

Emergence and Prediction of Heavy-Tailed Streamflow and Flood Distributions

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

Der

der Naturwissenschaftlichen Fakultät III

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von Herr Hsing-Jui Wang

geb. am 23 Juni 1987 in Tainan City

Promotionsausschuss:

Vorsitzender: Prof. Dr. Wolfgang Goßel

Gutachter: Prof. Dr. Ralf Merz

Gutachter: Prof. Dr. Kai Schröter

Tag der wissenschaftlichen Aussprache: 05. April 2024

Zusammenfassung

Hochwasser gehören weltweit zu den bedeutendsten Naturgefahren und stellen nach wie vor eine erhebliche Bedrohung für Leben und Eigentum dar. Die Vorhersage von starken oder extremen Überschwemmungen ist von essentieller Bedeutung für effektive Hochwasserschutzmaßnahmen und die Entwicklung von Anpassungsstrategien. Heavy-tail-Wahrscheinlichkeitsverteilungen von Hochwasserabflüssen, auch als endlastige Verteilungen oder Verteilungen mit schweren Rändern bekannt, sind ein wichtiger Indikator für das Auftreten solcher extremer Abflüsse. Das Ziel dieser Dissertation ist es, die Prozessursachen für das Auftreten von Heavy-tail-Hochwasserverteilungen zu analysieren und eine Methode zur zuverlässigen Bestimmung dieser Wahrscheinlichkeitsverteilungen zu entwickeln.

Der erste Teil dieser Arbeit untersucht die Rolle räumlicher Niederschlagsmuster für die Entstehung von endlastigen Abflussverteilungen. Mit Hilfe eines stochastischen hydrologischen Modells wurde die Abflussbildung einer Vielzahl synthetischer räumlich variabler Niederschlagsfelder in fünf deutschen Einzugsgebieten mit unterschiedlichen Klima- und Landschaftseigenschaften analysiert. Zusätzlich zu diesem modellgetriebenen Ansatz wurden auch beobachtete Niederschlags- und Abflussganglinien von 175 deutschen Einzugsgebieten untersucht. Die Ergebnisse beider Untersuchungsmethoden zeigen, dass eine zunehmende räumliche Niederschlagsvariabilität, insbesondere über einer bestimmten Schwelle, das Auftreten von endlastigen Abflussverteilungen verstärkt. Die Abflussbildung kann jedoch diesem Verhalten entgegenwirken und die Entstehung von extremen Abflüssen abfedern, was durch das Konzept der „Resilienz von Einzugsgebieten“ beschrieben werden kann. Die Ergebnisse dieser Arbeit zeigen, dass z.B. das Abflussrouting durch das Einzugsgebiet einen großen Einfluss hat, ob aus heterogenen Niederschlagsfeldern große Abflussmengen entstehen, wobei kleine und langgestreckte Einzugsgebiete eine geringere Resilienz aufweisen.

Im zweiten Teil dieser Dissertation wird ein neuer Index zur Beschreibung von Heavy-tail-Verhalten von Hochwasserverteilungen vorgestellt, der auf einer physikalischen Beschreibung der Abflussdynamik basiert. Der Index, abgeleitet aus der stochastischen Beschreibung des Niederschlags-Abflussverhaltens von Einzugsgebieten und aus dem Rezessionsverhalten täglicher Abflussganglinien ermittelt, erlaubt eine zuverlässige Beschreibung von Heavy-tail-Verhalten, auch bei kurzen Hochwasserreihen. Somit besitzt der Index einen entscheidenden Vorteil gegenüber bestehenden Indizes für Heavy-tail-Verhalten, denen einerseits eine physikalische Basis fehlt und die andererseits lange Datenreihen von Hochwasserabflüssen benötigen, um zuverlässig das Auftreten von Heavy-tail-Verteilungen vorhersagen zu können. Anhand von Beobachtungsdaten aus 98 deutschen Einzugsgebieten konnte gezeigt werden, dass der Index zuverlässig endlastige Verteilungen unter verschiedenen Klima- und Landschaftsbedingungen auch bei kurzen Beobachtungsreihen vorhersagen kann.

Im dritten Teil dieser Dissertation wird der neu entwickelte Indikator genutzt, um das Heavy-tail-Verhalten in einem Zusammenhang von Klima und Landschaftsbedingungen zu setzen. Die Analysen umfangreicher Datensätze aus Deutschland, dem Vereinigten Königreich (UK), Norwegen und den Vereinigten Staaten (US) zeigen unterschiedliche Muster. Deutschland und das Vereinigte Königreich weisen überwiegend endlastige Hochwasserverteilungen auf, während in den Vereinigten Staaten eine Mischung und in Norwegen hauptsächlich Verteilungen mit Light-tails gefunden wurden. Die zeitliche Variabilität in der Speicherfähigkeit innerhalb von Einzugsgebieten erweist sich als ein entscheidender Mechanismus, der stark nichtlineare Einzugsgebietsreaktionen begünstigt und damit das Auftreten einzelner extremer Überschwemmungen verstärkt, welche sich in Heavy-tail-Verteilungen widerspiegelt.

Wenn Hochwasserabflüsse in einem Gebiet durch die Mischung verschiedener Entstehungsprozesse, wie z.B. Schneeschmelze und Sommergewitter, entstehen, werden endlastige Hochwasserverteilungen wahrscheinlicher. Die gefundenen Muster weisen auf ein komplexes Zusammenspiel zwischen Klima, Landschaftsgegebenheiten und Überschwemmungsverhalten hin und unterstreichen die Nützlichkeit des neu entwickelten Index bei der Hochwassergefahrenbewertung.

Die Erkenntnisse dieser Dissertation eröffnen zahlreiche mögliche Anwendungen für eine zuverlässigere Erfassung und Beschreibung extremer Abflüsse in der Hochwasserbemessung, der Bewertung von Auswirkungen des Klimawandels und wasserwirtschaftlicher Maßnahmen auf das Hochwasserverhalten, die Klassifikation von Hochwassergefahren und der Entwicklung von Anpassungsstrategien.

Keywords Hochwässer, Heavy-Tail-Wahrscheinlichkeitsverteilungen, Hochwasserbemessung, Extremwertstatistik, Räumliche Niederschlagsvariabilität, Rezessionsverhalten täglicher Abflussganglinien

Abstract

Floods, recognized as one of the most pervasive natural hazards, continue to pose a significant global threat. The ability of societies to anticipate high or extreme floods is vital for effective hazard mitigation. Heavy-tailed behavior within flood frequency distributions serves as a critical indicator of this likelihood, emphasizing the importance of accurately identifying and understanding such behavior. This dissertation aims to advance the understanding of heavy-tailed streamflow and flood distributions and to introduce dependable methods for predicting and managing these phenomena within practical applications.

The dissertation comprises three main sections, along with an introduction and a concluding chapter summarizing the main findings and providing an outlook. The first section of this dissertation investigates the role of spatially variable rainfall in enhancing the tail heaviness of streamflow distributions. A diverse range of synthetic spatially variable rainfall inputs is generated and fed into a continuous probabilistic model of the catchment water transport to simulate streamflow in five German catchments with varying characteristics in size and topography. Concurrently, a data-based approach is employed to analyze rainfall and runoff records from 175 German catchments. The findings emphasize that streamflow distribution tails grow heavier with increased spatial rainfall variability, notably beyond a specific threshold. However, runoff generation may counteract, underscoring the concept of "catchment resilience," denoting the capacity of catchments to buffer increasing rainfall heterogeneities. The study revealed that small and elongated catchments exhibit less resilience, primarily influenced by runoff routing through river networks. This sheds light on the intricate relationships among spatial rainfall characteristics, catchment geometry, and the potential occurrence of high flows. Comparable data analyses ensure the validation of the simulation results and reinforce the transferability of these findings.

