated	Design flood	Error (%)
100	83.92	88.56	5.2	24.35	14.4	-69.1
200	96.73	99.2	2.5	27.98	16.4	-70.6
500	129.08	114	-13.2	35.22	19.2	-83.4
1000	141.8	125	-13.4	39.32	21.5	-82.9
10000	239.81	166	-44.5	63.04	29.9	-110.8

The comparison between simulated and design floods reveals generally close agreement at Gauge A (catchment outlet) across most return periods, with deviations within ±15% except for the 10,000-year event, where the flood is notably underestimated. In contrast, Gauge B shows substantial underestimation across all return periods, with errors exceeding 70%. The results reveal a stark contrast: for the larger outlet gauge, the current design flood values are largely in line with the T-year return period flood simulated using the extreme precipitation statistics, while this doesn't hold true for the smaller gauging stations. This also partly explains the overtopping of the headwater reservoirs upstream of Neckarbischofsheim (Manoj J et al., 2024) during the 2016 summer floods as detailed in the discussion. And it demonstrates that the current state of the art in flood protection does not adequately account for flood events influenced by rainfall intensity. Additionally, it underscores that while structural measures are often centralised and can be prone to failure, there is a critical need to incorporate non-structural, distributed measures that offer benefits regardless of extreme weather events (Guillaume et al., 2025).

3.4.4 *Impact of Nature-based solutions*

To assess the impact of nature-based solutions (NbS) during flash floods, we conducted a virtual experiment (see Table 3.4) that incorporated hedgerows and forest management practices in Subbasins 84 and 123. We compared the baseline scenario (Fig. B.9), which did not include any nature-based solutions, to a new scenario that implemented both these measures at the corresponding landuse class occurrences in the representative hillslope profile.

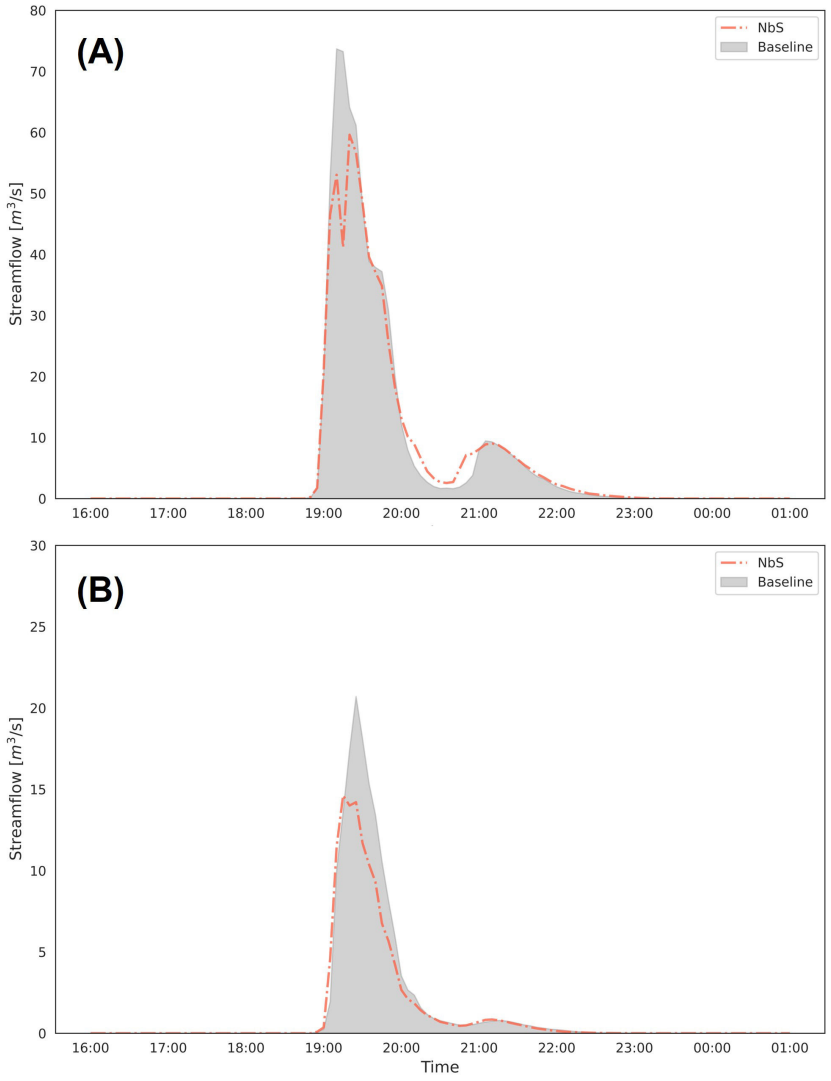

Figure 3.5: Impact of nature based solutions (NbS) at subbasins 84 (Fig. 3.5A) and 123 (Fig. 3.5B)

The run (Fig. 3.5) with the NbS showed reduced flood volume and peak discharge for both the subbasins. In subbasin 84, the relative volume reduction was around 6.5% while the peak discharge was reduced by around 19%. This again proves that while the NbS do

not necessarily change the major hydrologic characteristics leading to runoff generation of the area. It acts as a modulator (similar to a downstream flood reservoir) ensuring that the peak flood is attenuated at the local field scale before being released to the flood channel. For subbasin 123, the relative volume reduction was 19.6%, while the peak reduction was 29.3%. The reductions in peak runoff at these two headwater catchments led to a reduction of around 16 cumecs (11%) in the peak discharge and 23 cm (6%) in flood wave heights simulated at the Neckarbischofsheim gauge.

3.5 DISCUSSIONS

3.5.1 *Upscaling vectorised representation*

Nearly 40 years after Dooge's seminal paper (Dooge, 1986), which identified intermediate-scale catchments (up to 250 km^2) as systems of organised complexity, challenges remain in modelling flash flood responses (Collier, 2007) driven by Hortonian overland flow at these scales. These catchments are too organized to be represented solely through statistical methods, yet too large and heterogeneous to be described in a deterministic manner. A key requirement (Zehe et al., 2014) at such scales is a better understanding and representation of how different forms of spatial organisation affect storage and release of water and energy. This is particularly important for threshold or emergent behaviour (Zehe and Sivapalan, 2009) like onset of preferential flow (Beven and Germann, 1982).

In this study, we test a coupled hydrological - hydraulic model at such an intermediate scale (196 km^2) relying on a vectorised representation of sub-catchments as representative hillslopes that maintains the total flow potential of a subbasin and then coupled it to a 1D kinematic wave module for flood routing along the main catchment. Our main motivation is to account for the total runoff volume by conserving the energy gradient while accounting for preferential flow processes. We mainly use parameter transfer from one experimental catchment and literature values to ensure that our model has a parsimonious structure. While the theory of representative hillslopes has attracted increased attention (Fan et al., 2019) here we show that the approach can be coupled with standard flow routing methods for reliable rainfall-runoff simulations at the meso-catchment scale.

We revisited the devastating Kraichgau flood series in Germany in the summer of 1994 using our model setup. However, we faced the persistent hydrological issue of data availability. The normalised values of storm runoff obtained (flood maximum divided by catchment area) was in the range of the estimates for the same storm from

well-monitored test catchments (Villinger et al., 2022)) in the region. This again indicates the plausibility of our estimates.

Our choice of the 1D kinematic wave routing to connect the vectorised representation for subbasins, based on energy gradients, does not imply that the processes at the channel scale are in any way less important for a comprehensive assessment of flood dynamics during such storms. Rather, this choice is motivated by the fact that a large number of well-tested hydraulic models (Fenton, 2019) are available for channel routing. Often, the main limitation for achieving reliable predictions lies in the runoff generation aspect (Brunner et al., 2021), which determines how high-intensity rainfall during these events is translated into runoff. Our goal is to improve upon top-down approaches, such as the unit hydrograph and curve number methods, which have been found to be lacking in this context (Kirkby and Cerdà, 2021). It is worthwhile to note that the channel routing mainly modulates the flood peak and time to peak. The total volume of the storm (coming from the individual hillslopes) remains invariant.

3.5.2 *Enhancing current design approaches*

Significant advances have been made in climate physics and meteorology, particularly in modelling intense storms that align with the increased emission scenarios of anthropogenic climate change (IPCC, 2021). However, these advancements have not led to improved estimates of design flood hydrographs. This limitation is partly due to the lack of well-suited hydrological models and gaps in data records, especially at smaller headwater catchment scales, which are ironically the areas most affected by flash floods.

Unfortunately, the number of gauges in small headwater catchments has decreased in recent decades, and most gauges are now disproportionately located in larger perennial rivers (Krabbenhoft et al., 2022). This drastically impacts our ability to understand what has happened and can happen at such crucial scales (Michelon et al., 2021). Reconstruction can be a valuable tool for enhancing existing observations of high-intensity storm events.

Using our trained model, we reconstructed the flood values at one of the population centres, Neckarbischofsheim, within the catchment. The high successive discharge peaks are in line with the large scale destruction reported in the town (Fig. B.8 in Appendix B.2) during 27 - 28 June 1994. We also looked at the headwater catchments contributing to the runoff at the town to understand the hazard cascade for the event. Our main finding was that the spatial arrangement of settlements, located downstream of agricultural fields, makes them

more vulnerable to the impacts of overland flow and debris flow from upstream areas. Considering such land-use patterns and their spatial organisation while planning new settlements in the region could help reduce the risk of similar events in the future.

We then compared simulations of floods with different T-year return periods to the currently recommended design standards and found that the underestimation errors were larger at smaller spatial scales. This is an important consideration to be made in the future design of small to medium flood reservoirs in the region. While the comparison was made to the regionalised statistical model (LUBW, 2025) that relies on observed annual peak flood data and other catchment predictors currently commissioned in the federal state of Baden-Württemberg in Germany, similar models or statistical design floods exist for regions around the world.

While hydrologic non-stationarity (Milly et al., 2008) has been recognised for decades now, this hasn't yet translated into meaningful guidance for engineers and practitioners (Wasko et al., 2021). The intensification of sub-daily precipitation extremes (IPCC, 2021) is a significant challenge, as current design approaches primarily focus on the total volume of the storm rather than the intensity dependence of runoff, especially for small- to medium-sized reservoirs. The design floods for the reservoirs in the Elsenz Schwarzbach catchment relied on a hydrological model (Ihringer, 1994) employing the unit hydrograph and an empirical estimate of the runoff coefficient. This partly explains why quite a few of the reservoirs got overtopped during the intensity driven 2016 summer flash floods (caused by a 25-year rainfall), even though the reservoirs were designed for a 100-year flood.

3.5.3 *Revitalizing flood protection*

The headwater flood reservoirs in the Elsenz Schwarzbach, established after the disastrous flooding episodes in 1993-94 were overtopped (Fig. B.10) during the convective storm clustering of summer 2016. This again calls for a relook into traditional structural flood protection measures and their benefits compared to decentralized flood protection. Although the combined area of subbasins 84 and 123 is only 6.73 km², which accounts for approximately 19% of the total catchment area of 34.75 km² leading up to the gauging station at Neckarbischofsheim, the flood peaks during the 1994 flood, especially the larger initial flood wave, were significantly influenced by the rapid runoff response in these two subbasins.

We carried out virtual experiments on these two headwater sub-

basins, implementing nature based solutions (NbS) at the agricultural and forest landuse classes. This was modelled by changing key hydrological parameters in the hillslope model. We observed a pronounced reduction in flood peaks at both individual subbasins and the gauging station downstream. The total volume reduction of roughly 23,000 m^3 would have required a flood retention reservoir of comparable capacity, entailing substantial environmental and economic costs. This is in line with findings from other studies (Rosier et al., 2023, 2024) which report on the benefits of such NbS for localised runoff reduction.

Our decision to use representative hillslopes for the subbasins in our study area enables us to effectively map land use classes based on their positions relative to the topographic gradient driving overland flow. The runoff responses observed in subbasins 84 and 123 highlight the importance of considering both the relative positions of land use types and their arrangement within modelling units.

For better meso-scale flood management, it is crucial to avoid the superimposition of peaks arising from runoff generation at the different parts of the catchment. A rather homogeneous landuse distribution implies a strong connectivity in the overland flow. Measures such as hedgerows in agricultural fields reduce connectivity, providing localised impediments, slowing Hortonian runoff, and increasing local infiltrability. The related valuable delays in the time to peak help runoff from adjacent fields not to reach the flood channel simultaneously. This leads to downstream benefits along the river network.

3.5.4 *Limitations and Outlook*

Modelling of hydrological systems has benefitted immensely from different complementary (and yet sometimes competing) modelling philosophies (Fatichi et al., 2016; Hrachowitz and Clark, 2017). Our approach to scaling predictions to intermediate-scale catchments, using approximations that aim to conserve energy gradients, reproduces Hortonian runoff generation during a historical flood while addressing data scarcity. This method proves to be a valuable virtual laboratory for investigating design flood adequacies and different mitigation scenarios at such scales. However, certain limitations of the study warrant closer examination.

Our choice of hydrological model (CATFLOW - Zehe et al., 2001) is motivated by prior experience applying the model in similar landscapes. In theory, the concept of representative hillslopes is appropriate for use with any spatially distributed, physically based model (PARFLOW - Maxwell (2013), SERGHEI - Caviedes-Voullième

et al. (2023) or HydroGeoSphere - Jones et al. (2006)) that can explicitly account for the gradients represented by the hillslope geometry.

Regarding the study area, the Elsenz Schwarzbach was chosen as it provides an ideal testing ground for reconstructing a historical flood that had substantial socio-economic impacts, and because its population settlements are embedded within a mosaic of agricultural and forested land, which is well suited to evaluating the potential of nature-based solutions to mitigate flood-related damages. The locations of the catchment within the same hydro-pedological regime of an experimental catchment imply that we can opt for a transfer of soil parameters and rely on literature values instead of calibration. It is important to note that these parameters could also be derived using soil pedotransfer functions based on openly available textural data. However, in this case, more events may be required for calibration.

Concerning the experimental design, although the 1D kinematic wave for the hydraulic routing provides a simple yet computationally efficient framework that is adequate for reproducing the overall flow characteristics of the flood wave at the meso-catchment scale, this choice implies some important processes may not be sufficiently represented in our setup. During such intense events, overbank flow may occur, and water spreads onto the wide floodplains parallel to the channel. When the water leaves the main channel, there would be temporary storage on the floodplain and delayed return flow. These dynamics cannot be accurately captured by one-dimensional kinematic wave routing. This limitation may explain the more abrupt rise of the flood wave and the earlier recession seen in Figures 3 and 4a compared to observations. Leveraging advanced 2D flow routing solvers and linking them to the output hydrographs from the hillslope models could alleviate such shortcomings. This will also enhance the estimation of inundation maps for flood management planning scenarios, thereby bridging the gap between research and practice.

3.6 CONCLUSIONS

In this study, we developed a mesoscale catchment model to predict rainfall-runoff responses during high-intensity storms. To ensure that the model accurately conserves the local driving gradient of flow, we use approximations that average along the flow profile lines within each subbasin. These so-called representative hillslopes were then connected using standard hydraulic flow routing to model the evolution of the flood wave along the catchment. The results show that this is a feasible approach to flash flood modelling: using the gradient resolving approximation ensures that we can account for localised preferential flow processes while still balancing the practical-

ity of scaling up to the mesoscale. This addresses the computational complexity issues of fully distributed models and the simplistic linear assumptions of top-down models.

Our approach, which seeks to find the right balance between different modelling paradigms, has the potential to play a significant role in understanding flood-generating mechanisms and in future-proofing existing flood management policies for a warming climate driven by anthropogenic climate change. As shown in our study area in southwest Germany, severe underestimations exist at the smaller spatial scales; however, environmentally sound, non-structural flood mitigation measures, such as hedgerows, could effectively address these hazards and contribute to comprehensive flood reduction. This study contributes to the growing literature on flash flood modelling and lays the foundation for a more comprehensive risk assessment for future flood designs.

Part IV

CAN DISCHARGE BE USED TO INVERSELY CORRECT PRECIPITATION?

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

Manoj J, A., Loritz, R., Gupta, H., and Zehe, E. (2025) "Can discharge be used to inversely correct precipitation?" Hydrology and Earth System Sciences 29, pp.6115–6135. https://doi.org/10.5194/hess-29-6115-2025.

4

CAN DISCHARGE BE USED TO INVERSELY CORRECT PRECIPITATION?

ABSTRACT

This study explores the feasibility of using the information contained in observed streamflow measurements to inversely correct catchment-average precipitation time series provided by reanalysis products at the continental scale. We explore this possibility by training LSTM ensemble networks to inversely predict precipitation by using the streamflow of catchments as additional input. The first model uses discharge as an input feature along with other meteorological variables, while the second model uses only the meteorological predictors. Analysing the performance of both models showed that the discharge information not only led to an average improvement overall, but also resulted in a significant improvement (around 30 %) on days with precipitation amounts greater than 5 mm. An out-of-sample test showed that the inversely estimated precipitation is better able to reproduce small-scale, high-impact events that are poorly represented in the reanalysis product. Further, using the inversely generated precipitation time series for classical hydrological "forward" modeling resulted in improved estimates for streamflow and soil moisture. Given that the wealth of streamflow gauges around the world is currently underutilised for meteorological applications, our findings have significant implications for achieving better estimates of precipitation associated with high-impact flood events.

4.1 INTRODUCTION

The performance of hydrological models has traditionally been constrained by the availability and quality of observations covering various aspects of the water cycle. Among those, precipitation and streamflow observations are pivotal, as they represent cause-and-effect in the context of system dynamics. Long-term experimental data from well-studied research catchments, and data from operational monitoring networks, have thus long been the cornerstone of the hydrological sciences (Tetzlaff et al., 2017). The relevance of observed data and research observatories cannot be overemphasised, particularly due to the invalidity of stationarity assumptions (Milly et al., 2008) in the face of anthropogenic climate change and its impacts on water-related hazards and availability.

As the availability and quality of observations crucially constrain the "realism" of a hydrological model and thus the accuracy of predictions, data scarcity impedes accurate modelling and inference of hydrological processes. Global reanalysis products (Muñoz-Sabater et al., 2021; Onogi et al., 2007; Rienecker et al., 2011) can potentially, if of sufficient quality, complement the few existing ground-based observations by offering a valuable alternative when exhaustive local observations are not available. Further, they play a pivotal role in hydro-climatic research (Alexopoulos et al., 2023; Gu et al., 2023)), by providing a consistent, long-term view of the state of the global climate system via the assimilation of measurements and monitoring data into numerical weather models.

While previous studies (Essou et al., 2016; Tarek et al., 2020) have already shown the value of using reanalysis data as estimates for meteorological forcing data in regions with little or sparse ground-based weather station data, serious concerns about their quality remain when used in the context of hydrological modelling. The main issues include (Tarek et al., 2020)–(i) regional variations in data quality and (ii) limited representation of local hydro-meteorological processes, with both of these impacting/biasing model structures and simulated states and fluxes. Systematic biases are also critical obstacles to the broader applicability of such products (Clerc-Schwarzenbach et al., 2024). In the case of ERA5-Land, a component of the Copernicus Climate Change Service (C3S) provided by the European Centre for Medium Weather Forecasting (Muñoz-Sabater et al., 2021), there is a known tendency to significantly overestimate potential evapotranspiration (Clerc-Schwarzenbach et al., 2024; Kratzert et al., 2023; Xu et al., 2024). Deficiencies have also been documented in the representation of convective storms (Essou et al., 2016; Taszarek et al., 2021) with subsequent underestimation of precipitation magnitudes and intensi-

ties (Manoj J et al., 2024).

It is important to stress that "true" precipitation estimates are per default unknown at the catchment scale. We obtain estimates of them (with considerable uncertainty) by either interpolating data from stations in or surrounding the catchment or averaging gridded data from reanalysis/remote sensing products to the catchment scale. Such precipitation uncertainty is rarely considered when quantifying model output uncertainty; while studies are usually conducted to show how differences in simulated discharge can be as a consequence of changing precipitation input, they rarely look at how much improvement of the model performance would be possible by using different but plausible precipitation (Bárdossy et al., 2020, 2022).

Because precipitation forcing data plays a crucial role in rainfall-runoff modelling, several methods (Yumnam et al., 2022) have been suggested for correcting precipitation data. These range from the use of storm multipliers (Sun and Bertrand-Krajewski, 2013) to station-wise correction of data using a gauge-based precipitation network (Cornes et al., 2018). However, gauge-based methods require a sufficient number of weather stations (Agarwal et al., 2020), which is often not the case for most regions around the world. As seen from previous experience, the observation network is too sparse even in data rich regions, and the majority of high-impact rainstorms are simply not observed (Borga et al., 2008). This is particularly true for flash floods in response to convective storm activity (Manoj J et al., 2024; Meyer et al., 2022; Villinger et al., 2022) and well related to the classical "Predictions in Ungauged Basins - PUB problem" (Sivapalan et al., 2003). To overcome this problem, and in line with Kirchner (2009) on "doing hydrology backwards", this paper explores options for inverse estimation of precipitation using the information contained in observed streamflow. The goal is to determine whether inverse estimation at the catchment scale can refine precipitation estimates from reanalysis products, ensuring they are hydrologically consistent, especially for extreme events.

While the classical "forward rainfall–runoff generation problem" has received considerable attention over various decades (Montanari et al., 2013; Sivapalan et al., 2003), a smaller subset of studies (Brocca et al., 2013; Kirchner, 2009; Kretzschmar et al., 2014; Krier et al., 2012; Teuling et al., 2010) has investigated the feasibility of tackling the inverse problem. Kirchner (2009) reported an early and successful attempt to infer catchment average rainfall and evaporation time series from streamflow fluctuations and inspired several investigations examining the advantages and limitations of doing "hydrology backwards" in diverse catchments (Krier et al., 2012; Teuling et al., 2010). Although

these studies have established a robust mathematical foundation for addressing the inverse hydrological problem, they were limited to smaller, well-monitored research catchments. This raises questions about the applicability of this approach to larger catchments as well as to smaller, non-experimental ones.

Note that inversions of the catchment water balance are inherently ill-posed, making it near impossible to find a unique solution (Bishop, 2006). Adopting the concept of micro- and macro-states from statistical mechanics (Zehe and Blöschl, 2004), we argue that the exact micro-state, i.e. the "true" space–time pattern of precipitation in the catchment, is neither uniquely identifiable nor observable. Yet, we conjecture that streamflow data (being an integral response from a potentially large and heterogeneous data) can reduce the uncertainty associated with this process, because it provides valuable information on antecedent precipitation and the current state of the catchment. As streamflow remains a non-linear convolution of the catchment-average precipitation, we propose that machine learning is well suited to this problem. Deep learning has recently fertilised almost all fields of the natural sciences and engineering, showing great promise in solving a wide range of inverse problems, especially those related to imaging (Ongie et al., 2020). It has also been argued that such models can provide meaningful and general benchmarks for hypothesis testing (Klotz et al., 2022; Nearing and Gupta, 2015) and afford powerful avenues for generalisation using large datasets (Loritz et al., 2024b).

The overall objective of this study is to 'do hydrology backwards' using regional-scale long short-term memory (LSTM) network ensemble models trained on large-scale hydrological datasets. While ERA5 Land (Table 4.1: Muñoz-Sabater et al., 2021 has well-documented issues in representing the driving precipitation estimates for specific event scales (Essou et al., 2016; Manoj J et al., 2024), recent studies (Bandhauer et al., 2022; Goteti and Famiglietti, 2024) have shown that they hold considerable promise to tackle the "Predictions in Ungauged Basins - PUB problem". This makes it an ideal test candidate for an inverse correction using streamflow and observational precipitation estimates over the same region (E-OBS: Cornes et al., 2018). The underlying research question is, "How much information about the catchment-average precipitation is effectively encoded in the variability of the streamflow time series observed at the outlet?" To answer this question, we first look at the performance gain in using discharge for predicting precipitation by focusing on days with higher precipitation magnitudes and then investigate whether the approach can accurately replicate the spatial characteristics of the original observational dataset (by looking at various time series measures) across European catchments for an unseen testing period.

We then examine how the inverse model performs when moving to much smaller (50–200 km^2: Table 4.2) out-of-sample catchments. Here, we compare (using the event runoff coefficients) LSTM-based inverse estimates during flood events to the original reanalysis product (ERA5 Land) and rain gauge-based observational estimates over the same region (E-OBS). Finally, we use a conceptual hydrological model (HBV: Bergström and Forsman, 1973) and a process-based model (CATFLOW: Zehe et al., 2001) to assess the quality of the precipitation estimates for forward modelling of streamflow and soil moisture dynamics, respectively.