In the second section of this dissertation, the challenges of identifying heavy-tailed flood behavior due to limited data records and the absence of physical support for current indices are addressed. To overcome these limitations, a new index of heavy-tailed flood behavior, grounded in a physically-based description of streamflow dynamics, is introduced. This innovative index, embodied by the hydrograph recession exponent, allows for the inference of heavy-tailed flood behavior from daily flow records, even when data records are brief. A diverse array of case studies across Germany, representing various climatic and physiographic settings, validates this index. The findings underscore its effectiveness in reliably identifying cases with heavy or non-heavy tailed flood behavior from daily flow records. Furthermore, the index is capable of estimating the severity of tail heaviness and ranks it across cases, providing robust results even with limited data records. This index successfully addresses the primary limitations associated with current metrics, which often lack physical underpinnings and demand lengthy

data records for accurate tail behavior identification. The index also contributes valuable insights into the tail behavior of flood distributions and the associated flood hazard in river basins, utilizing readily available discharge data.

In the third section of this dissertation, hydrograph recession exponents derived from common streamflow dynamics are leveraged. These recession exponents have proven to be a robust indicator of flood tail propensity across regions with diverse data lengths. Through the analysis of extensive datasets from Germany, the United Kingdom (UK), Norway, and the United States (US), distinct patterns emerge. Germany and the UK exhibit prevalent heavy tails, while the US displays diverse behavior, and Norway primarily features non-heavy tails. This regional tail behavior is linked to the interplay between terrain and meteorological characteristics, further quantified through Köppen classifications. The temporal variability in catchment storage emerges as a critical mechanism driving highly nonlinear catchment responses favoring heavy-tailed floods, notably intensified by concurrent dry periods and high temperatures. Various flood generation processes, influenced by hydroclimatic seasonality and catchment scale, further shape this mechanism. These insights deepen the understanding of the intricate interplay between climate, physiographical settings, and flood behavior, underscoring the utility of hydrograph recession exponents in flood hazard assessment.

The findings of this study significantly contribute to our understanding of the mechanisms behind heavy-tailed flood behavior and introduce a physically-based, robust approach for assessing flood tail heaviness. This approach mitigates the sensitivity to limited data, a common challenge in such analyses, and enhances its applicability across regions. These outcomes open up numerous prospective applications in hazard research and engineering, such as flood frequency guidelines, assessing climate change-related impacts, flood hazard classification, quantitative flood hazard assessment, understanding the impact of human activities, and research in environmental sustainability.

Keywords Floods, Heavy-Tailed distribution, Extreme events, Extreme value statistics, Spatial rainfall variability, Hydrograph recession

Dedicated to God's Economy

Acknowledgements

I sincerely thank the Ministry of Education of Taiwan for initiating my PhD studies in Germany with the Government Scholarship to Study Abroad (GSSA). Additionally, I appreciate the financial support provided by the Helmholtz Association, Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE), and the German Research Foundation (DFG) through projects FOR 2416 "Space-Time Dynamics of Extreme Floods (SPATE)" and "Propensity of rivers to extreme floods: climate-landscape controls and early detection (PREDICTED)." Their generous funding was crucial to completing this cumulative dissertation developed at the Department of Catchment Hydrology, Helmholtz Centre for Environmental Research (UFZ), Halle (Saale).

I extend my deepest gratitude to Prof. Dr. Ralf Merz, Dr. Stefano Basso, and Assist. Prof. Dr. Soohyun Yang for their invaluable scientific guidance, unwavering support, and constant encouragement. Special thanks to the remarkable CATHYD team for their inspiring discussions, camaraderie, and the rejuvenating coffee breaks that vibrated my otherwise routine research life. I am also thankful to my collaborators in the TRACER group for broadening my horizons and providing opportunities to engage in fascinating research topics. Your contributions have been instrumental to the completion of this work.

I am immensely grateful to all my dear families, both my physical family and those within the household of God. Your countless prayers, brotherly love, and shepherding have been the bedrock that sustained me throughout this extraordinary four-year journey. Shuting, my beloved wife, your steadfast companionship, endurance, and unrequited support, despite the challenges, mean the world to me. Ultimately, I thank and praise our dear Lord Jesus, whose divine shepherding and faithful guidance led me to accomplish what felt like an impossible mission. As you've told us, "*I am the Alpha and the Omega, the First and the Last, the Beginning and the End*" (Rev. 22:13). In your amazing grace, you even sent two little girls who consistently prayed for their daddy's successful completion of his studies. Such is the magnitude of the grace I have received.

Table of Contents

Title Page	i
Zusammenfassung	iii
Abstract	v
List of Figures	xiii
List of Tables	xv
Abbreviations	xvii
1 Introduction	1
1.1 Definition, significance, and current state	1
1.2 Challenges and research prospects	4
1.3 Research objectives and dissertation structure	6
2 Impacts of Spatial Heterogeneity in Rainfall	9
2.1 Catchment and hydrological data	10
2.2 Observation-based stochastic framework for spatial heterogeneity in rainfall . .	11
2.3 Streamflow simulation	16
2.4 Alteration of tail heaviness induced by spatial rainfall variability	18
2.5 Identification of heavy-tailed streamflow distributions	19
2.6 Comparable data analyses	19
2.7 Chapter results and discussion	20
2.7.1 Spatial rainfall scenatio 1: fixed variability without autocorrelation . . .	20
2.7.2 Spatial rainfall scenatio 2: varying variability without autocorrelation .	23
2.7.3 Spatial rainfall scenatio 3: fixed variability with autocorrelation	24
2.7.4 Corroboration through real rainfall and streamflow data	25
2.7.5 Discussion	29
2.8 Chapter conclusions	32
2.9 Supporting information	33
2.10 Data availability	40

TABLE OF CONTENTS

3	Insights from Common Streamflow Dynamics	41
3.1	Asymptotic analysis of common streamflow dynamics	42
3.2	Dataset description	45
3.3	Hydrograph recession analysis	45
3.4	Benchmarking heavy-tailed case studies: recession exponents vs. empirical power law	46
3.5	Chapter results and discussion	46
3.5.1	Identification of heavy-tailed case studies	46
3.5.2	Evaluation of tail heaviness	47
3.5.3	Effectiveness in short-length record datasets	49
3.5.4	Discussion	53
3.6	Chapter conclusions	55
3.7	Supporting information	56
3.8	Data availability	58
4	Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation	59
4.1	Study areas and data	61
4.2	Inferring heavy tails of flood distributions from common streamflow dynamics	62
4.3	Validation of hydrograph recession exponents as an index of heavy-tailed flood behavior	63
4.4	Analyses of spatial and seasonal patterns of inferred flood tail behavior	64
4.5	Chapter results and discussion	64
4.5.1	Effectiveness of identifying heavy-tailed flood behavior using common discharge dynamics	64
4.5.2	Spatial patterns of inferred flood tail behavior	69
4.5.3	Seasonal patterns of inferred flood tail behavior	73
4.5.4	Factors associated with catchment scales and their role in flood tail behavior	76
4.5.5	Discussion	78
4.6	Chapter conclusions	82
4.7	Supporting information	84
4.8	Data availability	85
5	Conclusions and Future Avenues	87
5.1	Conclusions	87
5.2	Future avenues	90
	References	93
	Appendix A List of Publications and Author Contributions	119
	Appendix B Declaration under Oath	121
	Appendix C Curriculum Vitae	123

List of Figures

2.1	Study catchments in chapter 2	11
2.2	Schematic diagram of spatial rainfall scenarios	13
2.3	Estimated rainfall spatial variability of observations	14
2.4	An example of data-based stochastic spatial autocorrelation simulation of the rainfall field	16
2.5	Schematic diagram of the adopted hydrological model	18
2.6	Effects of increasing spatial rainfall variability (CV_{in}) on the tail heaviness of the streamflow distribution of the Ilm River at the gauging station of Niedertrebra	21
2.7	Effects of the catchment size	22
2.8	Effects of the catchment shape	23
2.9	Effects of varying rainfall variability (CV_{cross})	24
2.10	Effects of spatial autocorrelation of rainfall	27
2.11	Data-based analysis of the relation between streamflow-distribution tail heaviness and spatial rainfall variability with the impacts of catchment size and shape	28
2.12	Overall goodness-of-fit of probability distributions for the rainfall fields in REGNIE data set	35
2.13	Examples of observed rainfall fields with the increasing spatial CV of the Ilm River at the gauging station of Niedertrebra	37
2.14	The terrain map of Niedertrebra based on a digital elevation model	37
2.15	Definition of the index of relative tail heaviness	39
3.1	Accuracy of the proposed index	48
3.2	Empirical power law exponent b as a function of the proposed index of heavy- tailed behavior a	49
3.3	Stability of the categorization of case studies into heavy/nonheavy-tailed flood behavior for decreasing data lengths	52
3.4	A reference map of gauges across Germany used in chapter 3	57
4.1	Effectiveness of identifying heavy-tailed flood behavior using hydrograph recession exponents	66
4.2	Empirical power law exponent b as a function of the hydrograph recession exponent a (physically-based index of heavy-tailed flood behavior)	68
4.3	Spatial distribution of dominant flood behavior	70
4.4	Propensity of inferred flood tail behavior in diverse climate regions	71

LIST OF FIGURES

4.5	Consistency of inferred flood tail behavior across seasons	74
4.6	Seasonal variations in the percentage of inferred flood tail behavior between heavy and nonheavy case studies	75
4.7	Seasonal variations in recession exponents (inferring flood tail behavior) across diverse climate regions	76
4.8	Catchment nonlinearity as a function of catchment area	77
5.1	Summary of the emergence of flood tail heaviness	89
5.2	Streamflow recession patterns in recent decades in Norway	92

List of Tables

2.1	Information on study catchments for scenario simulations	12
2.2	Parameters for the Stochastic Rainfall Spatial Autocorrelation Generator . . .	15
2.3	Parameters for fixed variability scenarios (1, 3):	38
2.4	Parameters for the varying variability scenario (2):	38
4.1	Comparison of inferred flood tail behavior propensity with climate characteristics	72
4.2	Daily Hydrological Data Information	84

Abbreviations

<i>AIC</i>	Akaik e Informati on Criterion
<i>BfG</i>	Federal Institute for Hydrology
<i>BIC</i>	Bayesian Information Criterion
<i>CDF</i>	Cumulativ e Distribution Function
<i>DEM</i>	Digital Elevation Model
<i>DFG</i>	German Researc h Foundation
<i>DWD</i>	German Weather Service
<i>EVII</i>	Extreme Value Distribution of Maxima of Type II
<i>GEV</i>	Generalized Extreme Value Distribution
<i>GP</i>	Generalized Pareto Distribution
<i>GRDC</i>	Global Runoff Data Centre
<i>HESS</i>	Hydrology and Earth System Sciences Journal
<i>HRU</i>	Hydrologic Resp onse Unit
<i>KS</i>	Kolmogorov–Smirnov
<i>LfU</i>	Bavarian State Offi ce of Environment
<i>LM</i>	L-Moments
<i>ML</i>	Maxim um Likelihood
<i>MOE</i>	Ministry of Education of the Republic of China
<i>MOM</i>	Metho d of Moments
<i>NOAA</i>	National Oceanic and Atmospheric Administration
<i>PDF</i>	Probabilit y Densit y Function
<i>PHEV</i>	Physically-Based Extreme Value Distribution of River Flows
<i>PREDICTED</i>	Propensity of Rivers to Extreme Floods: Climate-Land scape Controls and Early Detection
<i>REGNIE</i>	Regionalisierung der Niederschlagshöhen
<i>SPATE</i>	Space-Time Dynamics of Extreme Floods

ABBREVIATIONS

SRTM Shuttle Radar Topography Mission

UFZ Helmholtz Centre for Environmental Research

UK United Kingdom

US United States

UTR Upper Tail Ratio

1

Introduction

Floods, as one of the most pervasive natural hazards, pose an ongoing threat to lives and properties across the globe (McDermott, 2022). The inherent risk of hazards, however, greatly depends on society's ability to anticipate the likelihood of high/extreme floods occurring (Merz et al., 2021). The presence of heavy-tailed behavior in flood frequency distributions signifies this likelihood, and therefore, accurately identifying such behavior is crucial in mitigating the hazards associated with extreme floods (Merz et al., 2022). This dissertation aims to enhance our comprehension of the emergence of heavy-tailed streamflow and flood distributions while proposing reliable methods for predicting and managing these behaviors in practical applications.

1.1 Definition, significance, and current state

When constructing an empirical frequency distribution of observations, the right tail of the distribution represents the likelihood of encountering extremely large magnitudes of the variable, providing valuable insights for hazard assessment. Mathematically, a distribution F is defined as a (right) heavy-tailed distribution if and only if,

$$\int_{-\infty}^{\infty} e^{\theta x} dF(x) = \infty, \text{ for all } \theta > 0 \quad (1.1)$$

The tail function of F can be defined as $\bar{F} = F(x, \infty)$, $x \in \mathbb{R}$, where \mathbb{R} denotes the real number line. Therefore, a heavy-tailed distribution is one of which the tail function cannot be bounded by any exponentially decreasing function. Though various sub-classifications of heavy-tailed distributions exist (Werner and Upper, 2002), a commonly accepted definition is that a distribution exhibits a heavy tail when its right tail decays more slowly than that of an exponential distribution (El Adlouni et al., 2008). This characteristic can also be observed as a heavier pattern in the tail of the distribution (i.e., the asymptotic behavior) when compared to an exponential distribution in their probability density functions (PDF).

Heavy-tailed behavior within a distribution indicates a higher likelihood of extreme values occurring compared to what would be expected in a light-tailed distribution. In a light-tailed

distribution, the tail of its PDF rapidly approaches zero as variable values become more extreme. In contrast, a heavy-tailed distribution exhibits a slower convergence towards zero as values become more extreme, indicating a greater probability of encountering exceptionally large values. As a result, heavy-tailed distributions emphasize the occurrence of rare events with significantly large magnitudes. This characteristic has been employed in research on hazards across diverse natural and human sciences, including seismic hazards (e.g., Felgueiras, 2012), economic risk (e.g., Ibragimov et al., 2015), rainfall extremes (e.g., Papalexiou and Koutsoyiannis, 2013), flood extremes (e.g., Smith et al., 2018), and social governance (Li et al., 2021). Particularly, given that human intuition often anticipates light-tailed behavior, processes/variables exhibiting heavy tail behavior often result in surprise and can cause severe hazards (Taleb, 2007). Conversely, accurately detecting heavy-tailed behavior helps prevent the underestimation of extremes (Mushtaq et al., 2022), which enables the enhancement of hazard assessment and management.