4.2 DATA AND METHODS

4.2.1 *Model Configuration*

LSTMs (Hochreiter, 1998) are a special type of recurrent neural network that makes use of cell states and so-called "gates" to control the information flow through the network. The LSTM model used in this study extends upon the work of Kratzert et al. (2018) and Acuña Espinoza et al. (2024). The LSTM architecture, which is commonly used for streamflow simulation in hydrology (Kratzert et al., 2018) uses a sequence of meteorological variables, such as precipitation and temperature as dynamic inputs, along with catchment attributes as static features, to predict the corresponding streamflow. In our setting, to establish an inverse model, we use the same general model architecture as in previous studies (Acuña Espinoza et al., 2024; Loritz et al., 2024a). The key difference is that future streamflow is now used along with other dynamic and static data as inputs (Table C.1 in Appendix C.1) in order to estimate the precipitation forcings of the catchments. To account for the time lag between precipitation and discharge response observed at the catchment outlet, the model was provided with 7 d lead time series for discharge.We explored ranges of hyperparameter settings on a smaller subset of the training dataset to establish relatively stable hyperparameter configurations (Fig. C.2 in the Appendix C.4), finally setting them according to Acuña Espinoza et al. (2024) with a reduced number (5) of training epochs. Table C.2 in Appendix C.1 indicates the values used for the LSTM network hyperparameters. Mean squared error was used as the training loss function. In accordance with standard practices in the deep learning community, we utilise an ensemble network for LSTM predictions. In all cases, three individual LSTM models (with different initialisation seeds) were trained, and we present the mean predictions for the remainder of this paper.

The codes for model building and training can be found online (Manoj J, 2025). The LSTM was trained as a regional model (single

network trained on all available catchments) based on the openly available datasets detailed in the next section (Section 4.2.1). For forward hydrological modelling using the inversely-generated precipitation timeseries estimates, we use two hydrological models (Appendix C.2) - the lumped conceptual HBV model (Hydrologiska Byråns Vattenbalansavdelning: Bergström and Forsman, 1973) and the spatially distributed process-based CATFLOW model (Zehe et al., 2001).

4.2.2 *Datasets*

This study utilized the Caravan dataset (Kratzert et al., 2023) to investigate our hypothesis regarding the inverse identifiability of precipitation from information about discharge dynamics. We trained our model on European catchments from the GRDC-Caravan (Färber et al., 2023) community extension and the original Caravan dataset, which includes catchments from CAMELS-GB (Coxon et al., 2020). The Caravan dataset uses the ERA5 Land (Muñoz-Sabater et al., 2021) as the primary meteorological forcing, while the catchment attributes include data from HydroATLAS (Linke et al., 2019). The discharge data is tapped from relevant state and national authorities and is accessible as open datasets. The observational E-OBS precipitation product (v31.0 – Cornes et al. 2018), which uses the station network of the European Climate Assessment & Dataset (ECA&D) project, was used as the training target for the model runs. Figure C.3 in the Appendix depicts the study catchments (1800 in total) in the training dataset.

We chose a training period of around 25 years between 1 October 1980 to 30 September 2005. Following the best practices in data-based modelling, the model was tested on an unseen testing period between 2006 and 2020 (2015 for CAMELS-GB catchments due to data unavailability). To investigate its generalizability across scales, we also tested the model on four catchments (Figs. C.4 and C.5 in the Supplement) that were not included in the original training set (Sect. 4.2.3). For the out-of-sample test, we made use of data from the Caravan Spain (Casado Rodríguez, 2023) and Caravan Switzerland (Höge et al., 2023) extensions, in addition to data from local data providers in Germany (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg – LUBW) and Luxembourg (Nijzink et al., 2025). To validate the inversely generated precipitation (Sect. 4.2.3) during forward modeling, we conducted hydrological model simulations in the Elsenz Schwarzbach and Lippe catchments (Fig. C.6 in Appendix C.4). Table 4.1 provides an overview of the main datasets used in this study, detailing their spatial and temporal resolutions, as well as their sources.

Table 4.1: Brief overview of the datasets used in this study, including their spatial and temporal resolution.

Dataset	Type & Source	Spatial Resolution	Temporal Resolution	Details
Caravan	Hydrometeorological dataset (Kratzert et al., 2023)	Catchment scale	Daily	Open community dataset that includes catchment forcing data and attributes along with streamflow.
GRDC-Caravan	Hydrometeorological dataset (Färber et al., 2023)	Catchment scale	Daily	Community extension to the Caravan dataset, incorporating data from the Global Runoff Data Centre (GRDC).
ERA5-Land	Reanalysis product (Muñoz-Sabater et al., 2021)	$0.1^\circ \times 0.1^\circ$	Hourly (aggregated to daily)	Reanalysis product produced by replaying the land component of ERA5 climate reanalysis.
E-OBS	Gridded observational precipitation product (Cornes et al., 2018)	$0.25^\circ \times 0.25^\circ$	Daily	Interpolated observational precipitation product utilizing the station network from the European Climate Assessment & Dataset (ECA&D) project.
Caravan Spain	Hydrometeorological dataset (Casado Rodríguez, 2023)	Catchment scale	Daily	Community extension to the Caravan dataset, incorporating data from Spain.
Caravan Switzerland	Hydrometeorological dataset (Höge et al., 2023)	Catchment scale	Daily	Community extension to the Caravan dataset, incorporating data from CAMELS-CH catchments.
Caravan Germany	Hydrometeorological dataset (Dolich et al., 2025)	Catchment scale	Daily	Community extension to the Caravan dataset, incorporating data from CAMELS-DE catchments.
MERRA-2	Reanalysis product (Gelaro et al., 2017)	$0.625^\circ \times 0.5^\circ$	Hourly (aggregated to daily)	Global atmospheric reanalysis by NASA Global Modeling and Assimilation Office (GMAO) using the Goddard Earth Observing System Model (GEOS).
GLDAS-2.2	Reanalysis product (Li et al., 2019)	$0.25^\circ \times 0.25^\circ$	Daily	NASA Global Land Data Assimilation System model outputs with data assimilation for the Gravity Recovery and Climate Experiment (GRACE-DA).

4.2.3 *Experimental Design*

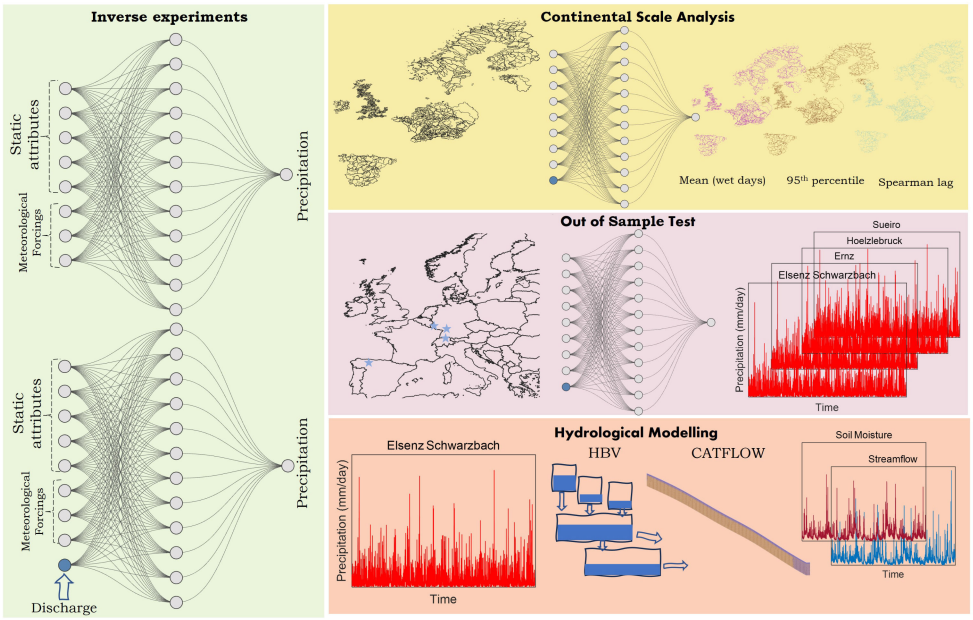

Figure 4.1: Schematic representation of our methodological approach. Each rectangular panel indicates different stages of our workflow. Initially, we train two LSTM ensemble networks to predict catchment average precipitation through inverse experiments. The trained models are then then utilized for a continental-scale analysis before being used for out-of-sample testing. Finally, a validation exercise for the inversely generated precipitation is conducted using various hydrological models.

4.2.3.1 Exploring information about precipitation encoded in streamflow

To shed light on the value of discharge for inversely predicting precipitation, we conducted a virtual experiment (Fig. 4.1) in which two LSTM ensemble models (Tables C.1 and C.2 in Appendix C.1) were trained using the same catchments and training period. The first model (*without_discharge*) used only ERA5 Land meteorological time series (total_precipitation, air temperature, solar and thermal radiation) and static attributes (area, ele_mt_sav, frac_snow, pet_mm_syr: Kratzert et al. (2023)), while the second model (*with_discharge*) included lagged discharge as an additional input variable. Both models were trained to predict daily catchment average precipitation sums from the observational EOBS product. Therefore, we only deal with spatially averaged timeseries for precipitation, assuming that these values represent the actual precipitation over the entire catchment.

We then used both the trained regional-scale models (*with_discharge* and *without_discharge*) to predict the precipitation time series inversely

for all the test catchments over the unseen testing period and evaluated (Appendix C.3) those using the mean wet day precipitation (MWD) – mm day^{-1}, 95th percentile limit (R95P) – mm day^{-1}, and Spearman autocorrelation values (SL) for each catchment, and then compared them to the values from ERA5 Land (the reanalysis product we want to improve) and E-OBS (observational product used as training target) at the continental scale.

4.2.3.2 *Out of sample precipitation inversions and their quality*

We further tested the feasibility of knowledge transfer to out-of-sample catchments and used the same regional-scale models (with_discharge and without_discharge) to inversely predict the intensity of driving rainstorms for selected flood events in four hydro-climatically diverse and much smaller catchments (not included in the original training dataset). These catchments (Table 4.2 and Figs. C.4 and C.55) were chosen based on the severity of the flooding and on the apparent inability of ERA5 Land forcings to accurately represent the storms that triggered the flood events.

Table 4.2: Attributes for the four catchments used for out-of-sample testing.

Catchment	Country	Area (km^2)	Mean precipitation (mm day^{-1})	Mean potential evapotranspiration (mm year^{-1})	Mean elevation (m)
Elsenz- Schwarzbach	Germany	196.5	2.51	812.85	246.7
Ernz	Luxembourg	69.3	2.31	724.04	345.5
Sueiro	Spain	132.5	3.31	873	381
Hoelzlebruck	Germany	47.1	4.14	658	980

4.2.3.3 *The potential of inverted precipitation for forward modelling*

To evaluate the value of generated precipitation data for forward modeling of streamflow, we calibrated the HBV conceptual hydrological model (Bergström and Forsman, 1973) over the Elsenz Schwarzbach (Manoj J et al., 2024) and Lippe (camelsde_DEA11130: Loritz et al., 2024a) catchments (Fig. C.6) using both the original ERA5 Land and the *with_discharge* LSTM-generated precipitation timeseries and compared the evaluation period performance of both model versions (Table C.3 in Appendix C.2). The HBV model (Appendix C.2) used in this paper requires precipitation (ERA5 Land/LSTM simulated), potential evapotranspiration, and air temperature as inputs. We follow the recommendations of Clerc-Schwarzenbach et al. (2024), similar to that of Loritz et al. (2024a), for the calculation of potential evapotranspiration, and use the temperature-based Hargreaves formula detailed by Adam et al. (2006).

Complementary to streamflow modelling, the performance of a hydrological model can also be judged by how well it replicates the catchment dynamics of a region. Soil moisture is a key variable controlling the partitioning of net radiation into sensible and latent heat (Seneviratne et al., 2010) or overland flow during a rainstorm (Zehe and Blöschl, 2004). We thus used each precipitation estimate (with_discharge LSTM and ERA5 Land) to run the process-based hillslope scale model CATFLOW (Appendix C.2), using a setup from Manoj J et al. (2024) used for uncalibrated predictions of local floods. Here, we focused on one of the headwater subcatchments (Catchment W32 in Fig. C.6) within the Elsenz Schwarzbach. The model simulated (Table C.3) the period from 1 January 2008 to 31 December 2015 using each of the ERA5 Land and with_discharge LSTM precipitation estimates, and the corresponding spatially averaged soil moisture states were compared against several soil moisture reanalysis products (Table 4.1: due to the unavailability of observed data). These include (a) ERA5 Land: Muñoz-Sabater et al., 2021, (b) GLDAS (NASA Global Land Data Assimilation System, GLDAS-2.2 GRACE DA: Li et al., 2019) and (c) MERRA (Modern-Era Retrospective analysis for Research and Applications version 2 – tavg1_2d_lnd_Nx: Gelaro et al., 2017).

4.3 RESULTS

4.3.1 *The information contained in streamflow about precipitation*

Figure 4.2 shows violin plots displaying the pairwise difference in the mean performance of the two LSTM models (Fig. C.1 in Appendix C.1) over the catchments (1800) in the test dataset for varying precipitation amounts (All days, days with daily precipitation greater than 1 mm and days with daily precipitation greater than 5 mm). Each point denotes the difference in NSE (Appendix C.3) for individual catchments while making predictions using the *with_discharge* model compared to the *without_discharge* model. A marked shift towards higher positive differences indicates that the model "*with_discharge*" has higher NSE values than the model "*without_discharge*". This holds true not only on average but also with respect to the best-performing catchments The median NSE metric value (Nash and Sutcliffe, 1970) for the regional LSTM model (considering entire time series) across the study catchments is about 13% higher when discharge is used as an additional predictor than when it is not. However, it is also observed that discharge information has worsened the performance in a few cases, likely due to the poor quality of streamflow data in these catchments. Analysing the performance improvement achieved by focusing on days with increasing precipitation amounts reveals that the gains are considerably greater on days with higher recorded precipitation

(increase in median NSE value of about 29% from 13% as we look only at days with more than 5 mm precipitation). This largely answers our main research question and shows that the variability of discharge as measured in the catchment outlets holds enormous information about the driving storms over the entire catchment area. Consequently, we can utilise this information by applying a data-driven LSTM network. The information gain is naturally higher for more extreme precipitation events, as average streamflow conditions do not provide much information about the catchment scale precipitation.

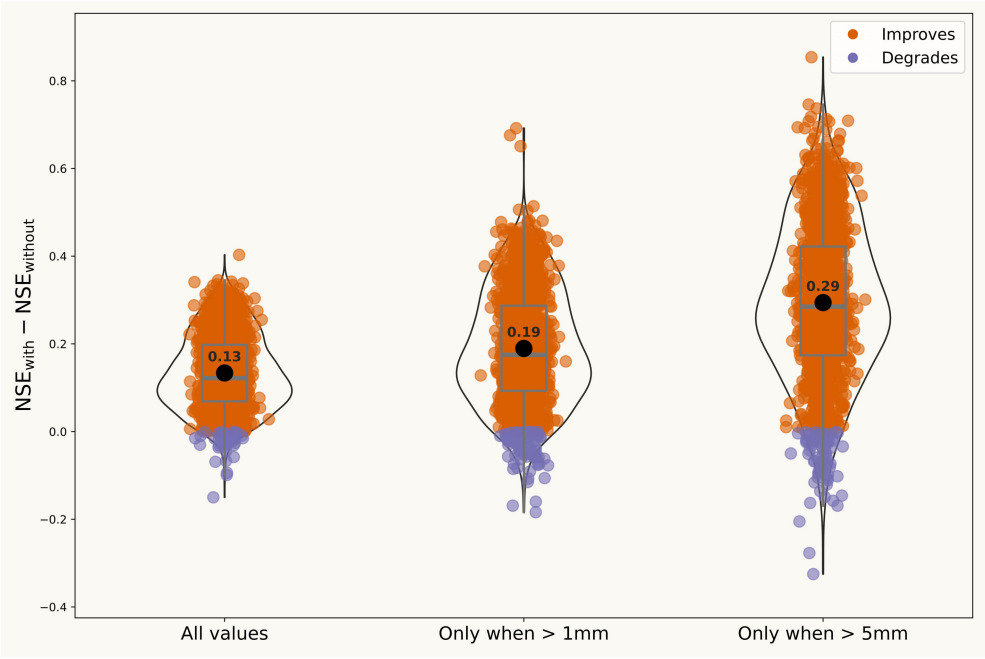

Figure 4.2: Comparison of performance gain for the *with_discharge* vs. *without_discharge* models in NSE for different precipitation amounts. The first violin plot illustrates the average improvement across all days in the testing period. The second and third plots display the mean performance gains over the catchments, specifically focusing on days where precipitation exceeded 1 and 5 mm, respectively

4.3.2 *Unraveling the Continental Scale Characteristics*

To examine the characteristics of the simulated time series from the *with_discharge* and *without_discharge* models over the testing period in detail, we computed three timeseries measures (Appendix C.3) namely mean wet day precipitation (MWD) – mm day^{-1}, 95th percentile limit (R95P) – mm day^{-1}, and Spearman autocorrelation values (SL) across all the catchments.

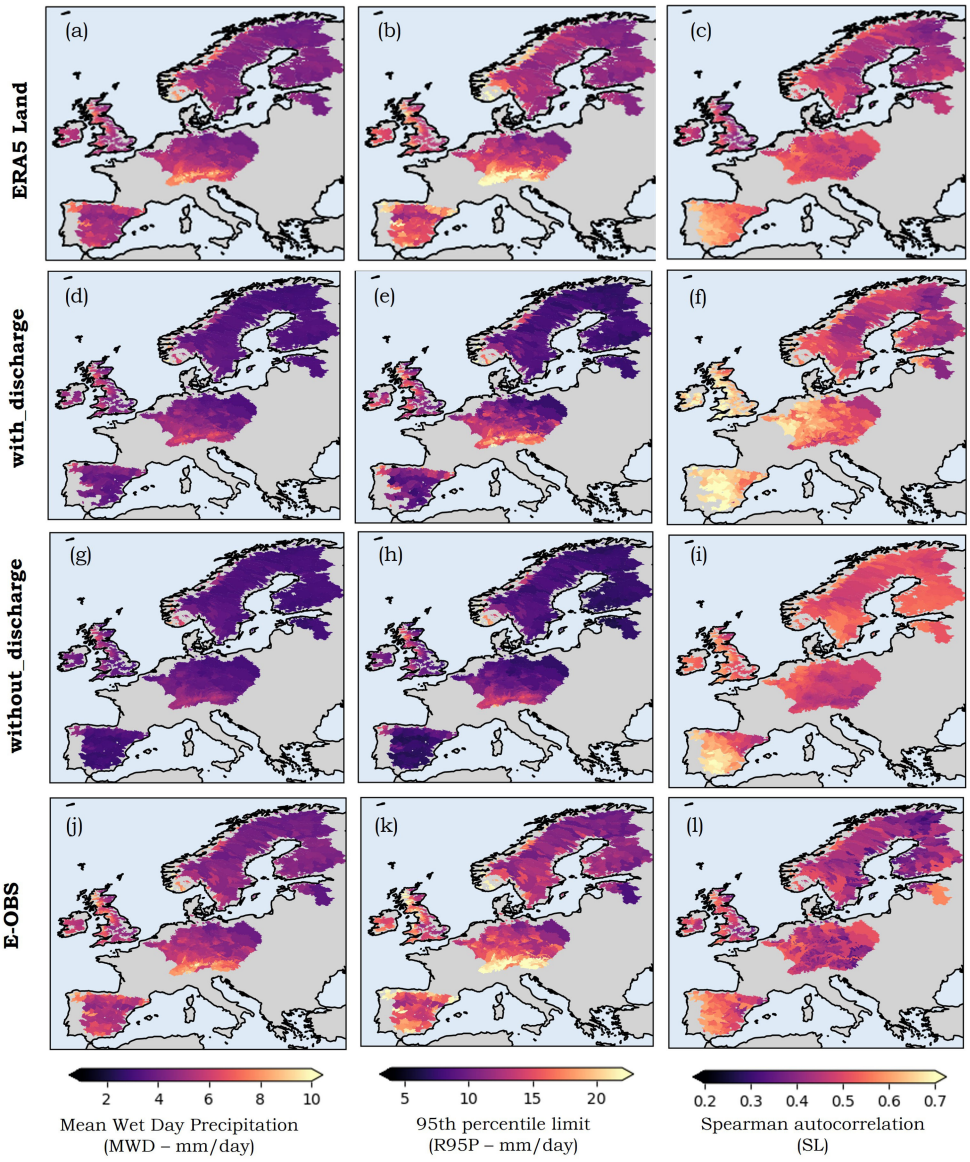

Figure 4.3: The spatial patterns of the different time series metrics (Appendix C.3) mean wet day precipitation (MWD) – mm day^{-1}, 95th percentile limit (R95P) – mm day^{-1}, and Spearman autocorrelation values (SL) over the study catchments for the different precipitation estimates – ERA5 Land (top row): (a–c), *with_discharge* LSTM model (second row): (d–f), *without_discharge* LSTM model (third row): (g–i) and E-OBS (bottom row): (j–l) from 2006 to 2020 (2015 for CAMELS-GB catchments).

The continental-scale analysis reveals distinct patterns for the major European climatic regions. The spatial patterns for the mean wet day precipitation (Fig. 4.3d: MWD) obtained using the *with_discharge* LSTM model are well aligned to the ones from ERA5 Land (Fig. 4.3a) and EOBS (Fig. 4.3j). Higher daily average values are observed towards the Alps, the Carpathian Mountain ranges, and the coast of Norway, consistent with the climatology of these regions. In addition,

we also see that the ERA5 Land largely matches the precipitation field's characteristics (wet day mean and 95th percentile limit) as in the observational E-OBS product. This indicates that both products contain complementary information at such larger spatial scales.

For the 95th percentile of wet days (R95P), we again see a robust representation of the spatial differences, along with an underestimation of the magnitudes (Fig. 4.3b–k). The Spearman autocorrelation coefficient values (SL: Fig. 4.3c–l) indicate that while the models underestimate the mean and 95th percentile limits, they overestimate the autocorrelation (which indicates the persistence in the precipitation time series) compared to the ERA5 Land and EOBS time series. Comparing the *with_discharge* and *without_discharge* models for MWD and R95P, we see that the addition of discharge information reduces the underestimation errors over the continental scale.

The higher autocorrelation values for the *with_discharge* (Fig. 4.3f) and *without_discharge* (Fig. 4.3i) may arise from model products incorporating catchment persistence, unlike the gridded observational E-OBS data. In the case of the *with_discharge* LSTM model, the higher values are likely due to the inclusion of strongly auto correlated streamflow data, which adds redundancy or a longer memory

4.3.3 *Out of sample predictions*

Figure 4.4 shows predicted event precipitation values over time for the four out-of-sample catchments. Again, we compare the inversely modelled values (*with_discharge* and *without_discharge*) to the ERA5 Land (the reanalysis product to be corrected) and the gauge-based E-OBS product (our training target). Table 4.3 lists the peak storm precipitation values reported by the different products along with the recorded flood values (both normalised to the catchment area in $\mathrm{mm\,day^{-1}}$). Also shown are the storm runoff coefficients for the respective events based on the different precipitation estimates and discharge data.

Figure 4.4a represents the summer flood in June 2016 in the Elsenz Schwarzbach catchment in Germany. This annual flood event was triggered by a series of convective rainfall events caused by persistent atmospheric conditions in Germany during the summer of 2016. Localised rainfall totals exceeded 100 mm in some catchments (Bronstert et al., 2018), triggering widespread flash floods. Our previous work (Manoj J et al., 2024) indicated that the ERA5 Land reanalysis product could not accurately replicate the characteristics of the convective storm that caused this annual flood event over the Elsenz Schwarzbach catchment. The *with_discharge* LSTM simulated precipitation for this

event was higher than the values reported by both ERA5 Land and the training target EOBS, while the *without_discharge* model performed even worse than ERA5 Land.