The significance of heavy-tailed distributions in hydroclimatology has garnered considerable attention in recent decades. Gumbel's (1958) guidebook on extreme statistics, which focused on floods and droughts, played a pivotal role in raising awareness among hydrologists about heavy-tailed distributions. Katz and Brown (1992) emphasized the importance of considering event magnitude variability instead of relying solely on averages in extreme statistics, especially in the context of climate change. Observational-based studies, as suggested by Katz (2001) for daily precipitation and by Anderson and Meerschaert (1998) for monthly flow rate, indicated that these variables should be recognized as having heavy-tailed behavior in their distributions. Studies over the past two decades have highlighted the prevalence of heavy-tailed behavior in hydroclimatology and its significance (Katz et al., 2002; Morrison and Smith, 2002; Koutsoyiannis, 2004a,b; Wilson and Toumi, 2005; Bernardara et al., 2008; El Adlouni et al., 2008; Papalexiou and Koutsoyiannis, 2013; Basso et al., 2015; Rossi et al., 2016; St. George and Mudelsee, 2019; Martinez-Villalobos and Neelin, 2021; Gupta and Chavan, 2023). As a result, selecting an appropriate distribution function to describe the heavy-tailed behavior of the observed values becomes essential. In the context of rainfall, the Extreme Value Distribution of Maxima of Type II (EV II) is recommended as a more suitable representation for annual maximum rainfall compared to the Gumbel distribution (Koutsoyiannis, 2004a,b). Additionally, the Frechet distribution is proposed for describing daily rainfall distributions (Papalexiou and Koutsoyiannis, 2013), both of which exhibit heavy-tailed behavior. Moreover, future rainfall patterns simulated by global climate models corroborate the prevalence of heavy-tailed behavior in rainfall distributions (Martinez-Villalobos and Neelin, 2021). In the context of floods, Generalized Extreme Value (GEV) distributions (e.g., Morrison and Smith, 2002) or Generalized Pareto (GP) distributions (e.g., Bernardara et al., 2008) have often been suggested as suitable representations of flood peaks, characterized by their heavy-tailed behavior. The importance of including extreme events in available data has been underscored to achieve dependable fitting of heavy-tailed distributions describing floods (St. George and Mudelsee, 2019). Furthermore, the presence of heavy-tailed behavior in streamflow distributions has been leveraged to enhance hazard risk assessment systems (Gupta and Chavan, 2023).

Given the significance and prevalence of heavy-tailed behavior in hydroclimatic parameters, two prominent research areas have captured substantial attention: the identifi-

cation/classification of heavy-tailed distributions and the investigation of the underlying mechanisms driving such behavior.

The methods for identifying heavy-tailed distributions can be subcategorized into graphical methods, nonparametric approaches, and parametric approaches. Estimating tail heaviness through graphical methods typically involves analyzing the slope of specific functions. For instance, a log-log plot can facilitate the recognition of heavy-tailed behavior by revealing a linear pattern in the exceedance probability curve (Embrechts et al., 1997; Beirlant et al., 2004). The generalized Hill's estimator employs the slope of the line in the Hill plot to assess the relative tail heaviness of distributions (Beirlant et al., 2004; Diop and Lô, 2009). Another approach involves the use of the mean excess function, which characterizes tail behavior by analyzing the slope of the regression line between the mean excess function and the lower boundary of the variable defining the tail (Embrechts et al., 1997; Nerantzaki and Papalexiou, 2019).

Additionally, nonparametric statistical methods have been applied to evaluate tail heaviness. The upper tail ratio (Villarini et al., 2011; Smith et al., 2018) divides the maximum record by the value corresponding to a specific return period (e.g., the 10-year return period). This ratio typically falls within a range above zero, with a larger ratio indicating a heavier tail. The Obesity index (Cooke and Nieboer, 2011) assesses tail heaviness by quantifying the concentration of large samples within the distribution, yielding values between 0 and 1, with higher values representing a heavier tail. Similarly, the Gini index (Eliazar and Sokolov, 2010) ranges from 0 to 1 and indirectly evaluates tail heaviness by measuring the effects of both small and large samples on the distribution.

Parametric methods, as the traditional approach, have commonly been employed to assess the tail heaviness of empirical data. This involves fitting a number of candidate distributions to empirical data and using goodness-of-fit tests to determine the appropriateness of the fitting, ultimately selecting the most suitable distribution. Different parameter estimation methods, including maximum likelihood (ML) (Engeland et al., 2004), L-moments (LM) (Bezak et al., 2014), and the method of moments (MOM) (Öztekin, 2005), can be used to fit the candidate distributions to data. Hybrid metrics of these methods have also been proposed to enhance the reliability of the fitting (Morrison and Smith, 2002). To evaluate goodness of fit, several methods are frequently utilized, such as the Kolmogorov–Smirnov (KS) test (Clauset et al., 2009), Q-Q plot (Engeland et al., 2004), chi-square test (Vivekanandan, 2015), L-moment ratio diagram (Rowinski et al., 2002), Akaike information criterion (AIC) (Engeland et al., 2004), and Bayesian information criterion (BIC) (Haddad and Rahman, 2011). A comprehensive review of the strengths and limitations of these methods for assessing tail heaviness in empirical distributions was conducted by Nerantzaki and Papalexiou (2022).

Many of these methods do not necessarily provide a direct definition of heavy-tailed behavior; instead, they offer a means of assessing relative tail heaviness, often requiring an additional benchmark to define the tail behavior. For instance, the obesity index for exponential distributions is suggested as 0.75, implying that any distribution with an obesity index above 0.75 should be classified as heavy-tailed (Cooke et al., 2014). One of the most widely accepted classification schemes for tail heaviness has been proposed for commonly used continuous probability distributions in hydrology (Werner and Upper, 2002; El Adlouni et al., 2008).

Based on this classification, the tail behavior from heaviest to lightest among distributions includes (but is not limited to): Pareto > Fréchet > Lognormal > Gumbel/Gamma/Pearson III > Exponential > Normal distributions.

Thanks to these quantitative methods, investigations into the underlying mechanisms of heavy-tailed flood behavior have been conducted across domains. While one intuitive assumption is that heavy-tailed floods result from heavy-tailed rainfall, findings have actually contradicted this notion. Contrary to expectations, heavy-tailed rainfall are not necessarily translated into heavy-tailed floods (McCuen and Smith, 2008; Sharma et al., 2018). Instead, the driving factors behind heavy-tailed floods are more complex (Merz et al., 2022).

Flood generation processes and catchment water balance are identified as more compelling controls of heavy-tailed flood behavior (Merz et al., 2022). For instance, Bernardara et al. (2008) found that Mediterranean regimes often exhibit dominant heavy-tailed flood behavior, while regions influenced by snowmelt or continental climates show less of such behavior. Tarasova et al. (2020a) dissected the contributions of various flood-type events to the flood frequency curve. Mushtaq et al. (2023) demonstrated the potential for incorporating mixed flood type analyses to more accurately depict heavy-tailed flood behavior. Moreover, the nonlinearity of catchment responses to rainfall events (Gioia et al., 2008; Rogger, Pirkl et al., 2012; Basso et al., 2015) and the drier state of catchment soil moisture (Merz and Blöschl, 2009; Botter, 2010; Berghuijs et al., 2014) have been found correlated with the presence of heavy-tailed flood behavior.