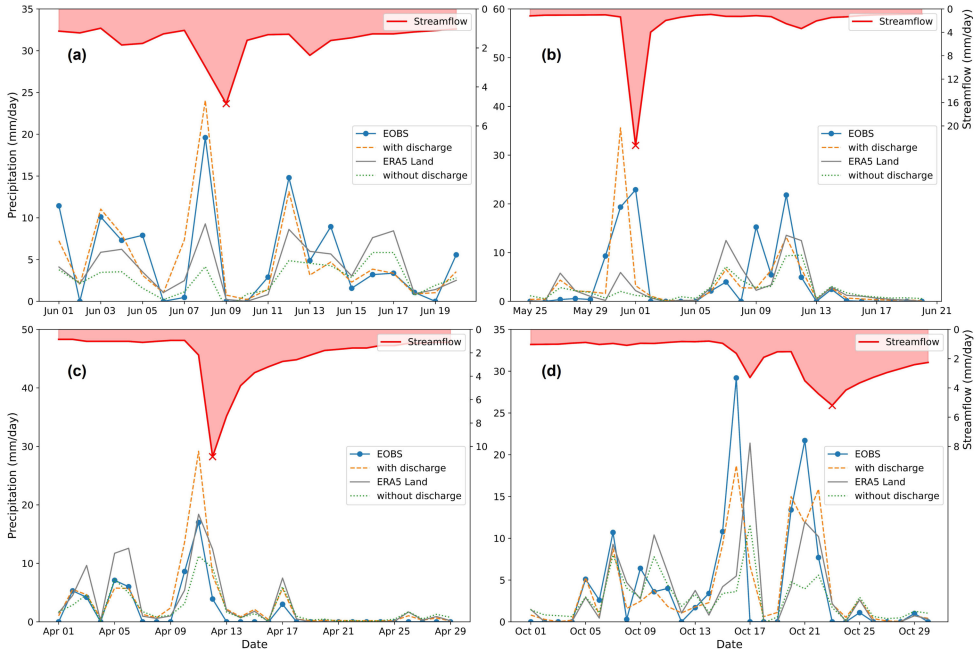

Figure 4.4: Precipitation estimates for flood events at four out-of-sample catchments: (a) Elsenz Schwarzbach, (b) Ernz, (c) Sueiro, and (d) Hoelzlebruck. The red line represents the observed daily streamflow, with a cross marking the day of the flood event. The orange curve indicates the precipitation predicted by the *with_discharge* LSTM model, while the green curve shows the precipitation predicted by the *without_discharge* model. The blue line reflects the original gauge-based EOBS time series, and the grey line represents the estimate from the ERA5 Land.

Table 4.3: Event characteristics (storm volume and runoff coefficients) for the four out of sample catchments.

Event Characteristics		Elsenz-Schwarzbach	Ernz	Sueiro	Hoelzlebruck
Precipitation (mm)	ERA5 Land	12.51	9.60	41.81	32.12
	with_discharge	32.79	42.75	58.53	50.85
	without_discharge	4.92	6.20	29.46	22.92
	E-OBS	20.07	51.72	29.50	44.90
Discharge (mm)		5.98	26.88	23.39	19.14
Runoff Coefficient	ERA5 Land	0.48	2.80	0.56	0.60
	with_discharge	0.18	0.63	0.40	0.38
	without_discharge	1.21	4.34	0.79	0.84
	E-OBS	0.30	0.52	0.79	0.43

A comparison of *with_discharge* LSTM-simulated precipitation values to radar estimates over the same region (Manoj J et al., 2024) revealed the estimates to be closer than those reported by the observational E-OBS product. The runoff coefficient (Table 4.3) for the event also decreased from 48% (ERA5 Land) to around 18% (*with_discharge*), which is consistent with estimates from Manoj J et al. (2024). The *with_discharge* LSTM model was also able to represent the second storm peak more accurately than ERA5 Land.

Next, the *with_discharge* model was used to estimate precipitation for another convective episode over the Ernz Catchment in Luxembourg (Fig. 4.4b) in the summer of 2018. Once again, we observed that the model overestimated the peak precipitation compared to the observational EOBS product used for training. However, the model benefited from integrating improved event timing information from ERA5 Land, which helped reduce timing errors compared to EOBS. Essentially, the model combined information from both ERA5 Land and discharge to produce a storm estimate that was more consistent with the hydrology of the flood, taking into account both the volume and timing of the event, than the observational EOBS product. In contrast, the *without_discharge* model again performed poorly for this event, resulting in an unrealistically high runoff coefficient of 4.34 (Table 4.3).

In the third catchment (Sueiro: camelses_1414 from Caravan Spain extension), the *with_discharge* estimate for storm forcing was higher than ERA5 Land and E-OBS (Fig. 4.4c). The corresponding runoff coefficients underline the reliability of the storm prediction from *with_discharge* (0.40) compared to E-OBS (0.79). For the Sueiro catchment (camelses_1414), the closest observational station is located more than 60 km away (Fig. C.5), explaining why the EOBS performs rather poorly in representing the driving forcings for the summer flood event.

In the Hoelzlebruck catchment (camelsch_4003 from Caravan Switzerland extension), two consecutive events occurred in October 2014. ERA5 Land was better than the *with_discharge* LSTM model in capturing the initial event magnitude, while the *with_discharge* model had better timing accuracy for the events (Fig. 4.4d). For the second event, which was the annual flood event, the *with_discharge* model, which incorporated streamflow information, was again able to reduce the relative errors in storm volume (Table 4.3). The *without_discharge* model showed the same timing error as ERA5 Land for the first storm; however, introducing discharge allowed the model to correct the timing bias.

4.3.4 *Forward Hydrological Modelling*

The precipitation estimates generated by the *with_discharge* LSTM model were then used to run classical hydrological models (HBV and CATFLOW: Table C.3) in a forward manner. To address the question of performance in differently sized basins, we run the conceptual HBV model in two catchments (Fig. C.6) – Elsenz Schwarzbach (Fig. 4.5: 196.5 km^2) and Lippe (Fig. 4.6: 3366.3 km^2).

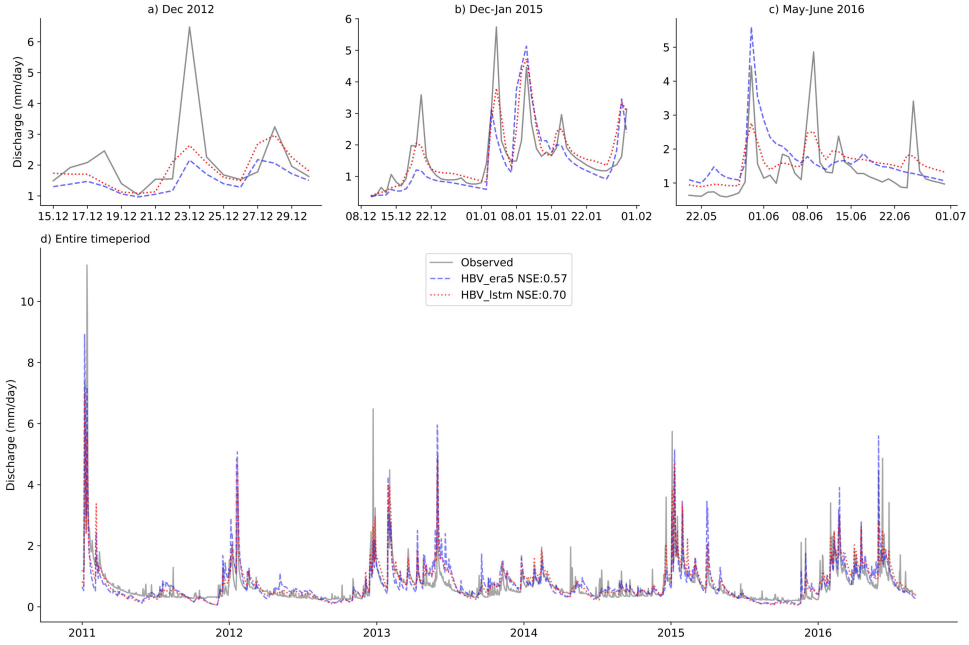

Figure 4.5: Observed (grey line) and simulated runoff (using the HBV model) at the Elsenz Schwarzbach catchment. The blue line denotes the streamflow simulated using the ERA5 Land precipitation product, while the red curve depicts the simulations using the inversely estimated precipitation obtained using the *with_discharge* LSTM model. Moreover, three rainfall–runoff events are highlighted and displayed separately.

Figure 4.5d illustrates that the HBV model, which utilized the inverted precipitation estimates, performed better (NSE = 0.70) during the evaluation period over Elsenz Schwarzbach compared to the model driven by the ERA5 Land (NSE = 0.57). To gain a better understanding of the differences between the models, we visually examined the results for three individual flood events, as shown in Fig. 4.5a–c. During the winter flood of December 2012 (23 Dec 2012, Fig. 4.5a), the model driven by ERA5 Land significantly underestimated both the peak and the volume of the flood event. When using *with_discharge* -simulated precipitation, the relative peak error decreased slightly Similarly, the model runs using *with_discharge* precipitation more accurately captured the post-event conditions (28 Dec 2012). In the winter of 2015 (Fig. 4.5b), the model using *with_discharge* precipitation again

demonstrated better performance. The model could more accurately represent the smaller flood peaks before the larger floods. This aligns with findings from other studies (Berghuijs et al., 2019; Manoj J et al., 2023) that emphasize the importance of initial conditions for floods across Europe.

During the convective summer storm event in June 2016 (Fig. 4.5c), neither model run successfully captured the flashy runoff response. Although the model that utilized ERA5 Land input predicted an earlier flood event in May 2016 with an overestimation bias, it did not accurately depict the dynamics of the annual flood event occurring a few days later. In contrast, the model with LSTM-generated precipitation (*with_discharge*) generally performed better in capturing both the magnitude and volume of the smaller storm peaks as well as the annual flood event on June 8, 2016. For the larger Lippe catchment, we

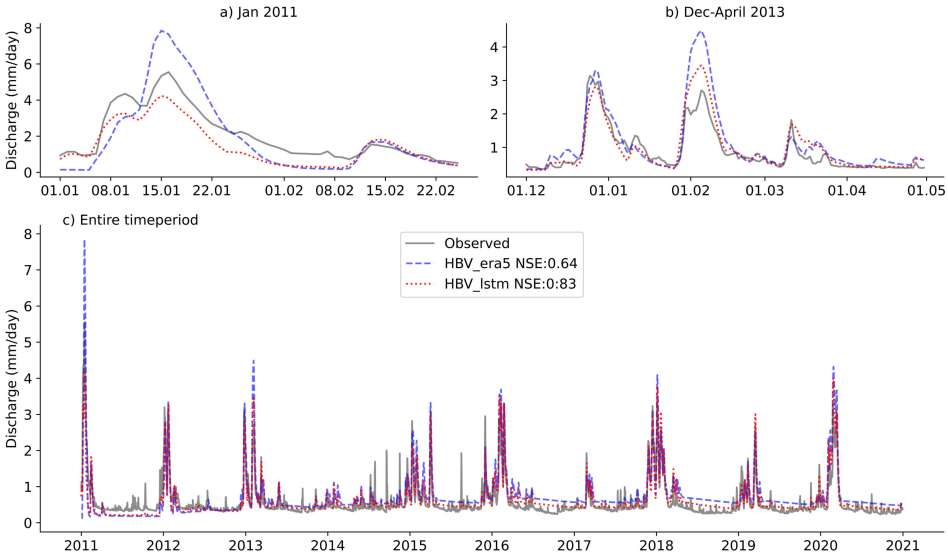

Figure 4.6: Observed (grey line) and simulated runoff (using the HBV model) at the Lippe catchment. The blue line denotes the streamflow simulated using the ERA5 Land precipitation product, while the red curve depicts the simulations using the inversely estimated precipitation obtained using the *with_discharge* LSTM model. Moreover, two rainfall–runoff events are highlighted and displayed separately.

again saw improved mean performance for the run with inversely generated precipitation (Fig. 4.6c). For the winter flood of 2011 (Fig. 4.6a), the HBV model, which used inversely generated precipitation, better matched the observed streamflow dynamics, whereas the ERA5 Land run exhibited significant overestimation errors. The inversely generated precipitation estimates again improved HBV model performance for replicating the discharge dynamics during the floods in December 2012 and February 2013 (Fig. 4.6b).

To understand the evolution of soil moisture dynamics while using the *with_discharge* LSTM-based precipitation estimates in physically based models, we conducted a hillslope-scale CATFLOW model simulation (Loritz et al., 2017; Manoj J et al., 2024) in one of the headwater catchments in Elsenz Schwarzbach (ERA5 Land vs *with_discharge* LSTM). The pairwise correlation values, as shown in Fig. 4.7, indicate that the use of the LSTM-based precipitation estimates does not lead to a loss of information regarding soil moisture dynamics in the region. In fact, we observe a slight increase in correlation when comparing the inversely derived precipitation estimates (referred to as CATFLOW_lstm) to MERRA and GLDAS (Table 4.1), in contrast with the correlation obtained for the run with ERA5 Land (referred to as CATFLOW_era5). As expected, the correlation value for the ERA5 Land run is slightly higher when assessed against soil moisture from the same ERA5 Land dataset, which may be attributed to model biases arising from using the same dataset for both precipitation and soil moisture.

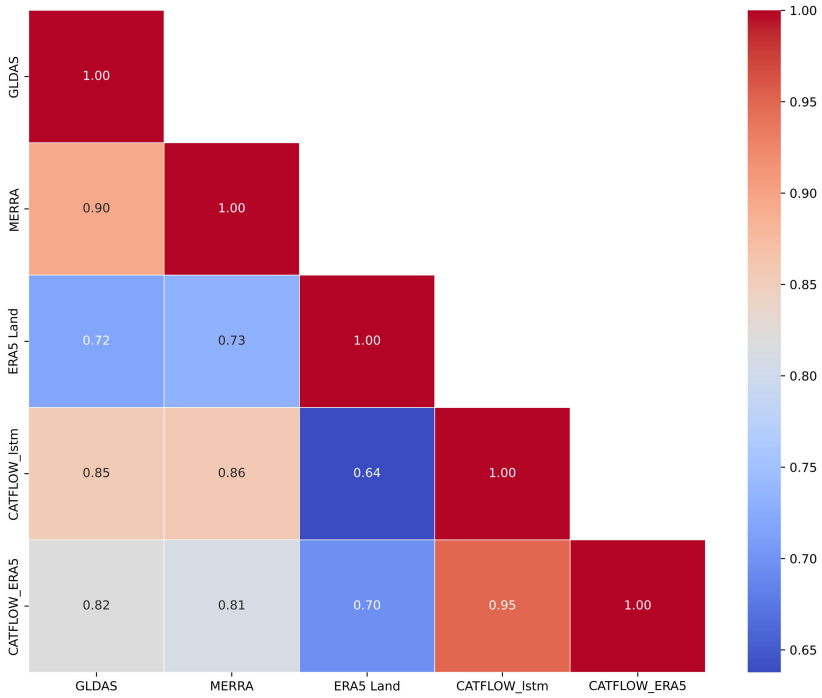

Figure 4.7: Correlation matrix plot illustrating the pairwise correlations between the different soil moisture estimates – GLDAS (NASA Global Land Data Assimilation System, GLDAS-2.2 GRACE DA: Li et al., 2019), MERRA (Modern-Era Retrospective analysis for Research and Applications version 2 – tavg1_2d_lnd_Nx: Gelaro et al. (2017)), ERA5 Land: Muñoz-Sabater et al., 2021, CATFLOW_lstm: model run using inversely estimated precipitation estimate from the LSTM model and CATFLOW_ERA5: model run using precipitation estimate from ERA5 Land product.

4.4 DISCUSSION

4.4.1 *Improved precipitation estimation using discharge*

Overall, our study reiterates that streamflow data can be exploited to obtain useful information about the nature of catchment-scale precipitation amounts: we can thus invert the cause using the effect as input to an LSTM. This is in line with, and steps beyond, previous studies (Brocca et al., 2013; Kirchner, 2009; Kretzschmar et al., 2014; Krier et al., 2012; Teuling et al., 2010) that explored the possibility of doing hydrology backwards using experimental catchments. Here, we successfully expanded this idea to large samples, cutting across the wide range of hydro-climatic conditions that characterise Europe. We found a largely 'normal' distribution of performance, with a few outliers, the latter indicating possible poor quality of discharge data.

Although ERA5 Land precipitation has known uncertainties, it provides continuous global spatial and temporal coverage, making it a useful training dataset. Our goal was not to generate a fully independent dataset but to improve the ERA5 Land precipitation estimates using the additional streamflow information. Reanalysis data, by definition, are a mix of observations and past short-range weather forecasts rerun with modern weather forecasting models. Different data assimilation methods are then employed (Li et al., 2019). The inversion technique could be used as another final layer of post-processing (using the LSTM in this case) for the model outputs to ensure that the final product is more consistent with the variabilities observed in the discharge record.

One limitation of our approach is that the LSTM model tends to underestimate the timeseries measures (MWD and R95P) at the continental scale. The LSTM's architecture is known to have a theoretical saturation limit, leading to the underestimation of some of the peak storm events. This so called 'saturation problem' (Baste et al., 2025; Chen and Chang, 1996) implies that irrespective of the input series, the predicted values can never exceed a theoretical limit (which is established during the training phase). Furthermore, the LSTM model looks for recurrence in patterns and mean conditions. This means that it can indeed account for consistent baseflow dynamics (as also indicated by analysis over the larger Lippe catchment, Figure 4.6). In extreme floods (Merz et al., 2021), the relative contributions of each component can vary significantly, depending on various factors such as the antecedent conditions of the catchment area. The model likely struggles to learn this variability while attempting to invert and obtain the driving precipitation values. Given the non-linear nature of the inverse problem, there are always multiple possible solutions. Since

the model is trained to minimize the mean squared error (Gupta et al., 2009), it may also tend to consistently predict lower values (on peaks) to effectively reduce the average error during training.

It is also important to acknowledge that "true" precipitation estimates don't exist at the catchment scale. We obtain estimates of forcing precipitation at such scales (with considerable uncertainty) by interpolating station data (e.g. EOBS) or averaging gridded data from reanalysis/remote sensing products (e.g. ERA5 Land).

In our out-of-sample simulations, we observed that the LSTM model, which included additional discharge information, overestimated the peak values reported by the observational product used for training. While such an overestimation is typically considered an artifact of imperfect model training and viewed as statistical white noise, we believe that the consistent overestimation of peaks in three out of the four catchments suggests that the LSTM model, trained globally on larger catchments with smaller observational uncertainties, is capable of learning the rainfall–runoff relationship and can adjust for observational errors at the out-of-sample sites. Although the model was not specifically trained on the timing characteristics of hydrological events, we found that it can still produce hydrologically consistent estimates for the time to peak for storms.

The performance comparison using the runoff coefficients was intended to provide insight into the feasibility of different precipitation estimates from a hydrological perspective. While we acknowledge the existence of even better regional products (e.g., HYRAS – German-Weather Service) for some of the study catchments compared to the continental scale EOBS, we believe that these various products should not be viewed as independent of one another. Instead, they contain complementary information as they represent the same physical truth i.e. precipitation occurring over a catchment, albeit with different uncertainties and errors. Although the EOBS data is only available over Europe, the trained model could be transferred to similar hydroclimatic regions worldwide that have discharge information to correct the globally available

4.4.2 *Catchment as a functional unit*

In the introduction, we argued that the catchment scale is crucial for improving our understanding of the factors that drive the water cycle and representing them more accurately in reanalysis products. Our findings across the four catchments highlight the benefit of using streamflow variations to rectify precipitation estimates. By leveraging the generalisation capabilities of the data-driven LSTM model, we

successfully transferred knowledge across different scales (Notably, only about 9% of the catchments in our training dataset had areas smaller than 100 km^2), indicating important implications for addressing the ever-evolving challenge of predictions in ungauged basins (PUB: Hrachowitz et al., 2013)

Although this approach can only be applied after the event has taken place, it has implications for generating coherent long-term statistical records for catchment forcings, which could be used for the design of small- to medium-purpose water resource projects. Employing daily precipitation sums from products like ERA5 Land and EOBS should ideally be a last resort for reproducing small-scale hydrological events, however, the scarcity of real-world data and the rarity of these events may sometimes necessitate a modelling decision to incorporate these coarser estimates. Using the streamflow fluctuations, it would be possible to identify localised rainfall cells or snowfall events that are poorly captured by traditional rain gauges (Kretzschmar et al., 2014). The approach also has potential for evaluating long-term rainfall estimates from Global Circulation Models for specific catchments using information about hydrological conditions (Fujihara et al., 2008).

While the LSTM-based precipitation estimates improved the representation of most events, there were still instances where the original ERA5 Land provided better accuracy for peak flood magnitudes (Fig. 4.5); this highlights the need for a blended approach that incorporates additional information rather than completely replacing one product with another. In regions around the world, the wealth of streamflow information remains underutilised in this aspect. For Germany alone (Loritz et al., 2024a), there are more than 1500 streamflow gauges, which represent a significantly higher representative area compared to precipitation stations.

The forward exercise using the HBV model showed that the precipitation estimates after inversion enhanced mean performance for streamflow simulation and helped improve the modelling of extreme individual floods. The ability to match the hydrograph differed between the different seasons. Compared to the storage-controlled winter floods (Dunne, 1978) , summer floods in these regions are usually driven by Hortonian flow (Horton, 1933) in response to high-intensity rainfall during convective storms. Previous studies (Kirchner, 2009; Krier et al., 2012) have discussed such storage-controlled dynamics and their impact on the inversion problem.

Previous experiences at the event scale (Beauchamp et al., 2013; Zehe and Blöschl, 2004) have also shown that inferring the antecedent soil moisture conditions remains a key challenge for accurate and

reliable flood simulations. By utilising the process-based CATFLOW model for soil moisture simulations in a small headwater catchment, we achieved high correlation values using the inverse precipitation estimate. This suggests that the approach can help represent the catchment's overall water dynamics and has the potential for reliable flood design estimations at the event scale, particularly in data-scarce regions.

4.4.3 *Limitations and Outlook*

It is important to stress that, as for any data-driven study, the results of our work are contingent on the quality of the training dataset. While we are aware of better regional products for individual countries, ERA5 Land provides consistent global coverage, and a permissive data sharing policy makes it one of the obvious choices for a continental scale modelling exercise. To evaluate the applicability of the commonly used LSTM network architecture, we decided to use the same architecture previously employed in hydrological studies instead of creating an experimental design with modified individual layers and training functions for inverse modelling. It is evident that exploring the impacts of different loss functions and deep learning model architectures like transformers would help advance the methodology discussed in this paper. This approach could also shed light on best-suited algorithms for the problem but is beyond the scope of the present work. The choice of Mean Squared Error (MSE) as the training function and Nash Sutcliffe Efficiency (NSE) as a performance metric is motivated by its success and applications in the forward problem (streamflow prediction), but this adds its own biases to the modelling exercise. In the present work, we tried to overcome this issue by relying less on the evaluation measure (NSE) and placing greater emphasis on the hydrological feasibility of the predictions (using the runoff coefficient). Additionally, we tried to complement this by calculating various other time series metrics commonly used in hydrometeorological studies. The four events for out-of-sample tests across various catchments were chosen based on the severity of the floods and ERA5 Land's inability to capture the characteristics of the driving storms. The choice of the hydrological models and calibration period also adds uncertainty to the forward simulations.

Our approach opens up many perspectives for future research. Transfer learning to data-scarce regions could help address the challenge of highly uncertain precipitation estimates in smaller catchments without precipitation gauges, improving hydrological modeling and the representation of extreme events such as convective storms, which are crucial for designing flood defense measures. Additionally, the inversion technique could serve as a final post-processing layer

for gridded reanalysis products, ensuring better consistency with discharge variability and enabling machine learning approaches to estimate spatial precipitation fields conditioned on discharge data (Bárdossy et al., 2020, 2022). Moreover, this methodology could be applied to reconstruct past floods by leveraging historical hydrological records, storm water level markings, and observational flood data (Bronstert et al., 2018; Seidel et al., 2009), providing valuable insights into the driving storms behind some of the devastating past flood events. The workflow could also be expanded for the generation of new precipitation products, merging multiple different precipitation sources alongside the streamflow inversion.