Other factors have also been suggested to play a role, such as catchment sizes (Villarini and Smith, 2010), spatial variability of catchments (Struthers and Sivapalan, 2007), confluences of river networks (Vorogushyn and Merz, 2013), and the construction of reservoirs (Assani et al., 2006). However, these factors either lack sufficient evidence to establish their roles clearly, or there is a lack of consensus regarding their influence. In summary, various hypotheses on heavy-tail behavior have been proposed, but a coherent, spatially comparable indicator based on hydrological process understanding is still missing.

1.2 Challenges and research prospects

Drawing upon our prior literature review, we conclude the challenges and propose prospects of heavy-tailed flood behavior research within three main aspects:

1. Understanding the influence of changing watershed systems on heavy-tailed flood behavior

Hydrological processes within watershed systems are expected to undergo significant transformations due to the global rise in temperature and extreme rainfall (Pendergrass, 2018; Qiu et al., 2021). It has been observed that one of the direct consequences of global warming on hydrological processes involves an augmentation in extreme rainfall intensity (Li et al., 2019). This intensification, however, does not manifest uniformly across space; instead, it exhibits a spatial distribution pattern on a global scale (Donat et al., 2016). In addition to heightened rainfall intensity, an increased frequency of extreme rainfall events has been proposed to contribute more significantly to the overall rise in total rainfall accumulation (Myhre et al., 2019). Furthermore, while both the intensity and frequency of extreme rainfall are on the rise, the spatial extent of these events is

diminishing. This suggests a tendency towards more localized and heterogeneous spatial distribution of extreme rainfall events (Wasco et al., 2016). Intriguingly, this shift in the spatial distribution of rainfall impacts high-return-period floods more substantially than the temporal alterations in rainfall (Peleg et al., 2017).

In the context of ongoing global warming, understanding the repercussions of these evolving rainfall patterns on flood extremes becomes increasingly paramount. However, providing a definitive answer to this question remains challenging due to the limited understanding of the underlying mechanisms. Particularly perplexing are the contradictory findings surrounding the influence of spatial heterogeneity on the emergence of heavy-tailed flood behavior (Struthers and Sivapalan, 2007; Harman et al., 2009; Rogger, Pirkel et al., 2012). These discrepancies pose constraints on accurately assessing the effects of heightened spatial heterogeneity in extreme rainfall on the emergence of flood hazards.

2. Challenges in reliable detection and prediction of heavy-tailed flood distributions

A significant challenge in current methods for identifying heavy-tailed flood behavior lies in the uncertainty of estimation caused by limited data availability. Given the global shortage of accessible long-term hydrological observations (Lins, 2008) and the reliance of current methods solely on statistical analysis of available data to estimate tail behavior, the reliability of such estimations comes into question. Papalexiou and Koutsoyiannis (2013) highlighted that accurate estimation of heavy-tailed rainfall behavior (achieved through GEV fittings) necessitates extensive data records. Smith et al. (2018) found a clear dependency between the estimated upper tail ratio and data length for floods. Wietzke et al. (2020) conducted a comparison of frequently used indices of heavy-tailed flood behavior, including the GEV shape parameter, upper tail ratio, Obesity index, and Gini index. They unveiled the general issue of data length sensitivity across various aspects in these indices.

To mitigate this uncertainty and enhance the reliability of estimations, especially in extreme estimations, the suggestion has been made to extract information more effectively from available data. For instance, ordinary dynamics can be employed to infer maximum behavior (Marani and Ignaccolo, 2015; Mushtaq et al., 2022). Furthermore, utilizing physically-based proxies, grounded in the knowledge of the underlying mechanisms of heavy-tailed flood behavior, potentially offers an alternative approach beyond solely relying on statistical analysis of samples. This parallels studies that have contributed to understanding heavy-tailed behavior in rainfall distributions (Wilson and Toumi, 2005). However, in the context of floods, the identification of suitable proxies remains an open question, despite the exploration of underlying mechanisms.

3. Comparative regional analysis of heavy-tailed flood behavior

Studying the link between spatial distributions of heavy-tailed flood behavior and regional physiographical settings helps unravel the complex interplay of factors influencing flood extremes. However, relatively limited research has been conducted in this area (Merz et al., 2022). Smith et al. (2018) explored heavy-tailed flood behavior in the United States using the upper tail ratio, identifying mountainous terrain and intense thunderstorms

as influential factors. They also observed distinct seasonal patterns between record floods and extreme floods. Similarly, Yang et al. (2021) applied the upper tail ratio to analyze flood distributions in China, pinpointing windward mountainous regions as areas prone to heavy-tailed flood behavior. They suggested the significance of the interaction between regional terrain and meteorological conditions, which contributes to prolonged rainfall and intensifies heavy-tailed flood behavior. Both studies utilized the upper tail ratio to estimate flood tail behavior. However, it's worth noting that this approach has been shown to be highly sensitive to the length of the available data, often resulting in underestimation when dealing with shorter datasets (Smith et al., 2018; Wietzke et al., 2020). As mentioned earlier, the sensitivity to sample sizes is a common challenge in the current methods used to assess the tail heaviness of distributions. To ensure comparability of analyses across catchments, it's crucial to use datasets with similar lengths (and overlapping periods) (Cunderlik and Burn, 2002). Given the global variability in the availability of hydrological observations (including streamflow) across different regions/gauges (Lins, 2008), conducting comparative regional analysis of heavy-tailed flood behavior is constrained. Due to this limitation, it becomes necessary to exclude river gauges with shorter record lengths in current studies. This practice, however, reduces the number of catchments available for analysis and can often make international comparisons challenging, or even disqualify regions with short record lengths from heavy-tail analyses altogether.

1.3 Research objectives and dissertation structure

This research aims to address acknowledged challenges in the field of heavy-tailed flood behavior research, as discussed in the preceding section. Specifically, Chapter 2 delves into the impact of rainfall-induced spatial variability on the emergence of heavy-tailed streamflow behavior. Chapter 3 introduces a robust proxy for inferring heavy-tailed flood behavior grounded in catchment hydrological processes, and Chapter 4 conducts an extensive geographical analysis of heavy-tailed flood behavior across diverse physiographical settings using the proposed indicator. The aim is to confirm and summarize the advancement of the understanding of heavy-tailed behavior by this new approach studied in this work.

These three chapters deal with different aspects by employing diverse methodologies and datasets to properly analyze the specific issues posed in each chapter. Each chapter is preceded by a deeper literature review on the chapter's focus, which discusses the respective state of the art and research gaps of the chapter topic. Based on this, research questions can be specified as follows:

Challenge 1 - Impacts of Spatial Heterogeneity in Rainfall (Chapter 2):

- What is the general role of the spatial rainfall variability on the emergence of heavy-tailed streamflow distributions?
- Does this role vary across catchments, and what would be the underlying controls of this role?

Challenge 2 - Insights from Common Streamflow Dynamics (Chapter 3):

- Is there an appropriate descriptor of flood distribution tail behavior within the intrinsic hydrological dynamics of the flow regime?
- How effective would the use of the proposed descriptor be in identifying heavy-tailed flood behavior?
- Will the proposed descriptor provides a reliable estimation of flood tail behavior across catchments with varying data lengths?

Challenge 3 - Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation (Chapter 4):

- Could the identified descriptor be validated through an analysis of an extensive dataset that encompasses a diverse range of climate and catchment characteristics to infer hotspots of heavy-tailed flood behavior?
- What insights can we gain about the influences of diverse physiographical settings on heavy-tailed flood behavior from the constructed geography of such behavior?

Individual results, discussions, and chapter conclusions corresponding to these focuses are presented directly in each chapter. Finally, Chapter 5 offers a condensed overview of the important findings obtained from the preceding chapters. It highlights the progress made in understanding the factors and mechanisms related to heavy-tailed flood behavior, showcasing the contributions of this study. This dissertation concludes by discussing the potential applications and future research directions related to heavy-tailed flood behavior, with the aim of advancing hazard prediction and prevention research.