4.5 CONCLUSIONS

Our main hypothesis was supported by the findings, which demonstrated that discharge has unused potential and can be inversely assimilated to adjust precipitation estimates derived from reanalysis products, while machine learning models are key to expanding this effort to large data sets spanning the scale of entire continents. As expected, the performance gain in using discharge information was significantly higher for days with increasing precipitation amounts. Insights from the out-of-sample catchments provided valuable information about the applicability of our method for estimating flood forcings and the generalizability of the model. Additionally, we have shown that the inversely estimated precipitation estimates can improve forward modelling of both streamflow and soil moisture dynamics, illustrating how the information gained can be integrated into existing modelling strategies.

Part V

SYNTHESIS AND CONCLUSIONS

In this section, I provide a synthesis of my main findings and the conclusions I could draw from each part of the thesis. Furthermore, I show, using an exemplary analysis, how to link the knowledge gained from each part for further research and its practical relevance for hydrological modelling of extreme events.

SYNTHESIS AND CONCLUSIONS

A systems approach in hydrology implies considering the catchment as an integrated system rather than a collection of isolated components. It conceptualises the hydrological cycle as a connected network of interacting water processes, with the water balance providing a fundamental constraint that links all constituent components. Importantly, changes in any single component will inevitably propagate through and affect others. This also implies linking the different classes of problems for a holistic hydrologic assessment of the system.

In Chapters 2–4, I presented three studies aimed at improving the representation of extremes (flash floods) in hydrological models. Across these studies, I employed a diverse set of modelling approaches—ranging from bottom-up, process-based models (CATFLOW), to top-down conceptual models (HBV), and data-driven methods (LSTM). I showed that by learning from the past—either through model and parameter transfer informed by previous field experiments, or by training data-driven models on large-scale hydrological datasets—we can augment the (often scarce) information available at the catchment scale.

Linking analysis, design and inverse

In Sections 5.1–5.4, I provide a concise summary of the key findings from each part of the thesis with respect to the objectives set out in Section 1.5 and demonstrate, through an exemplary analysis, what can be achieved by linking the findings from the three chapters.

5.1 ANALYSIS PROBLEM

The results of Chapter 2 reveal the feasibility of using gradient based approximations like representative hillslopes in process based models for reliable estimations of overland flow runoff during high intensity convective storms.
The key findings of Chapter 2 are:

Representative hillslopes can model flash floods in data-scarce settings

- Linking process-based models with climate reanalysis products helps addressing the model initialisation problem. This approach reduces uncertainties associated with the calibration of antecedent conditions and has important implications for operational flash-flood forecasting.
- Parameter transfer provides initial estimates of the flood characteristics in a nearly uncalibrated manner. Accounting for pref-

erential flow pathways calls for additional events and more detailed soil data.

- By using gauge-level markings at downstream flood control reservoirs, initial estimates of event responses in poorly gauged catchments can be obtained. Leveraging the information contained in such flood defence structures can substantially increase the training sample in hydrologically data-scarce regions.

5.2 DESIGN PROBLEM

Representative hillslopes can be scaled up

In my second study (Chapter 3), I implemented and tested the concept of representative hillslopes to an intermediate scale catchment (about 200 km^2). By relying on the vectorised representation for the subcatchments and linking them using standard hydraulic routing schemes, the framework enables the simulation of historical flash floods while explicitly accounting for dominant preferential flow mechanisms.
The key findings of Chapter 3 are:

- Calibrating the model at the larger catchment scale to capture hillslope scale preferential flow processes enables subsequent reconstruction of storm responses in severely impacted yet sparsely monitored headwater regions.

Nature-based flood defence measures deliver benefits

- Nature based solutions such as hedgerows in agricultural fields reduce hydrological connectivity by locally impeding Hortonian runoff and enhancing infiltration, thereby delaying the peak and desynchronising contributions from adjacent fields, with downstream benefits along the river network.

5.3 INVERSE PROBLEM

Discharge has potential to correct precipitation inversely

The third study of my thesis (Chapter 4) showed that discharge as measured at the catchment outlet retains sufficient memory of the driving precipitation at the catchment scale and that this information is invertible using machine learning methods.
The key findings of Chapter 4 are:

- The performance gain in using discharge information for inversely correcting precipitation was significantly higher for days with increasing precipitation amounts. This is logical from a hydrological perspective, as average conditions in the river do not provide much information about the driving precipitation; rather, it is the extreme precipitation events (usually missed in reanalysis products) that the discharge can help correct.

- The out-of-sample test showed that the method has potential to be used to tackle the predictions in the ungauged basin (PUB) problem. The approach also has potential to generate coherent, long-term statistical records (including historical storms) of catchment forcings.

Past can be corrected from knowledge of the present

- The inversely estimated precipitation estimates can be integrated into existing modelling strategies, thereby improving the forward modelling of streamflow and other catchment state fluxes.

5.4 WHAT CAN BE DONE WITH THESE FINDINGS?

Maintaining a coherent historical record of flood events (Blöschl et al., 2020) is crucial for characterising catchment response under extreme hydrological conditions, thereby alleviating some of the issues (Montanari et al., 2024) related to predictive uncertainty across a wide range of flow regimes. Flood level markings and other contemporary records provide information (Henkel, 1994; Vogt, 2024) about the impacts of such historical floods, which could be utilised to gauge the driving storms that led to the event.

Reconstructing historical storm of 1994

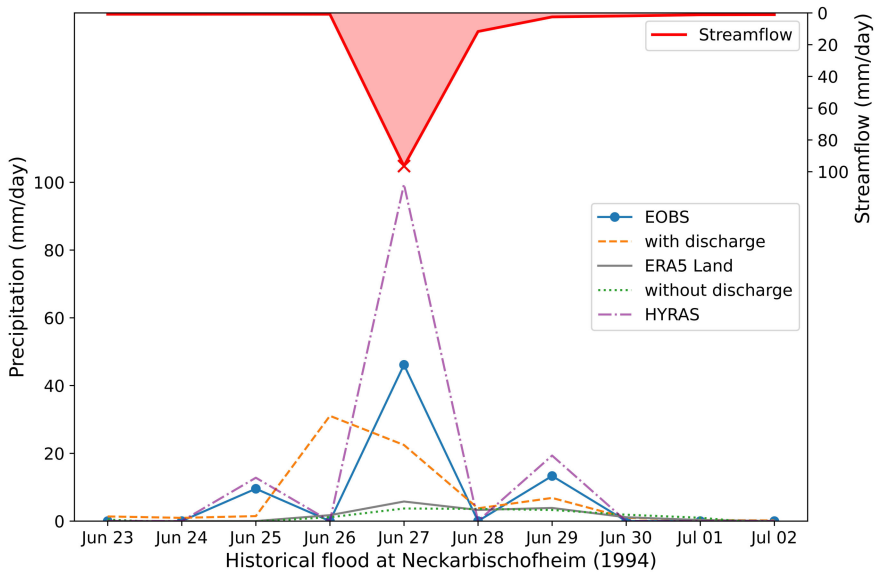

Figure 5.1: Inverse estimation for the historical Neckarbischofsheim flood. The red line represents the daily streamflow timeseries (reconstructed in Chapter 3), with a cross marking the flood day. The orange curve indicates precipitation predicted by the *with_discharge* LSTM model, while the green curve shows precipitation predicted by the *without_discharge* model (both trained in Chapter 4). The blue and purple line reflects the gauge-based E-OBS (used as training target in Chapter 4) and HYRAS respectively, and the grey line represents the original ERA5-Land estimate (also used for initialising our models in Chapter 2).

For such historical floods, reanalysis products like ERA5 Land has deficiences in representing the event magnitude. Consider the case (Fig. 5.1) of the historical Neckarbischofsheim flood investigated in Chapter 3, ERA5 Land catchment (Muñoz-Sabater et al., 2021) average predicts a daily sum of 5 mm, while the HYRAS observational product of the German Weather Services (Rauthe et al., 2013) totals a storm of around 100 mm

The inversion methodology that I proposed in Chapter 4 appears to be a viable approach for generating realistic estimates of historical storm events from reanalysis products. Because the model has already been trained on large-sample datasets from the GRDC (Färber et al., 2023) and other data providers, it can be readily applied to new and previously untested regions (transfer of knowledge by data-based learning). For Neckarbischofsheim, the problem is further compounded by gaps in the streamflow record; the observed flood level was far above the calibrated range of the gauging station rating curve, consequently leading to severe, disastrous impacts in the town (Henkel, 1994; Vogt, 2024), while the actual flood magnitude remained unknown.

As I showed in Chapter 3, by means of a coupled hydrological simulation with 1D routing, and representing subcatchments as representative hillslopes (a process-based version of a classical rainfall-runoff model), aided by model and parameter transfer from other experimental catchments (transfer of knowledge by hydrologic regionalisation), we are able to reconstruct the flood values for such data-scarce catchments.

Exemplary analysis linking my thesis

A natural approach would be to try to merge both these methods - using the reconstructed flood values (Chapter 3) to invert the driving precipitaion forcings (Chapter 4) and then to compare the new estimate with observational data. In this line, I ran (Fig. 5.1) the trained *with_discharge* and *without_discharge* LSTM models (three runs with different initialisation seeds as in Chapter 4) using the reconstructed discharge record till the gauging station at Neckarbischofsheim (Chapter 3) for the summer flood of June 1994. As expected and in line with my observations from Chapter 4, incorporating the discharge information led to improvement in the precipitation estimate from the reanalysis product. Also, the performance of the original ERA5 Land estimate was nearly comparable to the model run without including discharge information.

It is important to stress that since the continental scale model was trained on E-OBS data due to data availability considerations over entire Europe, we are not able to compare directly to the national HYRAS product which showed better accuracy in representing such

historical storm magnitudes over Germany. The exemplary analysis indicates that substantial insights can be gained by addressing hydrological system problems in an integrated manner rather than treating them in isolation, pointing to a promising direction for hydrological research.

5.5 OUTLOOK

Hydrology as an evolving science

Recently, increased attention (Kratzert et al., 2018; Shen et al., 2023) has been given to the performance of data-driven models in comparison with traditional hydrological models . This has led to differing perspectives (Beven, 2020; Nearing et al., 2021) within the hydrological community, with discussions reflecting a range of views on the respective roles and advantages of these modelling approaches, reminiscent of earlier well founded debates in hydrological science, such as those concerning determinism and stochasticity (Yevjevich, 1974) in the 20th century.

Catchment is an integral system

As shown in this thesis, it is important not to compartmentalise hydrology into isolated blocks. The interconnected nature of catchment systems means that a single modelling approach will rarely be sufficient. Instead, a range of complementary methods (Hrachowitz and Clark, 2017) is needed, with the most appropriate approach chosen for the problem at hand. For example, relying solely on machine learning models to reconstruct the Neckarbischofsheim flood would neglect valuable information from detailed historical field studies in the region. Conversely, a purely process-based approach for inverse modelling of precipitation would be hindered by high computational demands and ill-posedness.

The approaches explored in this work provide a foundation for more comprehensive future investigations -

Scope for future work

- The coupled modelling approach using representative hillslopes to tackle flash floods at the mesocatchment scale should be evaluated in regions beyond the Kraichgau and with alternative bottom-up models in addition to CATFLOW.
- The use of water levels from flood reservoirs to inversely get discharge timeseries could be expanded to other regions. Especially in the mountainous regions of Bavaria and Baden-Württemberg in Germany, there may be potential to gain more insight into the flash-flood formation in such small catchments.
- Historical high flood level markings are present in most towns and villages around the world, this could aid in the reconstruction of flood dynamics.

- To better characterise short-duration, high-intensity events, model responses should be analysed at finer temporal resolutions. In this context, the inversion methodology could also be applied to hourly time series.
- The observational benchmark used for inversion was chosen as EOBS to ensure data availability across Europe. Country-specific models could be developed, making use of higher quality national products.
- The inversion was performed on spatially averaged catchment time series; recent advancements in machine learning models now enable handling spatially distributed information. Thus, inverting the spatial field of precipitation, coherent with the current state-of-the-art gridded products.

The fact that engineers and earth system scientists, coming from two different schools of thought, have found common ground to advance hydrological sciences implies that much can be achieved through the synergy of complementary (*and yet sometimes competing*) philosophies.

Part VI

APPENDIX A

APPENDIX A (CHAPTER 2)

A.1 ENERGY CONSIDERATIONS IN THE DERIVATION OF THE REPRESENTATIVE HILLSLOPE CATENA

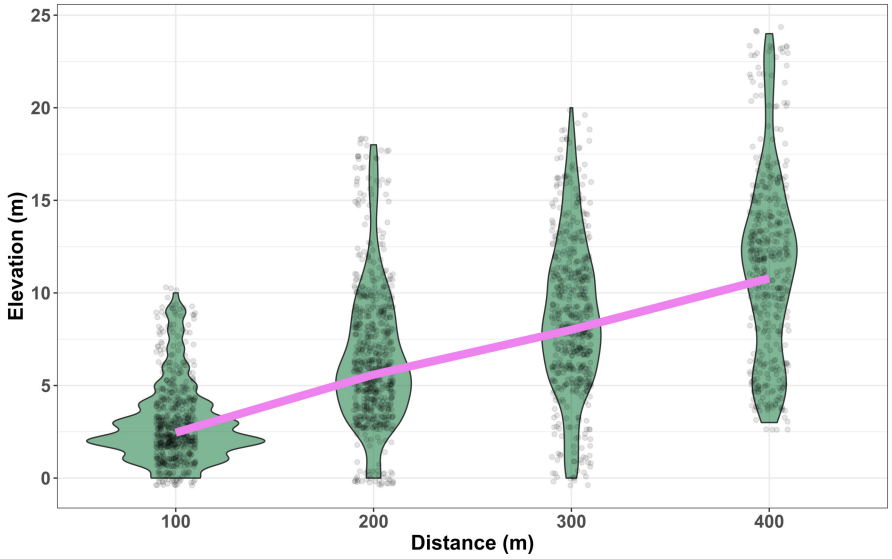

Figure A.1: Plots showing the distribution of elevation values of each cell within four distinct distance classes from Figure 2.3. The pink line denotes the representative hillslope profile derived from the mean elevation values using the approach detailed in Section 2.3.1.

From Newtonian mechanics, flow potential at a relative elevation (h) is defined as

$$E = m \times g \times h \tag{A.1}$$

Where E is the potential energy of the water on the hillslope (J), m is its mass (kg), g represents the gravitational acceleration (m s^{-2}), and h is the relative height of the water above a reference (m).

For each class (say $x = l$ m), the average flow potential due to elevation values is related to the sum of the individual flow potential of all the cells (j:1 to n) within the class

$$E_{\text{avg}}^{(x=l)} = \frac{E_{\text{total}}^{(x=l)}}{n} = \frac{\sum_{j=1}^{n} E_j^{(x=l)}}{n} \tag{A.2}$$

The flow potential in the representative hillslope element at $x = l$ m is given by:

$$\bar{E}^{(x=l)} = \bar{m} \times g \times \bar{h} \tag{A.3}$$

Where $\bar{h}$ is the estimate of weighted mean elevation for a class at distance l, calculated using Equation 2.2 in Section 2.3.1. Table A.1 shows these different energies for all the four classes illustrated in Fig. A.1. On average, the relative errors between flow potential in the classes and in the derived representative catena are seen to decrease as the distance from the stream increases.

Table A.1: The difference between the total flow potential in each class (Fig. A.1) and in the derived representative hillslope in terms of density, ρ and gravitational acceleration, g.

Classes	$l = 100$ m	$l = 200$ m	$l = 300$ m	$l = 400$ m
Energy in class , $E_{\text{avg}}^{(x=l)}$	$301\rho g$	$660\rho g$	$877\rho g$	$1152\rho g$
Energy in profile, $\bar{E}^{(x=l)}$	$245\rho g$	$558\rho g$	$801\rho g$	$1079\rho g$
Relative Error (%)	-22.8	-18.28	-9.4	-6.7

A.2 STORM EVENT ON 08.06.2016

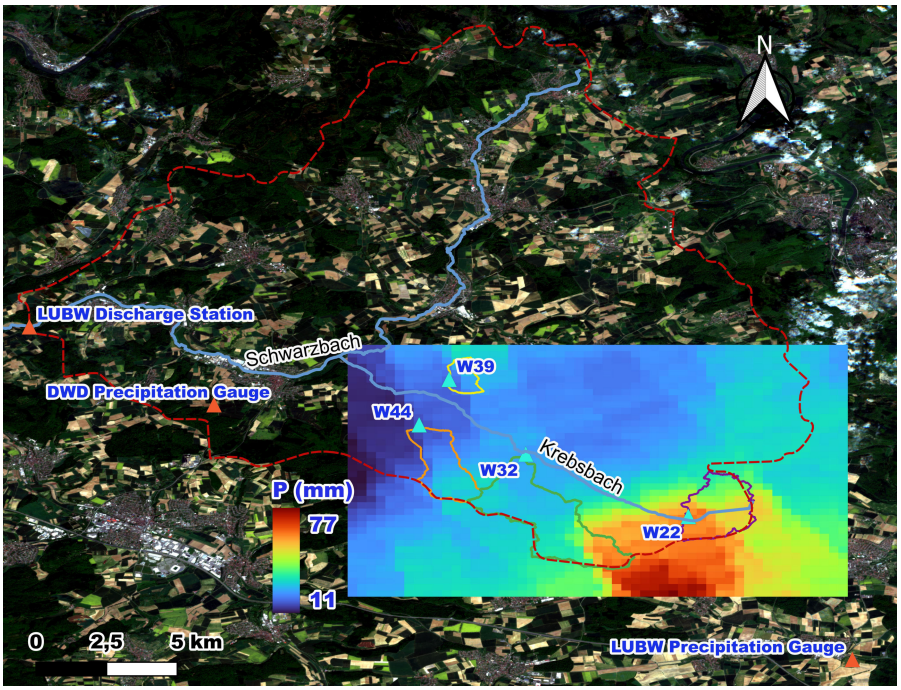

Figure A.2: Overview of the Schwarzbach catchment till the downstream streamflow station at Eschelbronn. An overlay layer displays the total accumulated precipitation (in mm) during the event across the four catchments. Additionally, the DWD and LUBW precipitation gauges are shown.

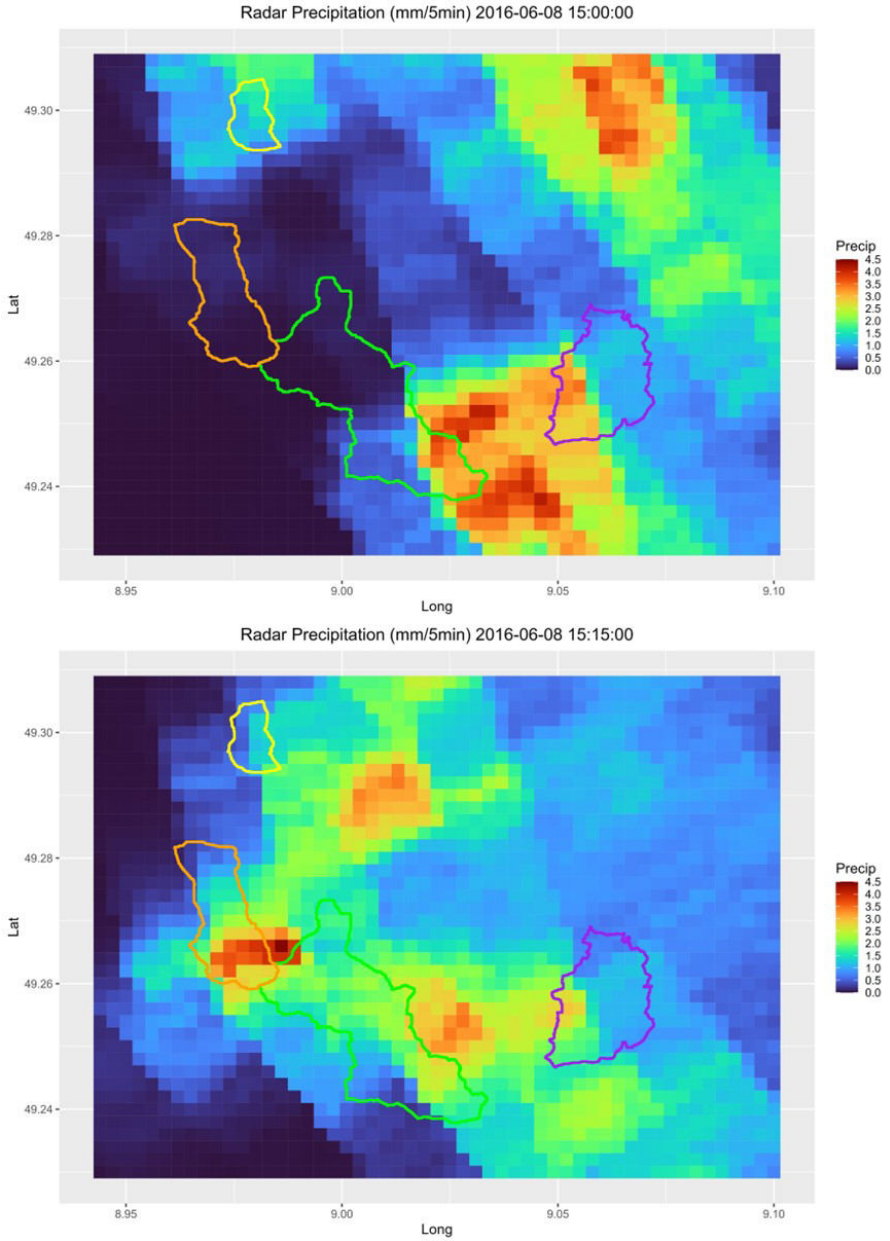

Figure A.3: Evolution of the convective storm event on 08.06.2016 over the Krebsbach as captured in the chosen radar based precipitation product (Kachelmannwetter, 2023)

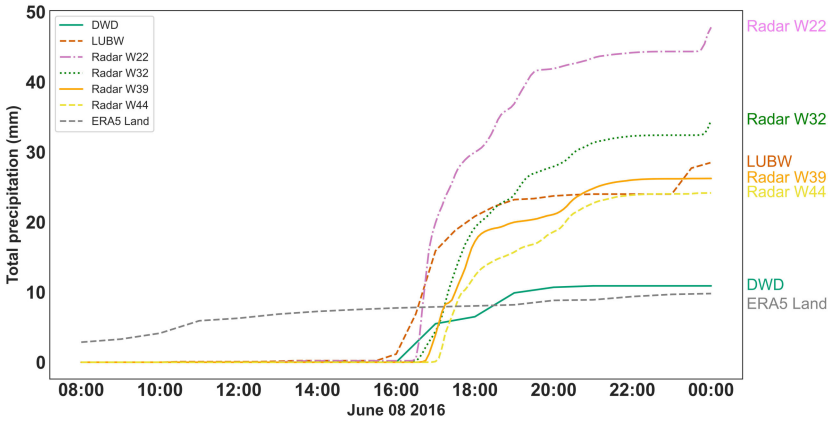

Figure A.4: Cumulative timeseries plots showing the hourly dynamics captured by the different precipitation estimates for the 08 June 2016 event.

Figure A.5: Impact of flash floods on 08.06.2016 over Catchment W22 (Zweckverband Hochwasserschutz Elsenz-Schwarzbach, 2016).

A.3 FLOOD ESTIMATION USING RESERVOIR MASS BALANCE

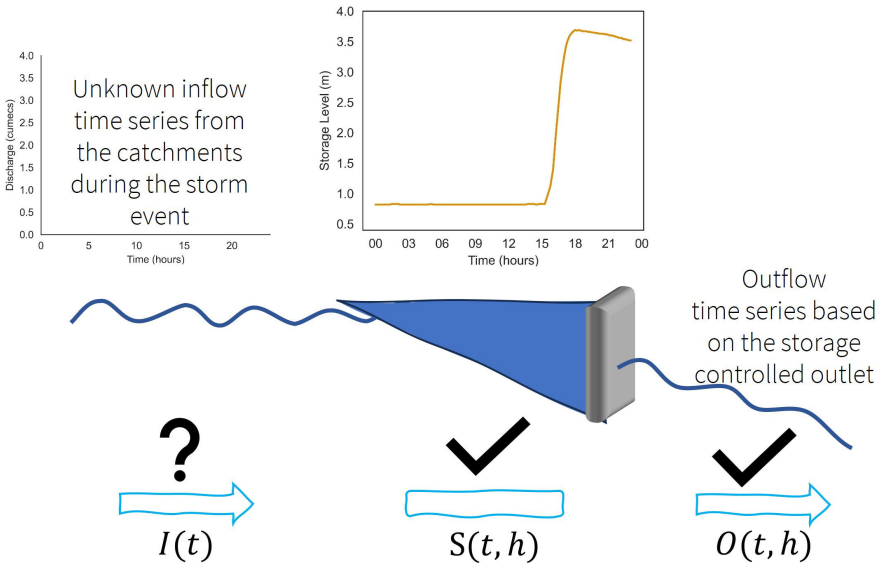

Figure A.6: Schematic representation of the reservoir mass balance inversion.

Mass conservation has long been the foundation of hydrological modeling. This basic physical law is usually expressed (for hydrological systems) in the form of:

$$\frac{dS}{dt} = I(t) - O(t) \tag{A.4}$$

where the change of a system's mass storage (S) with respect to time (t) is equal to total mass input, $I(t)$ minus total mass output, $O(t)$. This represents one of the most basic physical constraints placed on the functioning of any hydrological system.

Considering the mass balance of the downstream flood reservoirs in the four catchments (Fig. 2.1 and Fig. A.2) as shown in Fig. A.6, the storage in the reservoir at any time t being a function of the level (h). An automatic recorder measures the water level in the reservoir as shown in Fig. A.6. The outflow being again a function of the water level in the reservoir. Having knowledge of the reservoir geometry relations ($S = f(h)$) and the stage discharge relationship of the outlet ($O = g(h)$), we now need to estimate the inflow to the reservoir from the catchment due to the convective storm activity. Again, from Eq A.4:

$$\frac{dS}{dt} = I(t) - O(t) \tag{A.5}$$

$$\frac{S(t+\Delta t, h+\Delta h) - S(t,h)}{\Delta t} = I(t) - O(t) \tag{A.6}$$

Hence, the inflow is given by,

$$I(t) = \frac{S(t + \Delta t, h + \Delta h) - S(t, h)}{\Delta t} + O(t) \tag{A.7}$$

Now for the uncertainty analysis, we consider a relative error of 5% in the reservoir level measurements and again calculate the inflows using Eq. A.7. The inflow hydrograph obtained, and calculations are further shown for catchment W22 in Fig. A.7 and Table A.2 respectively.

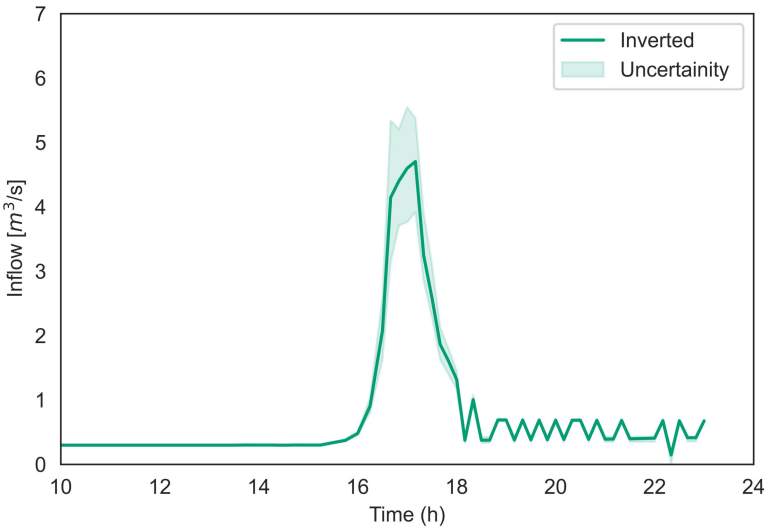

Figure A.7: Reconstructed inflow time series for catchment W22. All times in CET.

Table A.2: Reservoir mass balance calculations for catchment W22

S No	Time	Level (m)	Storage (m^3)	Change in Storage (m^3)	Outflow (m^3/s)	Dt (s)	Inflow (m^3/s)
63	08-06-2016 15:45	1.12	41.94883	16.42903	0.354373	900	0.372627
64	08-06-2016 16:00	1.4	107.4919	65.54307	0.407547	900	0.480373
65	08-06-2016 16:15	1.98	471.3575	363.8656	0.500051	900	0.904346
66	08-06-2016 16:30	2.48	1828.668	1357.311	0.565968	900	2.074091
67	08-06-2016 16:40	2.82	3952.833	2124.165	0.605154	600	4.145429
68	08-06-2016 16:50	3.06	6215.17	2262.337	0.631057	600	4.401619
69	08-06-2016 17:00	3.26	8583.66	2368.49	0.651033	600	4.598516
70	08-06-2016 17:10	3.43	11005.15	2421.491	0.666923	600	4.702742
71	08-06-2016 17:20	3.53	12547.01	1541.86	0.676144	600	3.245911
72	08-06-2016 17:30	3.6	13690.94	1143.932	0.68265	600	2.589203

A.4 CATFLOW SOIL MOISTURE SIMULATIONS

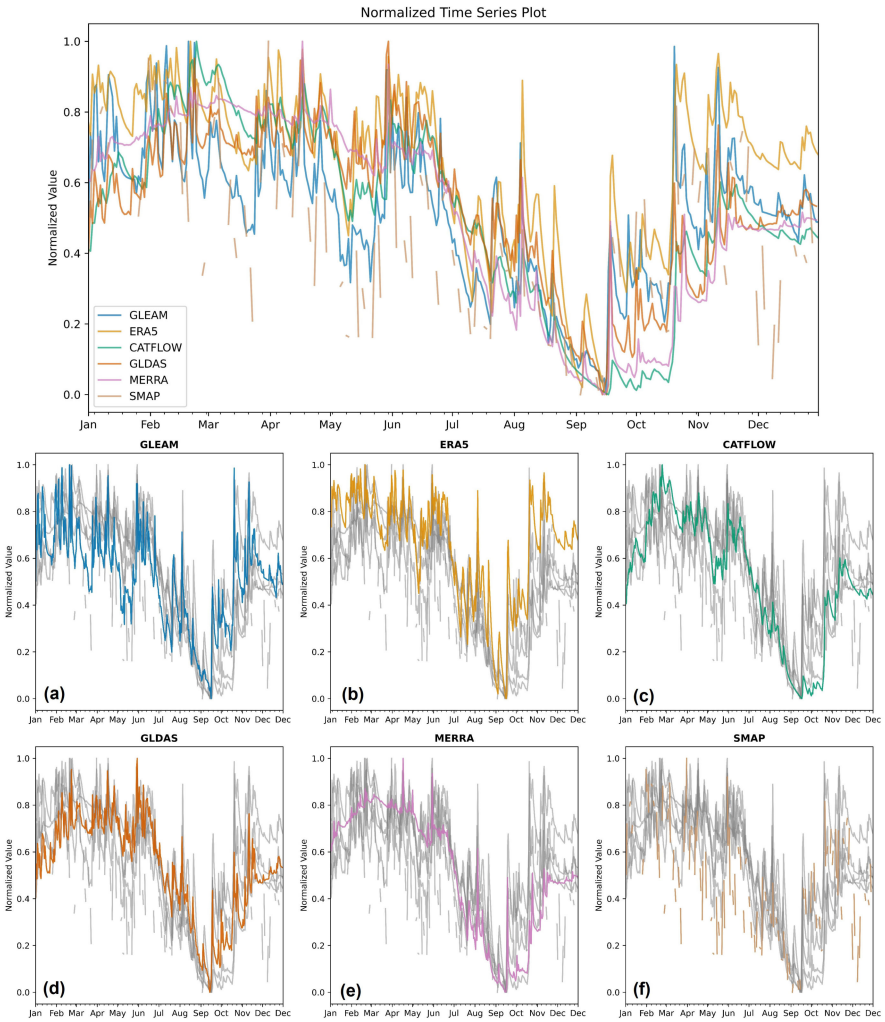

Figure A.8: Normalized time series plot for catchment W22 comparing the yearly dynamics of CATFLOW simulated soil moisture with the various soil moisture products. (a) GLEAM (Global Land Evaporation Amsterdam Model; (Miralles et al., 2011)), (b) ERA5 Land (Muñoz-Sabater et al., 2021), (c) CATFLOW, (d) GLDAS (NASA Global Land Data Assimilation System, GLDAS 2.2 GRACE DA; (Li et al., 2019)), (e) MERRA (Modern Era Retrospective analysis for Research and Applications, version 2, tavg1_2d_lnd_Nx; (Global Modeling and Assimilation Office, 2015)), and (f) SMAP (Soil Moisture Active Passive L Band radiometer Level 3 product; (Chan et al., 2018)).

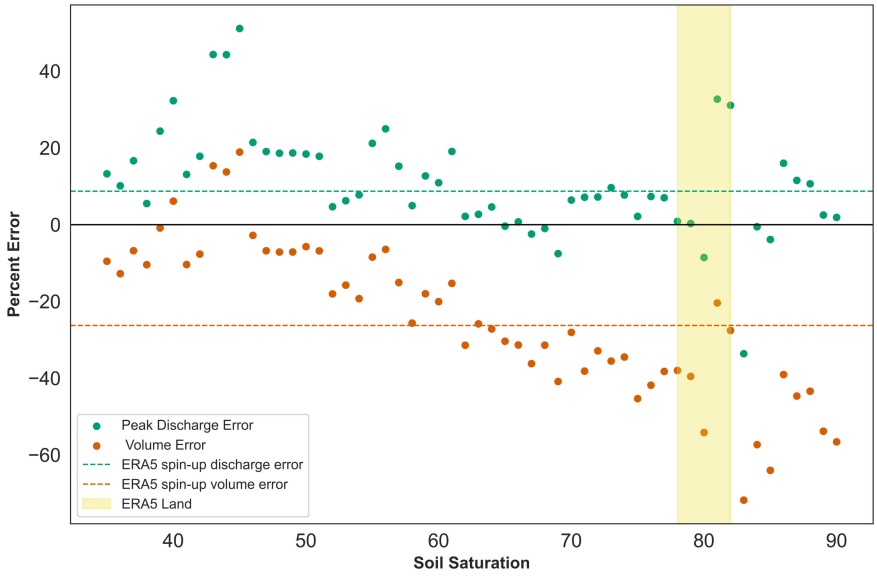

Figure A.9: Scatterplot demonstrating how changes in the initial soil saturation values affect the variation in peak discharge and volume errors for catchment W22. The dotted lines represent the errors in the original set-up, where soil moisture conditions were inferred from the ERA5 Land Initialization. The yellow-shaded region indicates the possible absolute value of soil saturation from the ERA5 Land directly.

Table A.3: Textural data required for estimating the hydraulic functions using the soil pedotransfer functions

Source	Sand (%)	Silt (%)	Clay (%)
LGRB (State Ministry for Geology, Natural Resources and Mining, Baden Württemberg (LGRB), 2024) – Sheet No 6820, Schwaigern)	1.7	83.4	14.9
EU LUCAS (European Union Land Use/Cover Area frame Survey (Ballabio et al., 2016), POINT_ID 42462906)	5	82	13

Part VII

APPENDIX B

APPENDIX B (CHAPTER 3)

B.1 ELSENZ SCHWARZBACH CATCHMENT

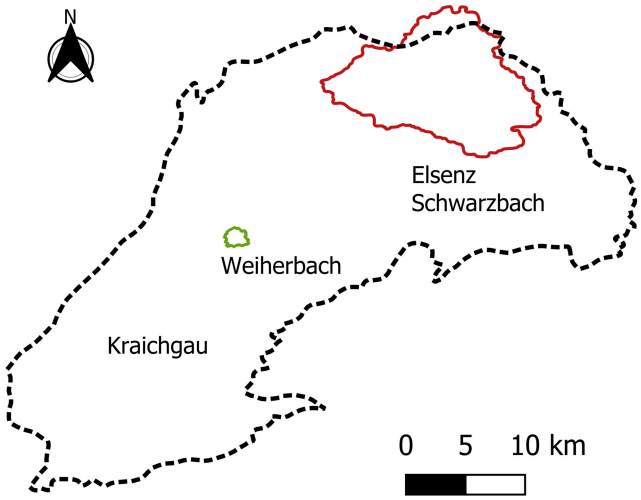

Figure B.1: Catchments Weiherbach and Elsenz Schwarzbach

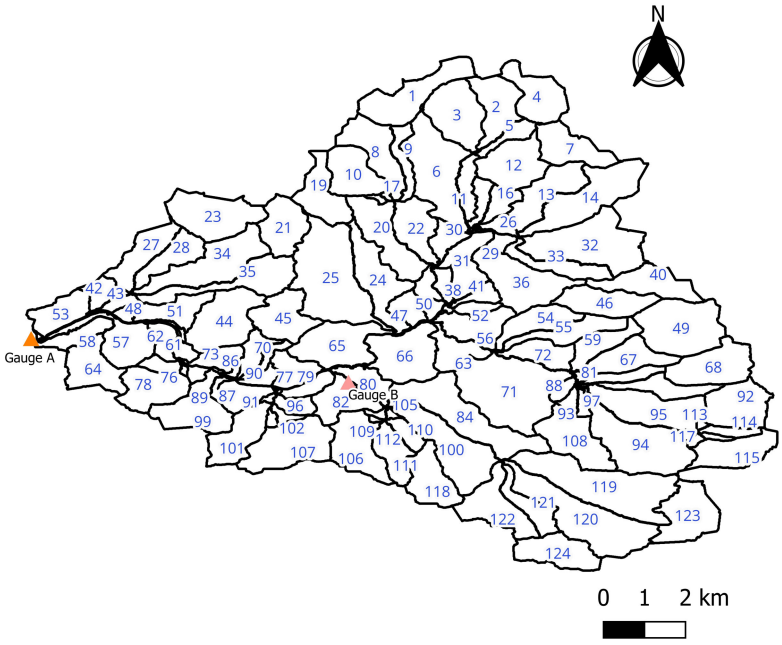

Figure B.2: Subbasins of the Elsenz Schwarzbach.

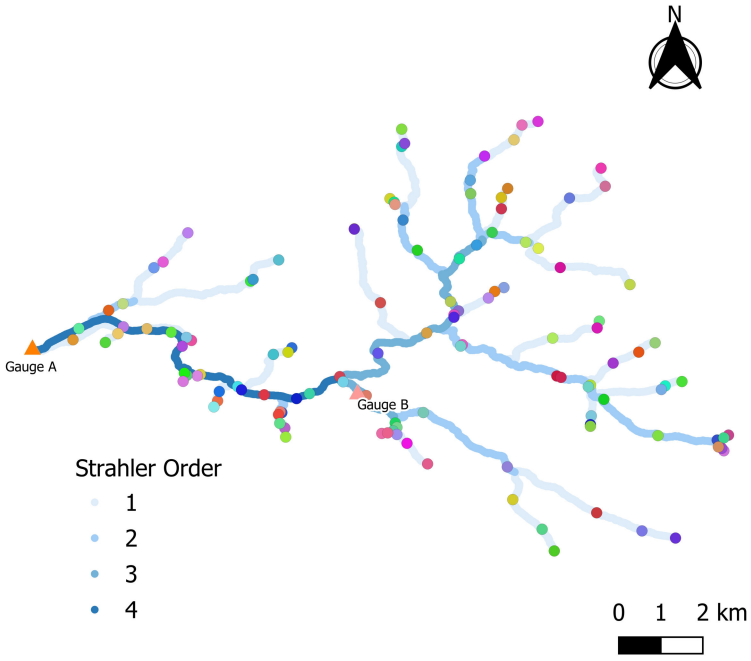

Figure B.3: Stream network used for routing within the catchment, along with the location of the outlets at which each representative hillslope is connected.

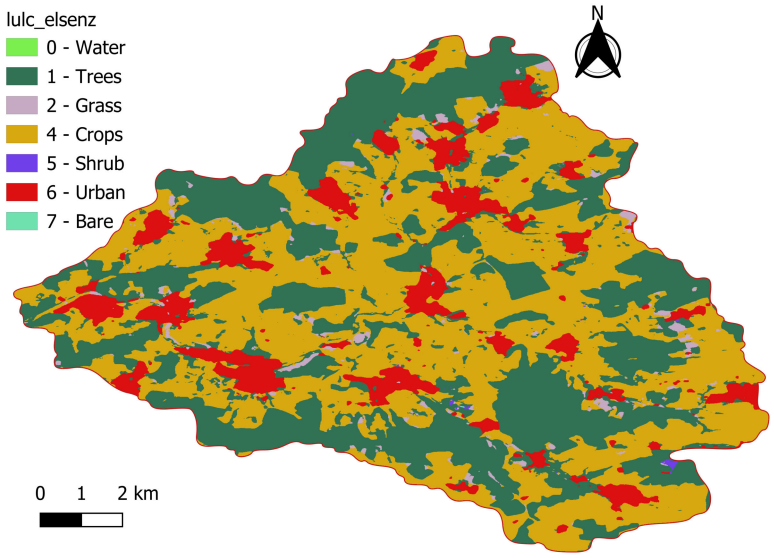

Figure B.4: Dominant landuse classes in the catchment as per the Dynamic World dataset.

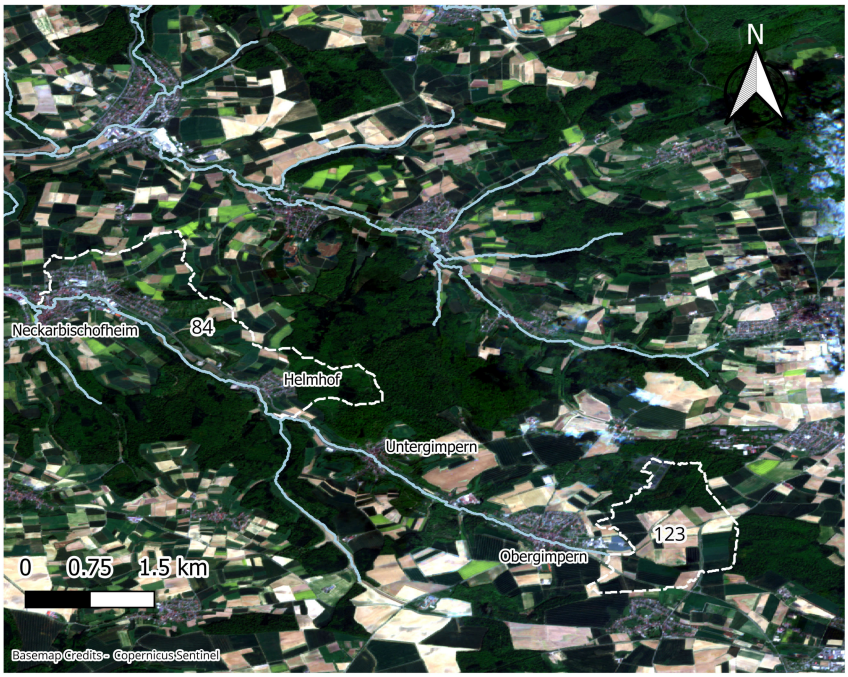

Figure B.5: Headwater subbasins 84 and 123, which drain to the city of Neckarbischofsheim

Figure B.6: Locations within the Elsenz–Schwarzbach catchment: (A) Eschelbronn near the outlet gauging station, (B) Aglasterhausen in the headwater region, and (C) a flood regulating structure at Waibstadt. (Field visit on 02 August 2025)

B.2 FLOOD EVENTS IN THE ELSENZ SCHWARZBACH

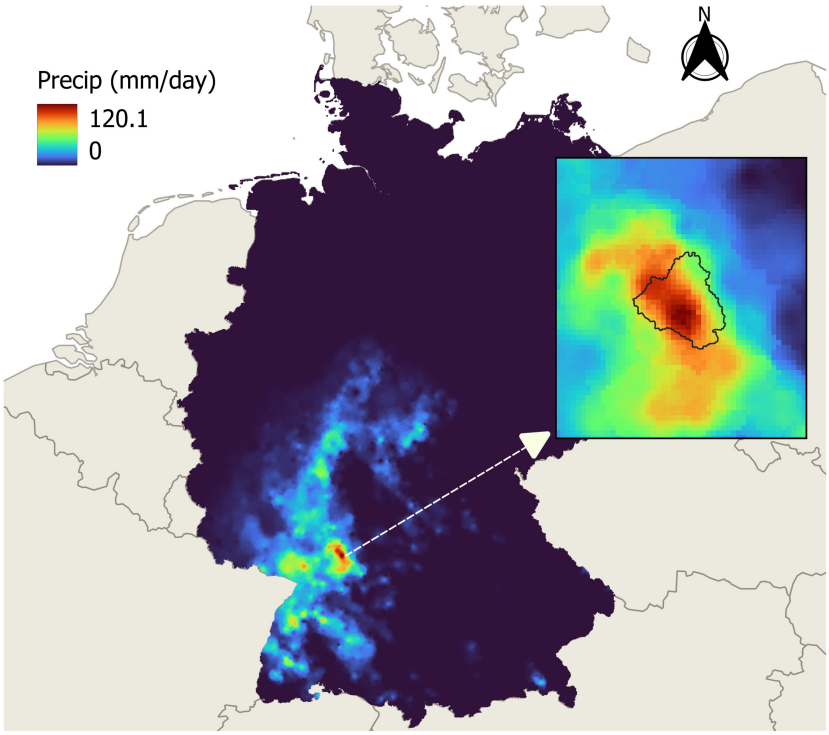

Figure B.7: Daily accumulated precipitation sums on 27 June 1994 over Germany. The inset figure shows the convective storm centre in the Elsenz Schwarzbach catchment.

Figure B.8: Aftermath of 27 June 1994 flash floods over the Elsenz Schwarzbach (Image credit: Archiv Neckarbischofsheim).

Figure B.9: Flood evolution (27.06.1994) over agricultural landscapes in the Elsenz–Schwarzbach catchment. Archival flood footage (Vogt, 2024) highlights extensive open fields and the absence of vegetative barriers in the affected area (Image credit: Archiv Neckarbischofsheim).

Figure B.10: Impact of flash floods on 08.06.2016 over Elsenz Schwarzbach (Image credit: Zweckverband Hochwasserschutz Elsenz–Schwarzbach)

B.3 MODEL PREPROCESSING

In this study, we utilize Docker containers (Boettiger, 2015) to preprocess terrain data for the CATFLOW hydrological model (Figure B.11). Docker is a containerization technology that packages software along with everything needed to run it—code, libraries, and system tools. This approach ensures consistent behavior across different computers and users, representing a significant step toward computational reproducibility (Mälicke, 2024; Nature Editors, 2012).

We first employed containerized tool *tool_whiteboxgis* (Manoj J et al., 2025) for the GIS preprocessing of the Digital Elevation Model (DEM) data. This tool wraps the open-source platform WhiteboxTools (Lindsay, 2014) for geospatial analysis. The *tool_whiteboxgis* prepares standard raster files, such as aspect, elevation, and distance, which can then be used directly to derive the required terrain files for the hydrological model CATFLOW.

These raster files are then used as inputs for tool_catflow (Manoj J and Dolich, 2025). To derive the hillslope profiles for each sub-catchment in the study area, we run tool_catflow in parallel to ensure fast, efficient processing. Each sub-catchment is processed in its own container, thus providing modularity (Boettiger, 2015) in both programming and hydrological sense. The relevant code and software are finally archived in Zenodo (Manoj J and Dolich, 2025; Manoj J et al., 2025) for reusability.

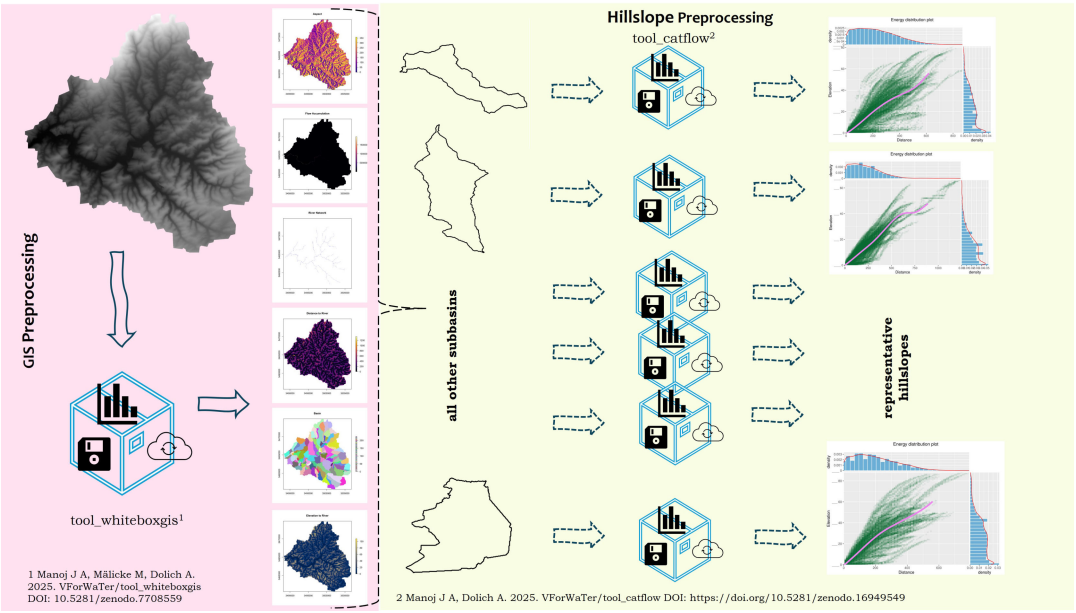

Figure B.11: Brief overview of the preprocessing workflow.