2

Impacts of Spatial Heterogeneity in Rainfall

Rainfall is often considered one of the primary drivers of hydrological responses within catchments (e.g., Bracken et al., 2008; Arnell and Gosling, 2016; Wasco and Sharma, 2017). Spatial rainfall variability has been proved as non-negligible to accurate streamflow prediction (e.g., Singh, 1997; Zoccatelli et al., 2011; Zhao et al., 2013). Particularly, it has been suggested as more important than the temporal rainfall structure (Peleg et al., 2017) to evaluate the impacts of climate changes on extreme streamflow events (Fatichi et al., 2016). The increasing trend in the occurrence of localized and extreme rainfall events has been found under global warming conditions (Donat et al., 2016; Pendergrass, 2018; Li et al., 2019; Myhre et al., 2019). A considerable rise in spatial rainfall variability caused by the increasing intensity of extreme rainfall and decreasing spatial extent have been discussed in studies based on observations (Wasco et al., 2016; Peleg et al., 2018) and modeling (Peleg et al., 2022). Furthermore, the change of spatial rainfall variability itself has been shown to be more impactful on floods than the component of rainfall intensity (Peleg et al., 2022).

However, traditionally, the focus of the effects of spatially variable rainfall on streamflow has been on the variability of a single rainfall storm or the magnitude of the resulting peak discharge (Borga et al., 2007; Viglione et al., 2010; Paschalis et al., 2014; Lu et al., 2017). Few researchers have addressed the issue of the effects of continuous spatially variable rainfall on streamflow distributions. Even though studies of flood frequency analysis have provided some insights into flood hazards based on peak discharges (e.g., Menabde and Sivapalan, 2001; Kay et al., 2006; Fiorentino et al., 2007; Kjeldsen et al., 2014), flood hazards based on variations between low and high flows remain poorly understood. Furthermore, long-term streamflow response (e.g., cumulative streamflow distributions) to rainfall is linked with multiple hydrological processes and complex interactions between catchment units. There is still some controversy surrounding the impacts of spatially heterogeneous rainfall on streamflow due to variable climate conditions (Viglione et al., 2010; Zhao et al., 2013), catchment sizes (Merz and Blöschl, 2009; Zhu et al., 2018), or the heterogeneities in other catchment attributes

(Struthers and Sivapalan, 2007; Harman et al., 2009; Rogger, Kohl et al., 2012). Hence, it remains a great challenge for hydrologists to understand the linkages between the changing characteristics of rainfall and high flows. Sharma et al. (2018) pointed out that one of the foremost issues, which is still missing, is the relationship between catchment size, catchment geometry, and storm characteristics (e.g., extent, duration, intensity). Meanwhile, providing a corresponding data validation to modeling results is challenging. A general investigation of the effects of spatial rainfall variability from both simulation scenarios and data analyses is needed.

To fill the knowledge gap, this study addresses two research questions: (1) what is the role of the spatial rainfall variability on heavier-tailed streamflow distributions, and (2) does this role be dependent on catchment sizes and shapes? I aim to investigate the research questions using both simulation scenarios and data analyses. Specifically, I applied a stochastic approach to generate synthetic rainfall fields and fed the results as the input of a non-gridded continuous hydrological model. The spatial context is based on catchment units (i.e., HRUs). The output of the long-term hydrological response was used for evaluating the tail heaviness of the streamflow distribution. Data analysis was presented in a comparable approach to the synthetic framework.

2.1 Catchment and hydrological data

The study is based on daily rainfall and streamflow observations for a set of 175 German catchments taken from Tarasova et al. (2018). The catchment areas vary from 36 to 23,700 km^2 , with a median value of 688 km^2 . Catchments influenced by large reservoirs or control gates (Lehner et al., 2011) and affected by visible anthropogenic streamflow disturbances are disregarded.

The length of daily streamflow and rainfall time series ranges between 37 and 63 years (between 1951 and 2013), with a median value of 61 years. We computed the specific (i.e., per unit catchment area) streamflow at each catchment outlet. Hereafter, we use the simpler term ‘streamflow’ instead of specific streamflow. Daily rainfall time series for each catchment was derived from the Regionalisierung der Niederschlagshöhen (REGNIE) dataset, which is a 1 km^2 resolution rainfall field interpolated from point observations through multiple regression, provided by the German Weather Service (Rauthe et al., 2013). This study referred to only summer periods to avoid bias from snow melting, which is not considered in our modeling framework (see Section 2.3). Summer season was identified as June to August, as suggested by Beurton and Thielen (2009) for German catchments. Daily temperature time series in the considered catchments were generated from the German Weather Service observations by means of external drift kriging, using elevation as an explanatory variable (Zink et al., 2017). We applied the temperature to estimate potential evapotranspiration using the Thornthwaite (1948) equation.

Five of the 175 catchments were selected for scenario simulations (Figure 2.1 and Table 2.1). The Delme River basin at Holzkamp (98 km^2 to represent small size), the Ilm River basin at Niedertrebra (887 km^2 to represent medium size), and the Amper River basin at Inkofen (2841 km^2 to represent large size) form the first group of basins used in the simulations,

for they have various catchment sizes while their drainage densities D_d (Horton, 1932) and elongation ratios R_e (Schumm, 1956) are similar. The Ilm River basin at Niedertrebra (more elongated, $R_e=0.45$), the Innerste River basin at Heinde (less elongated, $R_e=0.78$), and the Unstrut River basin at Erfurt-Moebisburg (circular, $R_e=0.90$) form the second group of basins used in simulations for they have various catchment shapes while their drainage densities and sizes are similar.

A 100-m resolution digital elevation model (DEM) retrieved from the Shuttle Radar Topography Mission (SRTM) was used to generate the river networks of the five select catchments. River networks were determined by means of the Arc Hydro tools with the recommended threshold (i.e., 1%) for stream determination (ESRI, 2009). Hydrologic response units (HRUs) were defined as the drainage area between neighboring nodes of the river network or as the area draining to the uppermost nodes.