Part VIII

APPENDIX C

APPENDIX C (CHAPTER 4)

C.1 LSTM CONFIGURATIONS

Table A1 details the static and dynamic inputs used for setting up the with_discharge and without_discharge LSTM ensemble models. The hyperparameter settings for both models are shown in Table A2, while Figure A1 provides the comparison results for both runs.

Table C.1: Model configurations for the LSTM model runs.

Model	Inputs		Output
	Static Attributes	Dynamic Attributes	
with_discharge	area (area of catchment – km^2) ele_mt_sav (spatial mean elevation – m above sea level) frac_snow (fraction of precipitation falling as snow) pet_mm_syr (potential evapotranspiration annual mean – mm)	total_precipitation_sum_era5 (precipitation daily sums – mm day^{-1}) temperature_2m_mean (daily mean temperature – °C) surface_net_solar_radiation_mean (shortwave radiation – Wm^{-2}) surface_net_thermal_radiation_mean (net thermal radiation at the surface – Wm^{-2}) qobs_lead (lead streamflow 7 days – mm day^{-1})	eobs_precipitation (precipitation daily sums – mm day^{-1})
without_discharge	area (area of catchment – km^2) ele_mt_sav (spatial mean elevation – m above sea level) frac_snow (fraction of precipitation falling as snow) pet_mm_syr (potential evapotranspiration annual mean – mm)	total_precipitation_sum_era5 (precipitation daily sums – mm day^{-1} temperature_2m_mean (daily mean temperature – °C) surface_net_solar_radiation_mean (shortwave radiation – Wm^{-2}) surface_net_thermal_radiation_mean (net thermal radiation at the surface – Wm^{-2})	eobs_precipitation (precipitation daily sums – mm day^{-1})

Table C.2: Hyperparameter settings for the LSTM network.

Hyperparameter	LSTM Network
Hidden Layer	1
Hidden cells	64
Batch size	256
Sequence length	365
Epochs	5
Drop out	0.4
Learning rate	0.001
Optimizer	Adam

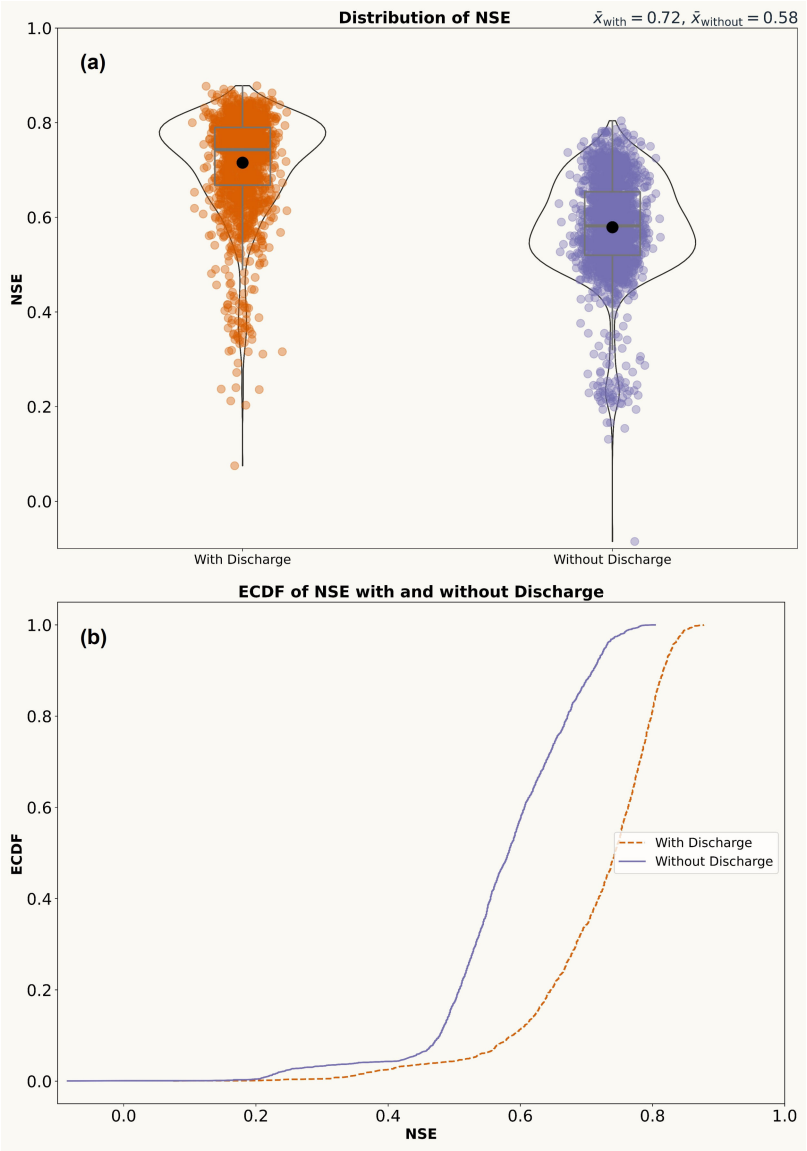

Figure C.1: Comparison of the mean performance of the two regional scale LSTM models (*with_discharge* and *without_discharge*). (a) Top panel depicts violin plots with included boxplots showing the distribution of performance (quantified by comparing the LSTM model simulated precipitation series to the observational EOBS time-series over the testing period: NSE) (b) Bottom panel displays cumulative distribution plots for the performance of the two models.

C.2 HYDROLOGICAL MODELLING

Hydrologiska Byråns Vattenbalansavdelning (HBV) The HBV model (Bergström and Forsman, 1973) is a so-called conceptual hydrological model that is used to simulate rainfall-runoff processes at the catchment scale. It makes use of different catchment water stores (storage

elements, also referred to as buckets). Each storage element represents a certain compartment of a catchment (e.g. groundwater, surface water bodies, soil zone). The main input requirements include precipitation, temperature and potential evapotranspiration. The model has several empirical parameters that need to be calibrated during the model training phase. A more detailed description of the model architecture and set up can be found in the studies by Seibert (2005) and Loritz et al. (2024a).

CATFLOW The physically based model CATFLOW for catchment water and solute dynamics was developed as part of the detailed process studies carried out from 1991 – 1996 in the Weiherbach catchment in South-West Germany (Zehe et al., 2001). The basic modeling unit is a 2-D hillslope, discretized by curvilinear orthogonal coordinates in the vertical and downslope directions. Soil water dynamics within the hillslopes are characterized using the potential based form of the 2D Darcy–Richards equation. Overland flow is simulated using the diffusion wave approximation of the Saint-Venant equation and explicit upstreaming, in combination with the Gauckler–Manning–Strickler formula. A detailed model description with the workflow required for setting up the model can be found in Manoj J et al. (2024).

C.3 PERFORMANCE METRICS

Nash-Sutcliffe Efficiency (NSE) – First proposed by Nash and Sutcliffe (1970), the Nash–Sutcliffe efficiency (NSE) is one of the most widely used similarity measures in hydrology for calibration, model comparison, and verification. It measures how well the simulated timeseries (y_{sim}) matches the observed values (y_{obs}).

$$\text{NSE} = 1 - \frac{\sum(y_{\text{obs}} - y_{\text{sim}})^2}{\sum(y_{\text{obs}} - \overline{y}_{\text{obs}})^2} \tag{C.1}$$

Values closer to 1 indicate excellent model performance (D. N. Moriasi et al., 2007), while NSE values near or below 0 suggest that the model, in fact, performs worse than simply using the mean of the observed values.

Mean Wet Day Precipitation (MWD: mm day^{-1}) – The Expert Team on Climate Change Detection and Indices (ETCCDI – World Climate Research program (WCRP), 2021) recommends evaluating the intensity of precipitation on wet days (defined as a day with a minimum of 1 mm precipitation) to understand systematic over or underestimation of precipitation amounts. This metric (Simple Daily Intensity Index as per ETCCDI) is reported as the mean daily precipitation on days where precipitation > 1 mm. Let P_i be the daily precipitation amount

Table C.3: Validation test cases for hydrological models.

Model	**HBV** (Bergström and Forsman, 1973)				**CATFLOW** (Zehe et al., 2001)	
Catchment (Fig. C.6)	Elsenz Schwarzbach (Manoj J et al., 2024)		Lippe (Loritz et al., 2024a)		Headwater catchment W32 in Elsenz Schwarzbach (Manoj J et al., 2024)	
Area (km^2)	196.5		3366.3		5.6	
Scenario	Using original ERA5 Land precipitation	Using inversely generated precipitation (*with_discharge*)	Using original ERA5 Land precipitation	Using inversely generated precipitation (*with_discharge*)	Using original ERA5 Land precipitation	Using inversely generated precipitation (*with_discharge*)
Forcings	ERA5 Land precipitation, potential evapotranspiration (Hargreaves), air temperature (ERA5 Land)	Inversely generated precipitation, potential evapotranspiration (Hargreaves), air temperature (ERA5 Land)	ERA5 Land precipitation, potential evapotranspiration (Hargreaves), air temperature (ERA5 Land)	Inversely generated precipitation, potential evapotranspiration (Hargreaves), air temperature (ERA5 Land)	ERA5 Land precipitation, potential evapotranspiration (ERA5 Land), plant and soil parameters from Manoj J et al. (2024)	Inversely generated precipitation, potential evapotranspiration (ERA5 Land), plant and soil parameters from Manoj J et al. (2024)
Calibration period	01 Jan 2000–31 Dec 2010		01 Jan 1990–31 Dec 2010		Uncalibrated predictions	
Validation period	01 Jan 2011–31 Dec 2016		01 Jan 2011–31 Dec 2020		01 Jan 2008–31 Dec 2015	
Outputs compared	Streamflow		Streamflow		Soil moisture	
NSE	0.57	0.70	0.64	0.83	0.70–0.82	0.64–0.86

on wet days, ($P_i > 1$ mm). If N represents the total number of wet days, then:

$$\text{MWD} = \frac{\sum_{i=1}^{N} P_i}{N} \tag{C.2}$$

95th Percentile Precipitation (R95P: mm day^{-1}) – This metric denotes the daily precipitation value at which 95% of all daily values (again only considering rainy days) are lower (top 5% events). This helps to assess the ability to capture extreme precipitation events. Let P_i be the daily precipitation amount on wet days, ($P_i > 1$ mm).

$$\text{R95P} = \text{Percentile}\left(\{P_i \mid P_i > 1 \text{ mm}\}, 95\right) \tag{C.3}$$

Spearman Rank Autocorrelation (SL) – The Spearman Rank Autocorrelation measures the monotonic relationship between daily precipitation values and their values on the preceding day (1-day lag). It is computed using the ranked values of the precipitation time series. For a precipitation timeseries (with total n observations) $P = \{P_1, P_2, \ldots, P_n\}$ with $R(P_t)$ and $R(P_{t+1})$ being the ranks of the precipitation values at times t and $t+1$,

$$\text{SL} = 1 - \frac{6\sum_{t=1}^{n-1}\left(R(P_{t+1}) - R(P_t)\right)^2}{n\left(n^2-1\right)} \tag{C.4}$$

This measure helps analyse persistence in precipitation patterns and whether the temporal structure of precipitation events are preserved.

C.4 SUPPLEMENTARY FIGURES

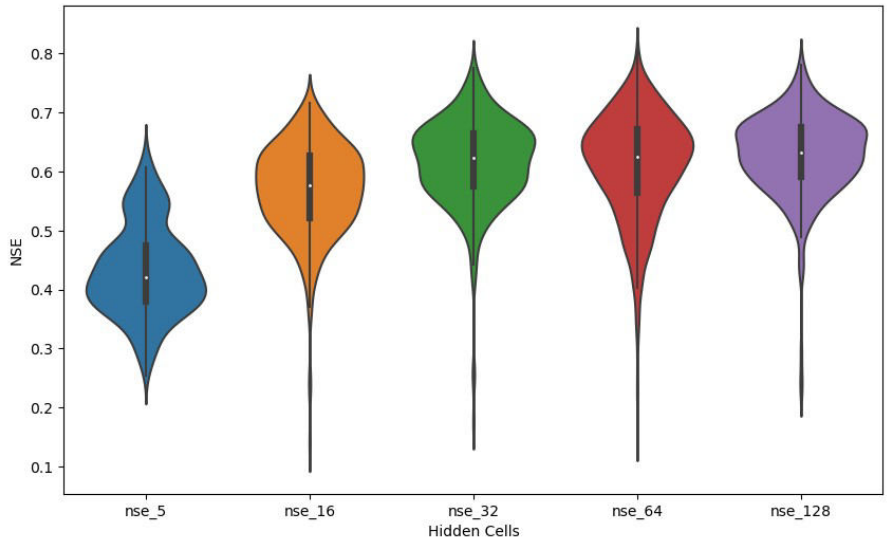

Figure C.2: Sensitivity analysis for the LSTM model by changing the number of cells in the hidden layer.

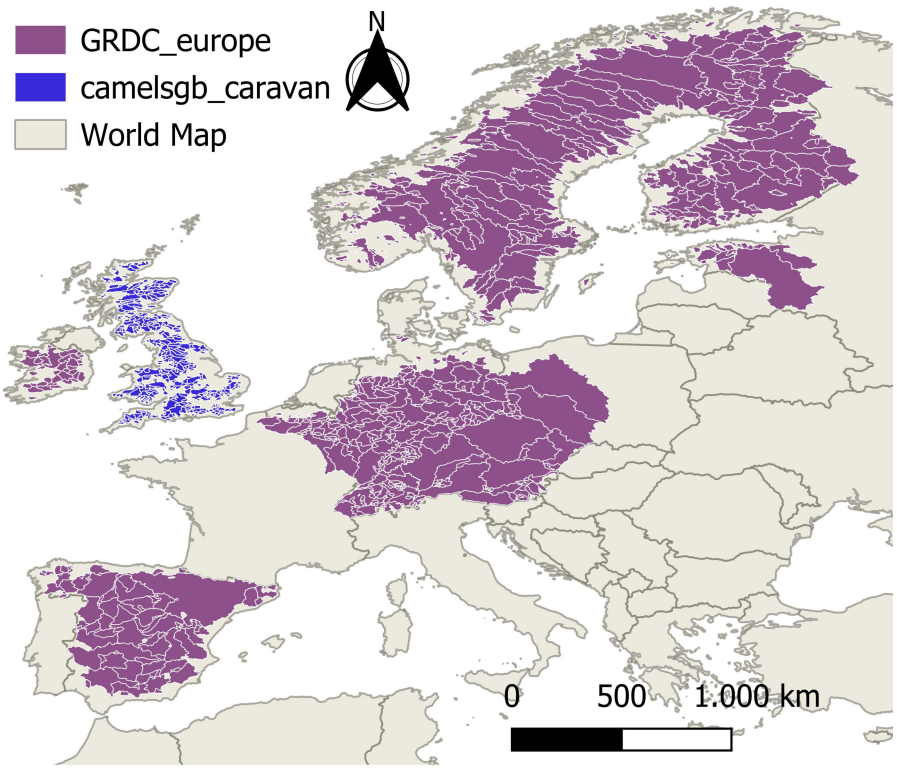

Figure C.3: Study area map depicting the catchments from GRDC and CAMELS-GB used for training the LSTM Model.

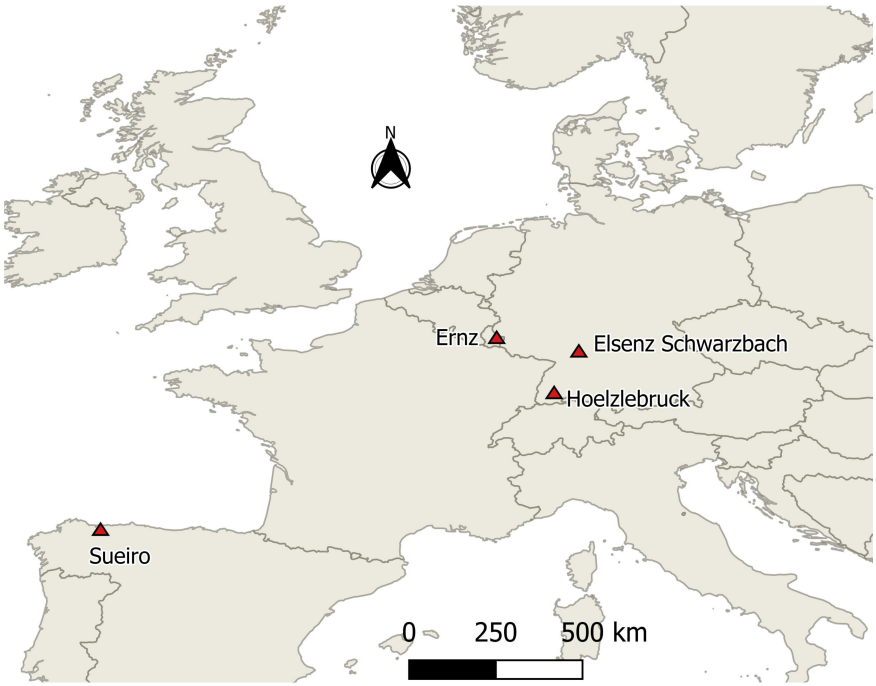

Figure C.4: Study area map depicting the four catchments (red triangles) used for out of sample testing.

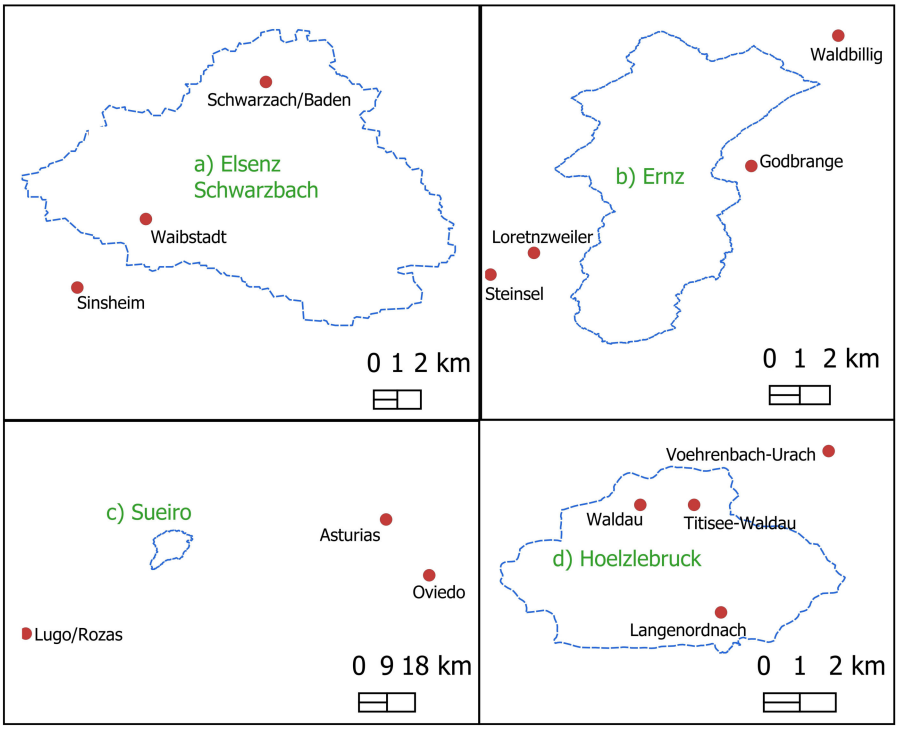

Figure C.5: Spatial maps showing the proximity of (red circles) observational stations (used for deriving the EOBS gridded product) to the four out of sample catchments considered in the present study.

Figure C.6: Catchments for the validation test cases using classical hydrological models

BIBLIOGRAPHY

Acuña Espinoza, E., R. Loritz, M. Álvarez Chaves, N. Bäuerle, and U. Ehret (2024). "To bucket or not to bucket? Analyzing the performance and interpretability of hybrid hydrological models with dynamic parameterization." In: *Hydrology and Earth System Sciences* 28.12, pp. 2705–2719. DOI: 10.5194/hess-28-2705-2024.

Adam, J. C., E. A. Clark, D. P. Lettenmaier, and E. F. Wood (2006). "Correction of Global Precipitation Products for Orographic Effects." In: *Journal of Climate* 19.1, pp. 15–38. DOI: 10.1175/JCLI3604.1.

Agarwal, A. et al. (2020). "Optimal design of hydrometric station networks based on complex network analysis." In: *Hydrology and Earth System Sciences* 24.5, pp. 2235–2251. DOI: 10.5194/hess-24-2235-2020.

Alexopoulos, M. J., H. Müller-Thomy, P. Nistahl, M. Šraj, and N. Bezak (2023). "Validation of precipitation reanalysis products for rainfall-runoff modelling in Slovenia." In: *Hydrology and Earth System Sciences* 27.13, pp. 2559–2578. DOI: 10.5194/hess-27-2559-2023.

Arias, P. et al. (2021). "Technical Summary." In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 33144. DOI: 10.1017/9781009157896.002.

Auerswald, K., J. Geist, J. N. Quinton, and P. Fiener (2025). "HESS Opinions: Floods and droughts – are land use, soil management, and landscape hydrology more significant drivers than increasing CO 2 ?" In: *Hydrology and Earth System Sciences* 29.9, pp. 2185–2200. DOI: 10.5194/hess-29-2185-2025.

Ballabio, C., P. Panagos, and L. Monatanarella (2016). "Mapping topsoil physical properties at European scale using the LUCAS database." In: *Geoderma* 261, pp. 110–123. DOI: 10.1016/j.geoderma.2015.07.006.

Bandhauer, M. et al. (2022). "Evaluation of daily precipitation analyses in E-OBS (v19.0e) and ERA5 by comparison to regional high-resolution datasets in European regions." In: *International Journal of Climatology* 42.2, pp. 727–747. DOI: 10.1002/joc.7269.

Bárdossy, A., F. Anwar, and J. Seidel (2020). "Hydrological Modelling in Data Sparse Environment: Inverse Modelling of a Historical Flood Event." In: *Water (Switzerland)* 12.11. DOI: 10.3390/w12113242.

Bárdossy, A., C. Kilsby, S. Birkinshaw, N. Wang, and F. Anwar (2022). "Is Precipitation Responsible for the Most Hydrological Model

Uncertainty?" In: *Frontiers in Water* 4.March, pp. 1–17. DOI: 10.3389/frwa.2022.836554.

Baste, S., D. Klotz, E. Acuña Espinoza, A. Bardossy, and R. Loritz (2025). "Unveiling the limits of deep learning models in hydrological extrapolation tasks." In: *Hydrology and Earth System Sciences* 29.21, pp. 5871–5891. DOI: 10.5194/hess-29-5871-2025.

Beauchamp, J., R. Leconte, M. Trudel, and F. Brissette (2013). "Estimation of the summer-fall PMP and PMF of a northern watershed under a changed climate." In: *Water Resources Research* 49 (6), pp. 3852–3862. DOI: 10.1002/wrcr.20336.

Beck, H. E. et al. (2018). "Present and future köppen-geiger climate classification maps at 1-km resolution." In: *Scientific Data* 5. DOI: 10.1038/sdata.2018.214.

Berghuijs, W. R., S. T. Allen, S. Harrigan, and J. W. Kirchner (2019). "Growing Spatial Scales of Synchronous River Flooding in Europe." In: *Geophysical Research Letters* 46.3, pp. 1423–1428. DOI: 10.1029/2018GL081883.