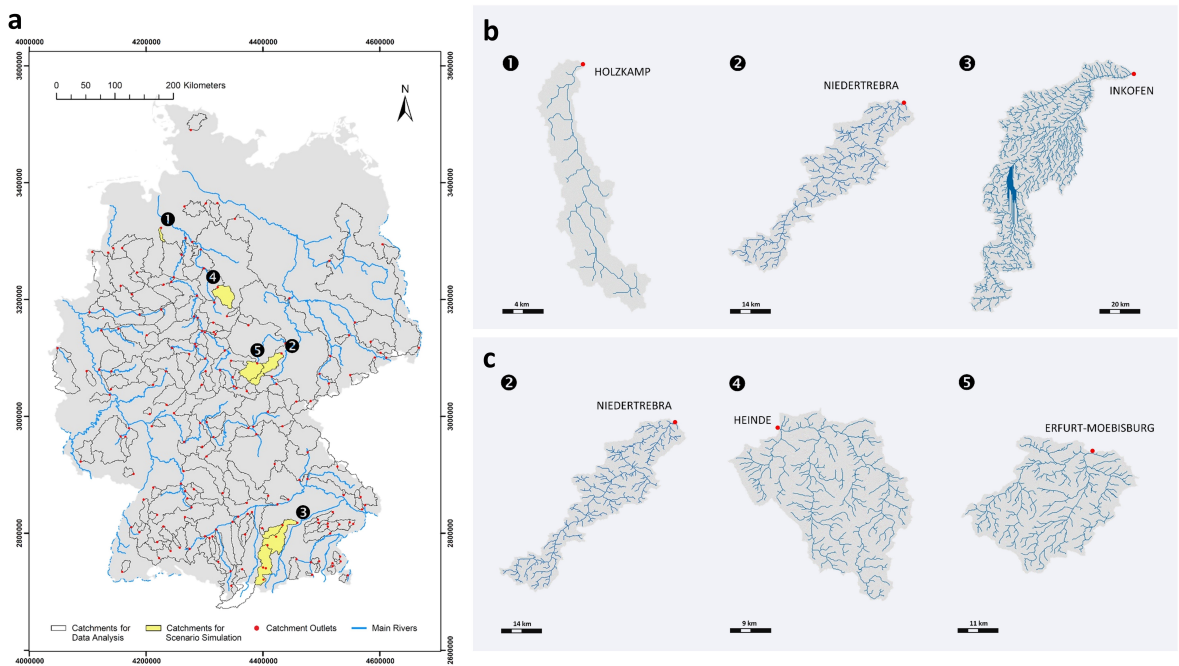


Figure 2.1: Study catchments in chapter 2 (a) Study catchments on the map of Germany. Black contours show the boundaries of 175 catchments used for data analyses. The yellow shaded areas labeled with numbers corresponding to panels b and c indicate the five catchments used for scenario simulation. (b) Catchments were selected for analyzing the effects of catchment size. The river networks and drainage areas of the Delme River basin at Holzkamp (1), the Ilm River basin at Niedertrebra (2), and the Amper River basin at Inkofen (3). (c) Catchments were selected for analyzing the effects of catchment shape. The river networks and drainage areas of the Ilm River basin at Niedertrebra (2), the Innerste River basin at Heinde (4), and the Unstrut River basin at Erfurt-Moebisburg (5).

2.2 Observation-based stochastic framework for spatial heterogeneity in rainfall

I first summarized the observed rainfall fields from the REGNIE dataset and identified the Gamma distribution as the fairest distribution to represent the spatial variabilities of rainfall (over the entire catchment) within five select catchments (see supporting information S1 and

2. Impacts of Spatial Heterogeneity in Rainfall

Table 2.1: Information on study catchments for scenario simulations Group 1 basins are used for the investigation of the effects of catchment size, and group 2 basins are used to study the effects of catchment shape. The elongation ratio close to 1 indicates high circularity.

Group	Catchment	Area [km^2]	Elongation Ratio, R_e [-]	Drainage Density, D_d [km/km^2]	HRU Number [-]	HRU Area (median/interquartile) [km^2]
1	Holzcamp	98	0.45	0.74	51	1.91/1.17
1,2	Niedertrebra	887	0.45	0.64	429	1.59/1.57
1	Inkofen	2841	0.47	0.88	1429	1.59/1.49
2	Heinde	898	0.78	0.69	438	1.60/1.70
2	Erfurt- Moebisburg	847	0.90	0.71	422	1.65/1.48

Figure 2.12). Then, we generated random variables from a pre-assigned Gamma distribution as the rainfall depths of each day, which were then allocated to each HRU to form the synthetic rainfall field. We systematically considered three critical components: the spatial rainfall variability, the varying degree of spatial rainfall variability in time, and the rainfall autocorrelation in space in our framework. Figure 2.2 exemplary illustrates the synthetic framework and the distinctions between components. Figure 2.2a shows the uniform case where the synthetic rainfall field is uniformly distributed in space with the average rainfall depth of observations (based on the REGNIE dataset) for the entire catchment and follows the observed time series. The hydrological response generated from this case is considered the reference point in this work. Figure 2.2b shows the first scenario which encompasses the spatially variable rainfall field with a fixed spatial variability (measured as coefficient of variation of rainfall) for all days, while every day exhibits randomly spatial rainfall patterns (alike Ayalew et al., 2014). Again, for each day, the catchment’s average synthetic rainfall depth equals the observed value. We implemented one more process, the varying variability in time, as the second scenario (Figure 2.2c). The spatial variability changes day by day in this case, with the catchment’s average synthetic rainfall depth each time equal to the catchment average calculated from REGNIE, and the spatial pattern of each day is random. In addition, I implemented another process, the rainfall autocorrelation in space, as the third scenario (Figure 2.2d). The rainfall field of each day is constructed by a stochastic spatial autocorrelation function, while the spatial rainfall variability is fixed in time, and the mean depth follows the observations. The framework starts from the most simplified case and adds different sophisticated components step by step to distinguish between effects on the tail of the streamflow probability distribution caused by different components.

The parameters of the Gamma distribution were assigned based on the requests of different scenarios. The shape parameter k of the Gamma distribution was either fixed for all days (scenarios 1 and 3) or allowed to vary randomly within a specified range (scenario 2). The scale parameter was then computed by setting the mean of the Gamma distribution equal to the average observed daily rainfall. Each realization applied one setting of the parameters of a Gamma distribution for the entire time series. For the rainfall fields with fixed variability, we imposed fifty-two different values of k between 10^3 and 10^{-4} (see supporting information Table 2.3), which resulted in a wide range of variabilities from extremely low (near to homogeneous, with the largest k) to extremely high (when the smallest k are utilized). For the rainfall fields

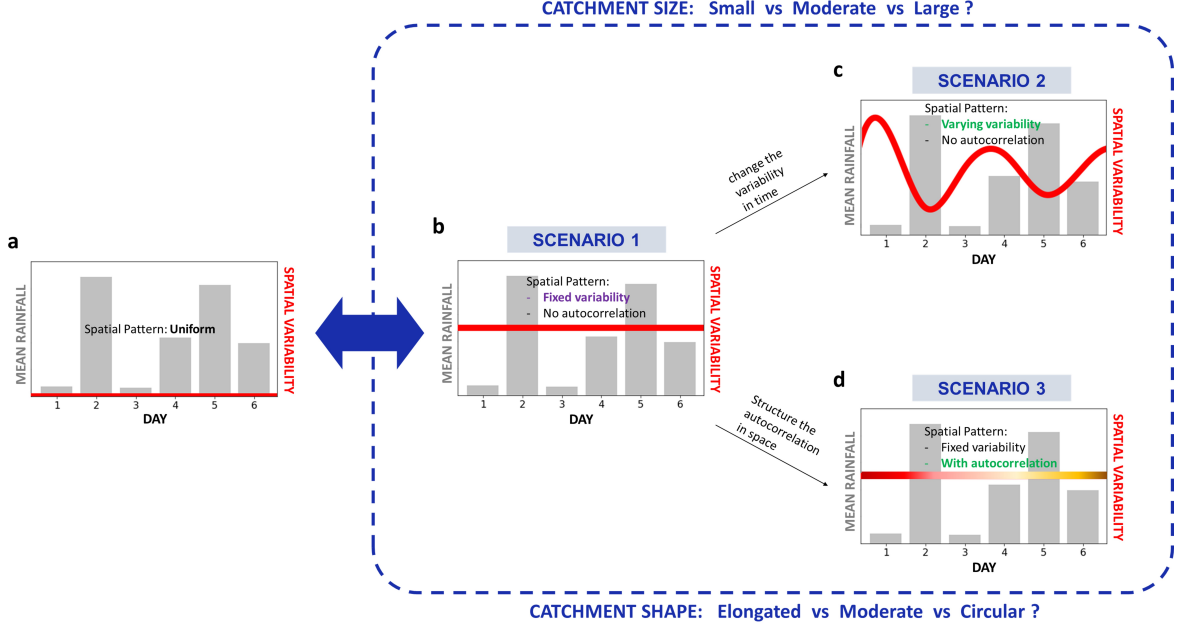


Figure 2.2: Schematic diagram of spatial rainfall scenarios The mean rainfall depth and event intermittence follow the observational time series in all cases. Settings for rainfall fields with (a) uniform depths and patterns in space, (b) fixed variability of depths in time (indicated by the horizontal red line for the value of spatial variability) and random patterns in space (no autocorrelation), (c) varying variability of depths in time and random patterns in space (no autocorrelation), and (d) fixed variability of depths in time and data-based stochastic autocorrelation structures in space.