Bergström, S. and A. Forsman (1973). "Development of a Conceptual Deterministic Rainfall-Runoff Model." In: *Nordic Hydrol.* 4.3 (1973), pp. 147–170. DOI: 10.2166/nh.1973.0012.

Beven, K. (2020). "Deep learning, hydrological processes and the uniqueness of place." In: *Hydrological Processes* 34.16, pp. 3608–3613. DOI: https://doi.org/10.1002/hyp.13805.

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

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York : Springer, [2006] ©2006.

Blöschl, G. et al. (2016). "The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory." In: *Hydrology and Earth System Sciences* 20 (1), pp. 227–255. DOI: 10.5194/hess-20-227-2016.

Blöschl, G., R. Merz, and C. Reszler (2007). "Floods in Austria." In: *Extreme Hydrological Events: New Concepts for Security*. Springer Netherlands, pp. 81–90. DOI: 10.1007/978-1-4020-5741-0_6.

Blöschl, G. et al. (2019a). "Changing climate both increases and decreases European river floods." In: *Nature* 573 (7772), pp. 108–111. DOI: 10.1038/s41586-019-1495-6.

Blöschl, G. et al. (2019b). "Twenty-three unsolved problems in hydrology (UPH) – a community perspective." In: *Hydrological Sciences Journal* 64.10, pp. 1141–1158. DOI: 10.1080/02626667.2019.1620507.

Blöschl, G. et al. (2020). "Current European flood-rich period exceptional compared with past 500 years." In: *Nature* 583 (7817), pp. 560–566. DOI: 10.1038/s41586-020-2478-3.

Boeing, F. et al. (2022). "High-resolution drought simulations and comparison to soil moisture observations in Germany." In: *Hydrology*

and Earth System Sciences 26.19, pp. 5137–5161. DOI: 10.5194/hess-26-5137-2022.

Boettiger, C. (2015). "An introduction to Docker for reproducible research." In: *ACM SIGOPS Operating Systems Review* 49.1, pp. 71–79. DOI: 10.1145/2723872.2723882.

Borga, M., E. Gaume, J. D. Creutin, and L. Marchi (2008). "Surveying flash floods: gauging the ungauged extremes." In: *Hydrological Processes* 22 (18), pp. 3883–3885. DOI: 10.1002/hyp.7111.

Brocca, L., T. Moramarco, F. Melone, and W. Wagner (2013). "A new method for rainfall estimation through soil moisture observations." In: *Geophysical Research Letters* 40.5, pp. 853–858. DOI: 10.1002/grl.50173.

Bronstert, A., D. Niehoff, and G. R. Schiffler (2023). "Modelling infiltration and infiltration excess: The importance of fast and local processes." In: *Hydrological Processes* 37 (4), pp. 1–20. DOI: 10.1002/hyp.14875.

Bronstert, A. et al. (2018). "Forensic hydro-meteorological analysis of an extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany." In: *Science of the Total Environment* 630, pp. 977–991. DOI: 10.1016/j.scitotenv.2018.02.241.

Brown, C. F. et al. (2022). "Dynamic World, Near real-time global 10 m land use land cover mapping." In: *Scientific Data* 9.1, pp. 1–17. DOI: 10.1038/s41597-022-01307-4.

Brunner, M. I., L. Slater, L. M. Tallaksen, and M. Clark (2021). "Challenges in modeling and predicting floods and droughts: A review." In: *Wiley Interdisciplinary Reviews: Water* 8.3, pp. 1–32. DOI: 10.1002/wat2.1520.

Brus, D. J. and J. J. H. van den Akker (2018). "How serious a problem is subsoil compaction in the Netherlands? A survey based on probability sampling." In: *SOIL* 4.1, pp. 37–45. DOI: 10.5194/soil-4-37-2018.

Bundesamt für Naturschutz (BfN) (2026). *Nördlicher Kraichgau. Landschaftssteckbrief (Landschafts-ID 12501)*. Bundesamt für Naturschutz (BfN). URL: https://www.bfn.de/landschaftssteckbriefe/noerdlicher-kraichgau.

Bürger, G., A. Pfister, and A. Bronstert (2019). "Temperature-Driven Rise in Extreme Sub-Hourly Rainfall." In: *Journal of Climate* 32 (22), pp. 7597–7609. DOI: 10.1175/JCLI-D-19-0136.1.

Casado Rodríguez, J. (2023). *CAMELS-ES: Catchment Attributes and Meteorology for Large-Sample Studies – Spain (1.0.2)*. Version 1.0.2. DOI: 10.5281/zenodo.8428374.

Caviedes-Voullième, D., M. Morales-Hernández, M. R. Norman, and I. Özgen-Xian (2023). "SERGHEI (SERGHEI-SWE) v1.0: a performance-portable high-performance parallel-computing shallow-water solver for hydrology and environmental hy-

draulics." In: *Geoscientific Model Development* 16.3, pp. 977–1008. DOI: 10.5194/gmd-16-977-2023.

Celia, M. A., E. T. Bouloutas, and R. 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.

Chan, S. et al. (2018). "Development and assessment of the SMAP enhanced passive soil moisture product." In: *Remote Sensing of Environment* 204, pp. 931–941. DOI: 10.1016/j.rse.2017.08.025.

Chen, C. T. and W. D. Chang (1996). "A feedforward neural network with function shape autotuning." In: *Neural Networks* 9.4, pp. 627–641. DOI: 10.1016/0893-6080(96)00006-8.

Chen, J. et al. (2021). "Magnitude amplification of flash floods caused by large woody in Keze gully in Jiuzhaigou National Park, China." In: *Geomatics, Natural Hazards and Risk* 12 (1), pp. 2277–2299. DOI: 10.1080/19475705.2021.1961882.

Clerc-Schwarzenbach, F. et al. (2024). "Large-sample hydrology – a few camels or a whole caravan?" In: *Hydrology and Earth System Sciences* 28.17, pp. 4219–4237. DOI: 10.5194/hess-28-4219-2024.

Cochrane, T. and D. C. Flanagan (2003). "Representative hillslope methods for applying the WEPP model with DEMs and GIS." In: *Transactions of the ASAE* 46.4, p. 1041. DOI: 10.13031/2013.13966.

Collier, C. G. (2007). "Flash flood forecasting: What are the limits of predictability?" In: *Quarterly Journal of the Royal Meteorological Society* 133.622, pp. 3–23. DOI: 10.1002/qj.29.

Cornes, R. C., G. van der Schrier, E. J. M. van den Besselaar, and P. D. Jones (2018). "An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets." In: *Journal of Geophysical Research: Atmospheres* 123.17, pp. 9391–9409. DOI: 10.1029/2017JD028200.

Coxon, G. et al. (2020). "CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain." In: *Earth System Science Data* 12.4, pp. 2459–2483. DOI: 10.5194/essd-12-2459-2020.

D. N. Moriasi et al. (2007). "Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations." In: *Transactions of the ASABE* 50.3, pp. 885–900. DOI: 10.13031/2013.23153.

DWD (2021). *RADOLAN*. URL: https://www.dwd.de/DE/leistungen/radolan/radolan.html.

DWD (2023). *KOSTRA-DWD*. URL: https://www.dwd.de/DE/leistungen/kostra_dwd_rasterwerte/kostra_dwd_rasterwerte.html.

DWD (2024). *HYRAS-DE-PRE: Raster data set of daily sums of precipitation in mm for Germany – HYRAS-DE-PRE*. URL: https://opendata.dwd.de/climate_environment/CDC/grids_germany/daily/hyras_de/precipitation/DESCRIPTION_HYRAS-DE-PR_V6-1_en.pdf.

Delbrück, M. (1997). "Großflächiges Bromidtracerexperiment zur räumlichen und zeitlichen Variabilität des Wassertransports an einem Lößhang." Dissertation. Universität Heidelberg.

Disse, M. and H. Engel (2001). "Flood events in the Rhine basin: Genesis, influences and mitigation." In: *Natural Hazards* 23.2-3, pp. 271–290. DOI: 10.1023/A:1011142402374.

Dolich, A., A. Maharjan, M. Mälicke, A. Manoj J, and R. Loritz (2025). *Caravan-DE: Caravan extension Germany - German dataset for large-sample hydrology*. DOI: 10.5281/zenodo.14755229.

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

Drusch, M. et al. (2012). "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." In: *Remote Sensing of Environment* 120, pp. 25–36. DOI: 10.1016/j.rse.2011.11.026.

Dunne, T. (1978). "Field studies of hillslope flow processes - 'Hillslope Hydrology'." In: ed. by M. J. Kirkby. John Wiley & Sons, pp. 227–293.

Essou, G. R., F. Sabarly, P. Lucas-Picher, F. Brissette, and A. Poulin (2016). "Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling?" In: *Journal of Hydrometeorology* 17.7, pp. 1929–1950. DOI: 10.1175/JHM-D-15-0138.1.

FAO/UNESCO (1988). *Soil Map of the World*.

Fan, Y. et al. (2019). "Hillslope Hydrology in Global Change Research and Earth System Modeling." In: *Water Resources Research* 55.2, pp. 1737–1772. DOI: 10.1029/2018WR023903.

Färber, C. et al. (2023). "GRDC-Caravan: extending the original dataset with data from the Global Runoff Data Centre." In: *Zenodo*. DOI: 10.5281/zenodo.10074416.

Fatichi, S. et al. (2016). "An overview of current applications, challenges, and future trends in distributed process-based models in hydrology." In: *Journal of Hydrology* 537, pp. 45–60. DOI: 10.1016/j.jhydrol.2016.03.026.

Fenton, J. D. (2019). "Flood routing methods." In: *Journal of Hydrology* 570.January, pp. 251–264. DOI: 10.1016/j.jhydrol.2019.01.006.

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.

Fujihara, Y., S. P. Simonovic, F. Topaloğlu, K. Tanaka, and T. Watanabe (2008). "An inverse-modelling approach to assess the impacts of climate change in the Seyhan River basin, Turkey." In: *Hydrological Sciences Journal* 53.6, pp. 1121–1136. DOI: 10.1623/hysj.53.6.1121.

Gelaro, R. et al. (2017). "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)." In: *Journal of Climate* 30.14, pp. 5419–5454. DOI: 10.1175/JCLI-D-16-0758.1.

van Genuchten, M. T. (1980). "A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils." In: *Soil Science Society of America Journal* 44.5, pp. 892–898. DOI: 10.2136/sssaj1980.03615995004400050002x.

Global Modeling and Assimilation Office (2015). *MERRA-2 tavg1_2d_lnd_Nx: 2d, 1-Hourly, Time-Averaged, Single-Level, Assimilation, Land Surface Diagnostics V5.12.4*. DOI: 10.5067/RKPHT8KC1Y1T.

Goteti, G. and J. Famiglietti (2024). "Extent of gross underestimation of precipitation in India." In: *Hydrology and Earth System Sciences* 28.14, pp. 3435–3455. DOI: 10.5194/hess-28-3435-2024.

Gu, L. et al. (2023). "How well do the multi-satellite and atmospheric reanalysis products perform in hydrological modelling." In: *Journal of Hydrology* 617, p. 128920. DOI: https://doi.org/10.1016/j.jhydrol.2022.128920.

Guillaume, B., A. Michez, and A. Degré (2025). "Leveraging soil diversity to mitigate hydrological extremes with nature-based solutions in productive catchments." In: pp. 4661–4688.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. 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.

Göppert, H (2018). "Evaluation of expired Heavy Rain Events via Radar Measurements." In: *Wasserwirtschaft* 108 (11), pp. 44–50. DOI: 10.1007/s35147-018-0223-8.

Güntner, A. and A. Bronstert (2004). "Representation of landscape variability and lateral redistribution processes for large-scale hydrological modelling in semi-arid areas." In: *Journal of Hydrology* 297 (1-4), pp. 136–161. DOI: 10.1016/j.jhydrol.2004.04.008.

Hager, W. H. (2015). "Albert Strickler: His Life and Work." In: *Journal of Hydraulic Engineering* 141 (7), pp. 1–5. DOI: 10.1061/(ASCE)HY.1943-7900.0001000.

Heldmyer, A. et al. (2022). "Evaluation of a new observationally based channel parameterization for the National Water Model." In: *Hydrology and Earth System Sciences* 26.23, pp. 6121–6136. DOI: 10.5194/hess-26-6121-2022.

Henkel, P. (1994). *Der Mensch und das Schalten und Walten der Natur*. Newspaper article - Frankfurter Rundschau on 29.07.1994.

Ho, S. Q.-G. and U. Ehret (2025). "Is drought protection possible without compromising flood protection? Estimating the potential dual-use benefit of small flood reservoirs in southern Germany." In: *Hydrology and Earth System Sciences* 29.13, pp. 2785–2810. DOI: 10.5194/hess-29-2785-2025.

Hochreiter, S. (1998). "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions." In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 06.02, pp. 107–116. DOI: 10.1142/S0218488598000094.

Höge, M. et al. (2023). "CAMELS-CH: hydro-meteorological time series and landscape attributes for 331 catchments in hydrologic Switzerland." In: *Earth System Science Data* 15.12, pp. 5755–5784. DOI: 10.5194/essd-15-5755-2023.

Hornik, K., M. Stinchcombe, and H. White (1989). "Multilayer feedforward networks are universal approximators." In: *Neural Networks* 2.5, pp. 359–366. DOI: https://doi.org/10.1016/0893-6080(89)90020-8.

Horton, R. E. (1933). "The Role of infiltration in the hydrologic cycle." In: *Eos, Transactions American Geophysical Union* 14.1, pp. 446–460. DOI: 10.1029/TR014i001p00446.

Hrachowitz, M. 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.

Hrachowitz, M. and M. 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.

Hundecha, Y. and A. Bárdossy (2004). "Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model." In: *Journal of Hydrology* 292 (1-4), pp. 281–295. DOI: 10.1016/j.jhydrol.2004.01.002.

IPCC (2021). "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen," in: *Cambridge University Press* (In Press), p. 3949.

Ihringer, J (1994). "Aufbau und Anwendung des Flußgebietsmodells "FGM"." In: *DVWK Fortbildungslehrgang Hydrologie "Niederschlag-Abfluß-Modelle Für Kleine Einzugsgebiete Und Ihre Anwendung*.

Jarvis, P. G. (1976). "The Interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field Author (s): P . G . Jarvis Source : Philosophical Transactions of the Royal Society of London . Series B , Biological Sciences , Publish." In: *Philosophical Transactions of the Royal Society of London* 273 (927), pp. 593–610.

Jones, J. P., E. A. Sudicky, A. E. Brookfield, and Y. Park (2006). "An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow." In: *Water Resources Research* 42.2. DOI: 10.1029/2005WR004130.

Jongman, B., P. J. Ward, and J. C. Aerts (2012). "Global exposure to river and coastal flooding: Long term trends and changes." In: *Global Environmental Change* 22.4, pp. 823–835. DOI: 10.1016/j.gloenvcha.2012.07.004.

Junghänel, T, H Ertel, and T Deutschländer (2017). "KOSTRA-DWD-2010R. Bericht zur Revision der koordinierten Starkregenregionalisierung und-auswertung des Deutschen Wetterdienstes in der Version 2010." In: *KOSTRA-DWD-2010R - Bericht zur Revision der koordinierten Starkregenregionalisierung und -auswertung des Deutschen Wetterdienstes in der Version 2010*.

Kachelmannwetter (2023). *Kachelmannwetter*. URL: https://kachelmannwetter.com/de.

Kirchner, J. W. (2009). "Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward." In: *Water Resources Research* 45.2, pp. 1–34. DOI: 10.1029/2008WR006912.

Kirkby, M. and A. Cerdà (2021). "Following the curve? Reviewing the physical basis of the SCS curve number method for estimating storm runoff." In: *Hydrological Processes* 35.11. DOI: 10.1002/hyp.14404.

Kling, H., M. Fuchs, and M. Paulin (2012). "Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios." In: *Journal of Hydrology* 424-425, pp. 264–277. DOI: 10.1016/j.jhydrol.2012.01.011.

Klotz, D. et al. (2022). "Uncertainty estimation with deep learning for rainfall-runoff modeling." In: *Hydrology and Earth System Sciences* 26.6, pp. 1673–1693. DOI: 10.5194/hess-26-1673-2022.

Koster, R. D. et al. (2009). "On the nature of soil moisture in land surface models." In: *Journal of Climate* 22 (16), pp. 4322–4335. DOI: 10.1175/2009JCLI2832.1.

Krabbenhoft, C. A. et al. (2022). "Assessing placement bias of the global river gauge network." In: *Nature Sustainability* 5.7, pp. 586–592. DOI: 10.1038/s41893-022-00873-0.

Kratzert, F., D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger (2018). "Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks." In: *Hydrol. Earth Syst. Sci* 22, pp. 6005–6022. DOI: 10.5194/hess-22-6005-2018.

Kratzert, F. et al. (2023). "Caravan - A global community dataset for large-sample hydrology." In: *Scientific Data* 10.1, p. 61. DOI: 10.1038/s41597-023-01975-w.

Kreibich, H. et al. (2022). "The challenge of unprecedented floods and droughts in risk management." In: *Nature* 608.7921, pp. 80–86. DOI: 10.1038/s41586-022-04917-5.

Kretzschmar, A., W. Tych, and N. A. Chappell (2014). "Reversing hydrology: Estimation of sub-hourly rainfall time-series from stream-

flow." In: *Environmental Modelling & Software* 60, pp. 290–301. DOI: 10.1016/j.envsoft.2014.06.017.

Krier, R. et al. (2012). "Inferring catchment precipitation by doing hydrology backward: A test in 24 small and mesoscale catchments in Luxembourg." In: *Water Resources Research* 48.10, pp. 1–15. DOI: 10.1029/2011WR010657.

LGRB) (2024). *Landesamt für Geologie Rohstuff und Bergbau (LGRB) Kartenviewer*.

LUBW (2007). *Arbeitshilfe zur DIN 19700 für Hochwasserrückhaltebecken in Baden-Württemberg*. Tech. rep.

LUBW (2023). *Extremereignis (T > 100 a) für Anlagensicherheit und Hochwasserschutz*. URL: https://registry.gdi-de.org/id/de.bw.lubw.mdk/EE363459-509A-4967-8A81-060A21D1E9D5.

LUBW (2025). *Abfluss-BW: Regionalisierte Abfluss-Kennwerte Baden-Württemberg*. URL: https://www.lubw.baden-wuerttemberg.de/wasser/hydrologische-abfluss-kennwerte.

Lavoisier, A. L. (1789). *Traité élémentaire de chimie (Elementary Treatise on Chemistry)*. French. Vol. 1. Paris: Chez Cuchet.

Li, B. et al. (2019). "Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges." In: *Water Resources Research* 55 (9), pp. 7564–7586. DOI: 10.1029/2018WR024618.

Lindsay, J. B. (2014). "The Whitebox Geospatial Analysis Tools Project and Open-Access GIS." In: *Proceedings of the GIS Research UK 22nd Annual Conference*. The University of Glasgow. DOI: 10.13140/RG.2.1.1010.8962.

Linke, S. et al. (2019). "Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution." In: *Scientific Data* 6.1, p. 283. DOI: 10.1038/s41597-019-0300-6.

Liu, L., Y. Ren, Z. Li, and L. Zhou (2025). "Enhancing hydrological modelling with reanalysis soil moisture: A data-driven approach for optimizing initial conditions through reanalysis integration." In: *Advances in Water Resources* 203.June, p. 105023. DOI: 10.1016/j.advwatres.2025.105023.

Loritz, R, M Bassiouni, A Hildebrandt, S. K. Hassler, and E Zehe (2022). "Leveraging sap flow data in a catchment-scale hybrid model to improve soil moisture and transpiration estimates." In: *Hydrology And Earth System Sciences* 26 (18), pp. 4757–4771. DOI: 10.5194/hess-26-4757-2022.

Loritz, R. et al. (2024a). "CAMELS-DE: hydro-meteorological time series and attributes for 1582 catchments in Germany." In: *Earth System Science Data* 16.12, pp. 5625–5642. DOI: 10.5194/essd-16-5625-2024.

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

Loritz, R. et al. (2024b). "Generalizing Tree–Level Sap Flow Across the European Continent." In: *Geophysical Research Letters* 51.6, e2023GL107350. DOI: https://doi.org/10.1029/2023GL107350.

Lumbroso, D. and E. Gaume (2012). "Reducing the uncertainty in indirect estimates of extreme flash flood discharges." In: *Journal of Hydrology* 414-415, pp. 16–30. DOI: 10.1016/j.jhydrol.2011.08.048.

Mälicke, M. (2024). "From method development to software integration: A comprehensive approach to geostatistical variogram uncertainty." PhD thesis. DOI: 10.5445/IR/1000172259.

Manoj J, A. (2025). *Ash-Manoj/lstm_backward: LSTM models for Manoj J et al. (2024)*. DOI: 10.5281/zenodo.14161027.

Manoj J, A. and A. Dolich (2025). *VForWaTer/tool_catflow: Release with option soil support (v0.9.5)*. DOI: https://doi.org/10.5281/zenodo.16949549.

Manoj J, A., R. K. Guntu, and A. Agarwal (2022). "Spatiotemporal dependence of soil moisture and precipitation over India." In: *Journal of Hydrology* 610 (May), p. 127898. DOI: 10.1016/j.jhydrol.2022.127898.

Manoj J, A., M. Mälicke, and A. Dolich (2025). *VForWaTer/tool_whiteboxgis: Updated input for mosaic tool*. DOI: 10.5281/zenodo.7708559.

Manoj J, A., T. Pérez Ciria, G. Chiogna, N. Salzmann, and A. Agarwal (2023). "Characterising the coincidence of soil moisture – precipitation extremes as a possible precursor to European floods." In: *Journal of Hydrology* 620 (September 2022). DOI: 10.1016/j.jhydrol.2023.129445.

Manoj J, A. et al. (2024). "Toward Flash Flood Modeling Using Gradient Resolving Representative Hillslopes." In: *Water Resources Research* 60.6. DOI: 10.1029/2023WR036420.

Marchi, L., M. Borga, E. Preciso, and E. Gaume (2010). "Characterisation of selected extreme flash floods in Europe and implications for flood risk management." In: *Journal of Hydrology* 394 (1-2), pp. 118–133. DOI: 10.1016/j.jhydrol.2010.07.017.

Marchi, L., M. Cavalli, W. Amponsah, M. Borga, and S. Crema (2016). "Upper limits of flash flood stream power in Europe." In: *Geomorphology* 272, pp. 68–77. DOI: 10.1016/j.geomorph.2015.11.005.

Maxwell, R. M. (2013). "A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling." In: *Advances in Water Resources* 53, pp. 109–117. DOI: 10.1016/j.advwatres.2012.10.001.

Merz, B. et al. (2021). "Causes, impacts and patterns of disastrous river floods." In: *Nature Reviews Earth and Environment* 0123456789. DOI: 10.1038/s43017-021-00195-3.

Meyer, J., M. Neuper, L. Mathias, E. Zehe, and L. Pfister (2022). "Atmospheric conditions favouring extreme precipitation and flash floods in temperate regions of Europe." In: *Hydrology and Earth System Sciences* 26.23, pp. 6163–6183. DOI: 10.5194/hess-26-6163-2022.

Michelon, A., L. Benoit, H. Beria, N. Ceperley, and B. Schaefli (2021). "Benefits from high-density rain gauge observations for hydrological response analysis in a small alpine catchment." In: *Hydrology and Earth System Sciences* 25.4, pp. 2301–2325. DOI: 10.5194/hess-25-2301-2021.

Milly, P. C. et al. (2008). "Climate change: Stationarity is dead: Whither water management?" In: *Science* 319.5863, pp. 573–574. DOI: 10.1126/science.1151915.

Miralles, D. G. et al. (2011). "Global land-surface evaporation estimated from satellite-based observations." In: *Hydrology and Earth System Sciences* 15.2, pp. 453–469. DOI: 10.5194/hess-15-453-2011.

Miser, H. J. (1980). "Operations Research and Systems Analysis." In: *Science* 209.4452, pp. 139–146. DOI: 10.1126/science.209.4452.139.

Mishra, A. et al. (2022). "An Overview of Flood Concepts, Challenges, and Future Directions." In: *Journal of Hydrologic Engineering* 27.6, pp. 1–30. DOI: 10.1061/(asce)he.1943-5584.0002164.

Mohr, S. et al. (2023). "A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe – Part 1: Event description and analysis." In: *Natural Hazards and Earth System Sciences* 23.2, pp. 525–551. DOI: 10.5194/nhess-23-525-2023.

Montanari, A., B. Merz, and G. Blöschl (2024). "HESS Opinions: The sword of Damocles of the impossible flood." In: *Hydrology and Earth System Sciences* 28.12, pp. 2603–2615. DOI: 10.5194/hess-28-2603-2024.

Montanari, A. et al. (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.

Montz, B. E. and E. Gruntfest (2002). "Flash flood mitigation: Recommendations for research and applications." In: *Environmental Hazards* 4.1, pp. 15–22. DOI: 10.3763/ehaz.2002.0402.

Mualem, Y. (1976). "A new model for predicting the hydraulic conductivity of unsaturated porous media." In: *Water Resources Research* 12.3, pp. 513–522. DOI: 10.1029/WR012i003p00513.

Mueller, E. N. and A. Pfister (2011). "Increasing occurrence of high-intensity rainstorm events relevant for the generation of soil erosion in a temperate lowland region in Central Europe." In: *Journal of Hydrology* 411.3-4, pp. 266–278. DOI: 10.1016/j.jhydrol.2011.10.005.

Munich Re (2016). *Munich Re*. URL: https://www.munichre.com/en/company/media-relations/media-information-and-corporate-

news/media-information/2016/2016-07-12-media-information.html.

Muñoz-Sabater, J. et al. (2021). "ERA5-Land: a state-of-the-art global reanalysis dataset for land applications." In: *Earth System Science Data* 13.9, pp. 4349–4383. DOI: 10.5194/essd-13-4349-2021.

Nash, J. and J. Sutcliffe (1970). "River flow forecasting through conceptual models part I — A discussion of principles." In: *Journal of Hydrology* 10.3, pp. 282–290. DOI: https://doi.org/10.1016/0022-1694(70)90255-6.

Nature Editors (2012). "Must try harder." In: *Nature* 483.7391, p. 509. DOI: 10.1038/483509a.

Nearing, G. S. and H. V. Gupta (2015). "The quantity and quality of information in hydrologic models." In: *Water Resources Research* 51.1, pp. 524–538. DOI: 10.1002/2014WR015895.

Nearing, G. 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.

Nearing, G. S. et al. (2021). "What Role Does Hydrological Science Play in the Age of Machine Learning?" In: *Water Resources Research* 57.3, e2020WR028091. DOI: https://doi.org/10.1029/2020WR028091.

Niehoff, D., U. Fritsch, and A. Bronstert (2002). "Land-use impacts on storm-runoff generation: scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany." In: *Journal of Hydrology* 267.1, pp. 80–93. DOI: https://doi.org/10.1016/S0022-1694(02)00142-7.

Nijzink, J. et al. (2025). "CAMELS-LUX: Highly Resolved Hydro-Meteorological and Atmospheric Data for Physiographically Characterized Catchments around Luxembourg." In: *Earth System Science Data Discussions* 2025, pp. 1–34. DOI: 10.5194/essd-2024-482.

Ongie, G. et al. (2020). "Deep Learning Techniques for Inverse Problems in Imaging." In: *IEEE Journal on Selected Areas in Information Theory* 1.1, pp. 39–56. DOI: 10.1109/jsait.2020.2991563.

Onogi, K. et al. (2007). "The JRA-25 Reanalysis." In: *Journal of the Meteorological Society of Japan. Ser. II* 85.3, pp. 369–432. DOI: 10.2151/jmsj.85.369.

Pall, P., M. R. Allen, and D. A. Stone (2007). "Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming." In: *Climate Dynamics* 28.4, pp. 351–363. DOI: 10.1007/s00382-006-0180-2.

Paniconi, C. and M. Putti (2015). "Physically based modeling in catchment hydrology at 50: Survey and outlook." In: *Water Resources Research* 51.9, pp. 7090–7129. DOI: 10.1002/2015WR017780.

Pérez Ciria, T., D. Labat, and G. Chiogna (2019). "Detection and interpretation of recent and historical streamflow alterations caused by river damming and hydropower production in the Adige and Inn river basins using continuous, discrete and multiresolution

wavelet analysis." In: *Journal of Hydrology* 578.August, p. 124021. DOI: 10.1016/j.jhydrol.2019.124021.

Pérez, A., R. Abrahão, J. Causapé, O. Cirpka, and C. 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., T. L. Hohenbrink, S. C. Iden, and W. Durner (2021). "A Simple Model to Predict Hydraulic Conductivity in Medium to Dry Soil From the Water Retention Curve." In: *Water Resources Research* 57.5. DOI: 10.1029/2020WR029211.

Pfister, L and J. W. Kirchner (2017). "DebatesHypothesis testing in hydrology: Theory and practice." In: *Water Resources Research* 53.3, pp. 1792–1798. DOI: 10.1002/2016wr020116.

Piper, D. et al. (2016). "Exceptional sequence of severe thunderstorms and related flash floods in May and June 2016 in Germany - Part 1: Meteorological background." In: *Natural Hazards and Earth System Sciences* 16.12, pp. 2835–2850. DOI: 10.5194/nhess-16-2835-2016.

Plate, E. J. and E. Zehe, eds. (2008). *Hydrologie und Stoffdynamik kleiner Einzugsgebiete*. Stuttgart, Germany: Schweizerbart Science Publishers.

Poschlod, B. (2022). "Attributing heavy rainfall event in Berchtesgadener Land to recent climate change – Further rainfall intensification projected for the future." In: *Weather and Climate Extremes* 38.February, p. 100492. DOI: 10.1016/j.wace.2022.100492.

Rakovec, O. et al. (2022). "The 2018–2020 Multi-Year Drought Sets a New Benchmark in Europe." In: *Earth's Future* 10.3, e2021EF002394. DOI: https://doi.org/10.1029/2021EF002394.

Rauthe, M., H. Steiner, U. Riediger, A. Mazurkiewicz, and A. Gratzki (2013). "A Central European precipitation climatology - Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS)." In: *Meteorologische Zeitschrift* 22.3, pp. 235–256. DOI: 10.1127/0941-2948/2013/0436.

Reggiani, P. and J. Schellekens (2003). "Modelling of hydrological responses: the representative elementary watershed approach as an alternative blueprint for watershed modelling." In: *Hydrological Processes* 17.18, pp. 3785–3789. DOI: https://doi.org/10.1002/hyp.5167.

Remmers, J. O. E. et al. (2025). "HESS Opinions: Reflecting and acting on the social aspects of modeling." In: *Hydrology and Earth System Sciences* 29.20, pp. 5371–5382. DOI: 10.5194/hess-29-5371-2025.

Remson, I, G. M. Hornberger, and F. J. Molz (1971). *Numerical Methods in Subsurface Hydrology*. With an Introduction to the Finite Element Method. Wiley-Interscience.

Richet, J.-B., J.-F. Ouvry, and M. Saunier (2017). "The role of vegetative barriers such as fascines and dense shrub hedges in catchment management to reduce runoff and erosion effects: Experimen-

tal evidence of efficiency, and conditions of use." In: *Ecological Engineering* 103, pp. 455–469. DOI: 10.1016/j.ecoleng.2016.08.008.

Rienecker, M. M. et al. (2011). "MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications." In: *Journal of Climate* 24.14, pp. 3624–3648. DOI: 10.1175/JCLI-D-11-00015.1.

Rosier, I., J. Diels, B. Somers, and J. Van Orshoven (2023). "The impact of vegetated landscape elements on runoff in a small agricultural watershed: A modelling study." In: *Journal of Hydrology* 617.PC, p. 129144. DOI: 10.1016/j.jhydrol.2023.129144.

Rosier, I., J. Diels, B. Somers, and J. Van Orshoven (2024). "Maximising runoff retention by vegetated landscape elements positioned through spatial optimisation." In: *Landscape and Urban Planning* 243.December 2023, p. 104968. DOI: 10.1016/j.landurbplan.2023.104968.

Roy, R., S. Hinduja, and R. Teti (2008). "Recent advances in engineering design optimisation: Challenges and future trends." In: *CIRP Annals* 57.2, pp. 697–715. DOI: https://doi.org/10.1016/j.cirp.2008.09.007.

Ruiter, M. C. et al. (2020). "Why We Can No Longer Ignore Consecutive Disasters." In: *Earth's Future* 8.3. DOI: 10.1029/2019EF001425.

Ruiz-Villanueva, V et al. (2012). "Extreme flood response to short-duration convective rainfall in South-West Germany." In: *Hydrology and Earth System Sciences* 16.5, pp. 1543–1559. DOI: 10.5194/hess-16-1543-2012.

Sarhadi, A., M. C. Ausín, M. P. Wiper, D. Touma, and N. S. Diffenbaugh (2018). "Multidimensional risk in a nonstationary climate: Joint probability of increasingly severe warm and dry conditions." In: *Science Advances* 4.11, eaau3487. DOI: 10.1126/sciadv.aau3487.

Schäfer, D (1999). *Bodenhydraulischen Funktionen eines Kleineinzugsgebiets –Vergleich und Bewertung unterschiedlicher Verfahren.* Dissertation, Institute of Hydromechanics, University of Karlsruhe, Germany.

Schalko, I., L. Schmocker, V. Weitbrecht, and R. M. Boes (2018). "Backwater Rise due to Large Wood Accumulations." In: *Journal of Hydraulic Engineering* 144.9, pp. 1–13. DOI: 10.1061/(asce)hy.1943-7900.0001501.

Scherer, U., E. Zehe, K. Träbing, and K. Gerlinger (2012). "Prediction of soil detachment in agricultural loess catchments: Model development and parameterisation." In: *Catena* 90, pp. 63–75. DOI: 10.1016/j.catena.2011.11.003.

Schierholz, I, D Schafer, 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, F. and A. Don (2019). "Root-restricting layers in German agricultural soils. Part I: extent and cause." In: *Plant and Soil* 442.1-2, pp. 433–451. DOI: 10.1007/s11104-019-04185-9.

Schroers, S et al. (2022). "Morphological controls on surface runoff: an interpretation of steady-state energy patterns, maximum power states and dissipation regimes within a thermodynamic framework." In: *Hydrology And Earth System Sciences* 26.12, pp. 3125–3150. DOI: 10.5194/hess-26-3125-2022.

Seibert, J (2005). "HBV light." In: *HBV light version 2 User's Manual* November.

Seidel, J., F. Imbery, P. Dostal, D. Sudhaus, and K. Burger (2009). "Potential of historical meteorological and hydrological data for the reconstruction of historical flood events-the example of the 1882 flood in southwest Germany." In: *Natural Hazards and Earth System Science* 9.1, pp. 175–183. DOI: 10.5194/nhess-9-175-2009.

Seneviratne, S. I. et al. (2010). "Investigating soil moisture-climate interactions in a changing climate: A review." In: *Earth-Science Reviews* 99.3-4, pp. 125–161. DOI: 10.1016/j.earscirev.2010.02.004.

Shen, C. et al. (2023). "Differentiable modelling to unify machine learning and physical models for geosciences." In: *Nature Reviews Earth and Environment* 4.8, pp. 552–567. DOI: 10.1038/s43017-023-00450-9.

Sivapalan, M. 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.

Sivapalan, M. (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.

Spreitzer, G., J. Tunnicliffe, and H. Friedrich (2019). "Using Structure from Motion photogrammetry to assess large wood (LW) accumulations in the field." In: *Geomorphology* 346, p. 106851. DOI: 10.1016/j.geomorph.2019.106851.

Steinbrich, A., H. Leistert, and M. Weiler (2016). "Model-based quantification of runoff generation processes at high spatial and temporal resolution." In: *Environmental Earth Sciences* 75.21, pp. 1–16. DOI: 10.1007/s12665-016-6234-9.

Strahler, A. N. (1957). "Quantitative analysis of watershed geomorphology." In: *Eos, Transactions American Geophysical Union* 38.6, pp. 913–920. DOI: 10.1029/TR038i006p00913.

Sun, S. and J. L. Bertrand-Krajewski (2013). "Separately accounting for uncertainties in rainfall and runoff: Calibration of event-based conceptual hydrological models in small urban catchments using Bayesian method." In: *Water Resources Research* 49.9, pp. 5381–5394. DOI: 10.1002/wrcr.20444.

Suresh, S. and F. Hossain (2025). "Has Hydropower Made the World More Flood-Prone?" In: *Earth's Future* 13.9, e2025EF006648. DOI: https://doi.org/10.1029/2025EF006648.

Szabó, B., M. Weynants, and T. K. Weber (2021). "Updated European hydraulic pedotransfer functions with communicated uncertainties in the predicted variables (euptfv2)." In: *Geoscientific Model Development* 14.1, pp. 151–175. DOI: 10.5194/gmd-14-151-2021.

Tang, Y. K. and R. W. Skaggs (1977). "Experimental evaluation of theoretical solutions for subsurface drainage and irrigation." In: *Water Resources Research* 13.6, pp. 957–965. DOI: 10.1029/WR013i006p00957.

Tarek, M., F. P. Brissette, and R. Arsenault (2020). "Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America." In: *Hydrology and Earth System Sciences* 24.5, pp. 2527–2544. DOI: 10.5194/hess-24-2527-2020.

Taszarek, M., J. T. Allen, M. Marchio, and H. E. Brooks (2021). "Global climatology and trends in convective environments from ERA5 and rawinsonde data." In: *npj Climate and Atmospheric Science* 4.1, pp. 1–11. DOI: 10.1038/s41612-021-00190-x.

Terzaghi, K., R. B. Peck, and G. Mesri (1996). *Soil Mechanics in Engineering Practice*. 3rd ed. Wiley-Interscience publication. John Wiley & Sons. 592 pp.

Tetzlaff, D., S. K. Carey, J. P. McNamara, H. Laudon, and C. Soulsby (2017). "The essential value of long-term experimental data for hydrology and water management." In: *Water Resources Research* 53.4, pp. 2598–2604. DOI: 10.1002/2017WR020838.

Teuling, A. J., I. Lehner, J. W. Kirchner, and S. I. Seneviratne (2010). "Catchments as simple dynamical systems: Experience from a Swiss prealpine catchment." In: *Water Resources Research* 46.10, pp. 1–15. DOI: 10.1029/2009WR008777.

Troch, P. A., C. 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. DOI: 10.1029/2002WR001728.

Troy, T. J., N. Devineni, C. H. Lima, and U. Lall (2025). "Can runoff modeled at coarse resolution simulate floods at finer resolutions? A case study over the Ohio River Basin." In: *Advances in Water Resources* 206.August 2024, p. 105151. DOI: 10.1016/j.advwatres.2025.105151.

United States Geological Survey (1989). *Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains*. Tech. rep.

United States Geological Survey (2022). *The Water Cycle*.

Vahrenkamp, N., T. Asfour, and R. Dillmann (2013). "Robot placement based on reachability inversion." In: *2013 IEEE International Conference on Robotics and Automation*, pp. 1970–1975. DOI: 10.1109/ICRA.2013.6630839.

Ven Te Chow (1959). *Open Channel Hydraulics*. McGraw-Hill Professional.

Villinger, F., R. Loritz, and E. Zehe (2022). "Physikalisch-basierte Simulation einer abgelaufenen Sturzflut mittels "repräsentativer Hänge" in einem ländlichen Einzugsgebiet." In: *Hydrologie & Wasserbewirtschaftung* 66.6. DOI: 10.5675/HyWa_2022.6_1.

Vogt, H.-J. (2024). *Das Hochwasser in Neckarbischofsheim 1994*. URL: https://www.youtube.com/watch?v=cPxJu-9h17A.

Wasko, C. et al. (2021). "Incorporating climate change in flood estimation guidance." In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2195. DOI: 10.1098/rsta.2019.0548.

Weiler, M. (2005). "An infiltration model based on flow variability in macropores: development, sensitivity analysis and applications." In: *Journal of Hydrology* 310.1, pp. 294–315. DOI: https://doi.org/10.1016/j.jhydrol.2005.01.010.

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.

Wikipedia contributors (2025). *Kraichgau*. URL: https://de.wikipedia.org/wiki/Kraichgau (visited on 01/05/2026).

World Climate Research program (WCRP) (2021). *Expert Team on Climate Change Detection and Indices (ETCCDI)*. URL: https://www.wcrp-climate.org/etccdi.

Xu, C., W. Wang, Y. Hu, and Y. Liu (2024). "Evaluation of ERA5, ERA5-Land, GLDAS-2.1, and GLEAM potential evapotranspiration data over mainland China." In: *Journal of Hydrology: Regional Studies* 51, p. 101651. DOI: https://doi.org/10.1016/j.ejrh.2023.101651.

Yevjevich, V. (1974). "Determinism and stochasticity in hydrology." In: *Journal of Hydrology* 22.3-4, pp. 225–238. DOI: 10.1016/0022-1694(74)90078-X.

Yumnam, K., R. Kumar Guntu, M. Rathinasamy, and A. Agarwal (2022). "Quantile-based Bayesian Model Averaging approach towards merging of precipitation products." In: *Journal of Hydrology* 604.July 2021, p. 127206. DOI: 10.1016/j.jhydrol.2021.127206.

Zehe, E., T. Maurer, J. Ihringer, and E. 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, E. and M. Sivapalan (2009). "Threshold behaviour in hydrological systems as (human) geo-ecosystems: manifestations, controls, implications." In: *Hydrology and Earth System Sciences* 13.7, pp. 1273–1297. DOI: 10.5194/hess-13-1273-2009.

Zehe, E. 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, E. 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, E., R. Becker, A. Bárdossy, and E. 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, E. and G. Blöschl (2004). "Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions." In: *Water Resources Research* 40 (10), pp. 1–21. DOI: 10.1029/2003WR002869.

Zeimetz, F., B. Schaefli, G. Artigue, J. G. Hernández, and A. J. Schleiss (2018). "New Approach to Identifying Critical Initial Conditions for Extreme Flood Simulations in a Semicontinuous Simulation Framework." In: *Journal of Hydrologic Engineering* 23 (8), pp. 1–9. DOI: 10.1061/(asce)he.1943-5584.0001652.

Zhang, Y. and M. G. Schaap (2017). "Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3)." In: *Journal of Hydrology* 547, pp. 39–53. DOI: 10.1016/j.jhydrol.2017.01.004.

Zscheischler, J. et al. (2020). "A typology of compound weather and climate events." In: *Nature Reviews Earth & Environment* 1.7, pp. 333–347. DOI: 10.1038/s43017-020-0060-z.

Zweckverband Hochwasserschutz Elsenz-Schwarzbach (2016). *Elsenz-Schwarzbach Water Board*. URL: https://www.zvhws.de/.

ACKNOWLEDGMENTS

Whatever is good in this work is owed to the guidance, support, and wisdom of my mentors and teachers along the way; any shortcomings remain entirely my own.

First and foremost, I would like to thank all my colleagues at the Chair of Hydrology for their support over the years and for ensuring that I never felt far from home throughout the duration of my PhD. Doing science is never straightforward, and a healthy workspace that allows you to express yourself freely and openly makes all the difference. I am grateful to the German Research Foundation for funding my doctoral stay in Germany, because enthusiasm for science alone does not quite cover the groceries at Edeka. I also acknowledge all my research collaborators, editors, and peer reviewers for the time and effort they invested in strengthening my work and advancing science, even in times of uncertainty.

My deepest gratitude belongs to my family, whose love and support have shaped me throughout my life, my mother, who always saw potential where I could not and never stopped believing in me; my sister, whose patience was always boundless; and my (late) father, whose quiet, practical wisdom continues to guide me. I also thank all my friends in Europe and India for the endless adventures I was privileged to share with them during my PhD days.

Lastly (but most importantly), I owe this thesis to my supervisor, Prof. Dr. Erwin Zehe. None of this would have been possible without your constant motivation and guidance. You were the best mentor I could ask for during some of the most darkest days of my life. You are one of the reasons I am still eager to continue in academia, I look forward to share your wisdom, passion, and excitement for science with my future students (though perhaps sparing them your love for nuances of administration). Vielen vielen Dank, Erwin!

OWN PUBLICATIONS

FIRST AUTHOR; PEER-REVIEWED INTERNATIONAL PUBLICATIONS

Manoj J, A., Loritz, R., Villinger, F., Mälicke, M., Koopaeidar, M., Göppert, H., and Zehe, E. (2024). "Toward flash flood modeling using gradient resolving representative hillslopes." *Water Resources Research*, 60, e2023WR036420. https://doi.org/10.1029/2023WR036420.

Manoj J, A., Loritz, R., Gupta, H., and Zehe, E. (2025) "Can discharge be used to inversely correct precipitation?" *Hydrology and Earth System Sciences* 29, pp. 6115-6135 https://doi.org/10.5194/hess-29-6115-2025.

CONFERENCE CONTRIBUTIONS (ORAL)

Manoj J, A., Loritz, R., Gupta, H., and Zehe, E. (2025): "Can discharge be used to inversely correct precipitation?", European Geosciences Union General Assembly 2025 (EGU25), Vienna, Austria.

Manoj J, A., Loritz, R., Gupta, H., and Zehe, E. (2024): "Can Data-Driven Models be Used to Correct Precipitation Forcing Data Inversely from Observed Discharge Time Series?", American Geophysical Union Fall Meeting 2024 (AGU24), Washington D.C., USA.

Manoj J, A., Loritz, R., Villinger, F., Mälicke, M., and E. Zehe (2023): "Towards Flash Flood Modeling in Data Scarce Catchments Using Representative Hillslopes in Gradient Resolving Models", American Geophysical Union Fall Meeting 2023 (AGU23), San Francisco, USA.

RESEARCH SOFTWARE

Manoj J, A. and Dolich, A. (2025). VForWaTer/tool_catflow: Preprocessing for the hydrological model CATFLOW. Repository available at https://github.com/VForWaTer/tool_catflow. Zenodo. https://doi.org/10.5281/zenodo.7861332

Manoj J, A., Mälicke, M., and Dolich, A. (2025). VForWaTer/-tool_whiteboxgis: GIS Preprocessing tool. Repository available at https://github.com/VForWaTer/tool_whiteboxgis. Zenodo. https://doi.org/10.5281/zenodo.7708559

DECLARATION

Authors contributions:

Chapter 2 Ashish Manoj J et al. (2024) "Toward flash flood modeling using gradient resolving representative hillslopes." *Water Resources Research*, 60, e2023WR036420. https://doi.org/10.1029/2023WR036420.

Ashish Manoj J (AMJ) conducted the hydrological modelling using CATFLOW, implemented the general methodological framework, and analysed the resulting data through the development of R and Python codes. AMJ prepared the manuscript, which was subsequently revised and improved by Erwin Zehe (EZ), Ralf Loritz (RL), and Franziska Villinger (FV). The final version of the manuscript was reviewed by all co-authors, and AMJ managed the manuscript throughout the review process.

Chapter 3 Ashish Manoj J et al. (Under Review) "Leveraging representative hillslopes for enhanced flood risk management in mesoscale catchments." *Hydrological Processes*, Preprint available at: https://doi.org/10.22541/au.176797104.47760019/v1.

Ashish Manoj J (AMJ) performed the model preprocessing and established the coupled model using CATFLOW, and analysed the resulting data using Python scripts. AMJ drafted the manuscript, which was later revised and enhanced by Erwin Zehe (EZ) and Franziska Villinger (FV). The final manuscript was reviewed and approved by all co-authors.

Chapter 4 Ashish Manoj J et al. (2025) "Can discharge be used to inversely correct precipitation?" *Hydrology and Earth System Sciences* 29, pp.6115–6135. https://doi.org/10.5194/hess-29-6115-2025.

Ashish Manoj J (AMJ) authored the manuscript, implemented the various models, and performed the analyses. Ralf Loritz (RL), Hoshin Gupta (HG), and Erwin Zehe (EZ) contributed to the development of the theoretical framework and supported the interpretation of the results and the revision of the manuscript. AMJ managed the manuscript throughout the review process.

Eidesstattliche Versicherung gemäß § 13 Absatz 2 Satz 1 Ziffer 4 der Promotionsordnung des Karlsruher Instituts für Technologie (KIT) für die KIT-Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften:

1. Bei der eingereichten Dissertation zu dem Thema *An Integrative System Approach to Flash Flood Hydrology* 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, 2026

Ashish Manoj Jaseetha

COLOPHON

This document was typeset using the typographical look-and-feel classicthesis developed by André Miede. 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/

Final Version as of May 28, 2026 (classicthesis v4.6).