with varying variability, the shape parameter k of the Gamma distribution was changed day by day. In particular, we allowed it to fluctuate randomly around a central value k_0 within a specified range $(-b, b)$. I generated rainfall fields characterized by low to high degrees of varying variability by changing the width of this range and assigning to b twenty different values between $0.05k_0$ and k_0 . We set $k_0 = 10^{-1}$, i.e., an intermediate value among those analyzed for k in the fixed variability scenario (see supporting information Table 2.4). With this approach, we generated variations of the spatial variability across events while controlling the average spatial variability within events to highlight effects due to the changing of spatial rainfall variability only. A hundred stochastic realizations for each specified shape parameter were generated to assess the uncertainty of the results.

I adopted three distinct indices (Equations 2.1 to 2.3) to evaluate the degree of variability of the synthetically generated spatial rainfall fields. The first index is the coefficient of variation in space (CV_{space}) of the rainfall depth assigned to all catchment units in each day, which we termed $CV_{space,t}$ for day t :

$$CV_{space,t} = \frac{\sqrt{\frac{1}{U} \sum_{u=1}^U (d_{u,t} - \bar{d}_t)^2}}{\bar{d}_t} \quad (2.1)$$

where $d_{u,t}$ [mm] is rainfall depth at catchment unit u on day t , \bar{d}_t is the spatial mean of rainfall on day t , and U is the total number of catchment units.

2. Impacts of Spatial Heterogeneity in Rainfall

The second index is the temporal mean of the first index, which characterizes the average spatial variability of rainfall across all wet days (i.e., days with rainfall > 0) of the time series. It is computed as:

$$CV_{in} = \frac{1}{T} \sum_{t=1}^T CV_{space,t} \quad (2.2)$$

where T is the total number of wet days in the rainfall time series of a catchment.

The third index is the temporal coefficient of variation of the first index, which represents the fluctuation of the spatial rainfall variability from one day to the other. It is calculated as the coefficient of variation of CV_{space} across all wet days (i.e., days with rainfall > 0) in the time series:

$$CV_{cross} = \frac{\sqrt{\frac{1}{T} \sum_{t=1}^T (CV_{space,t} - CV_{in})^2}}{CV_{in}} \quad (2.3)$$

Note that for the case of uniform rainfall, both CV_{in} and CV_{cross} are equal to zero. In the fixed variability scenario, CV_{cross} is always equal to zero and different values of CV_{in} are investigated, whereas CV_{cross} is different from zero in the varying variability scenario. Meanwhile, Figure 2.3 shows the range of the event-based spatial rainfall variability (panel a) and the coefficient of variation of the spatial rainfall variability in time (panel b) from observations. The blue lines indicate the synthetic indices generated in this framework that match the majority of the cases and thus are representative of the real-world rainfall variabilities. In addition, examples of observed rainfall fields with different spatial rainfall variabilities within the considered range are displayed in the supporting information Figure 2.13.

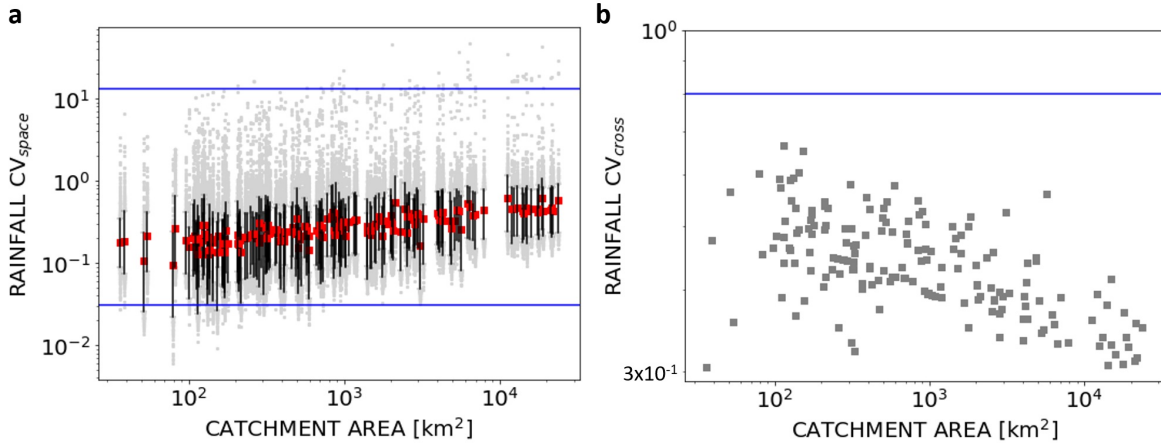


Figure 2.3: Estimated rainfall spatial variability of observations The computation is based on spatial rainfall depths in 175 German catchments between 1951 and 2013 summer wet days from the REGNIE dataset. (a) Event-based correlation of variations (CV_{space}) as a function of the basin area. The mean value (red squares), 5-95 percentiles (black boxes), and outliers (gray dots) of the CV_{space} are displayed. Horizontal blue lines indicate the lower/upper boundary generated in the synthetic work of this study. (b) Temporal correlation of variations across events (CV_{cross}) as a function of the basin area. The horizontal blue line indicates the upper boundary (from zero) generated in the synthetic work of this study.

Rainfall fields with a stochastic spatial autocorrelation were generated by a rearrangement matrix (Peleg and Morin, 2014). We first derived the spatial rainfall autocorrelation function

of the observations from the average of the integrated rainfall field (of all the wet days) for each catchment (Bacchi and Kottegoda, 1995). The Pearson correlation coefficient of rainfall depths between each HRU and the N -th nearest HRU was computed as a function of the spatial lag (i.e., N -th nearest HRU) (Grillenzoni, 2022). The correlogram curve between 0-th to N_{max} -th nearest HRU was used as the spatial rainfall autocorrelation function, in which N_{max} was determined by the corresponding correlation coefficient equal to or below zero. We used the method proposed by Peleg and Morin (2014) to generate the stochastic rainfall field with spatial autocorrelation. The synthetic rainfall depths (generated by an assigned Gamma distribution) were rearranged for each HRU to encompass the observed rainfall spatial autocorrelation property by using two generated matrices, M_1 and M_2 . M_1 is a matrix with the same dimensions as the synthetic rainfall field composed of random variables of a standard normal distribution (i.e., mean = 0 and standard deviation = 1). M_2 is a reference matrix for rearrangement with the same dimensions as M_1 . The value of each element in M_2 is determined by the following equation (which encompasses the observed rainfall spatial autocorrelation within D spatial lag):

$$M_2(x, y) = \frac{\sum_{|(X,Y)-(x,y)| \leq D} [M_1(X, Y) \cdot \rho(|(X, Y) - (x, y)|)^{d^*}]}{\sum_{|(X,Y)-(x,y)| \leq D} [\rho(|(X, Y) - (x, y)|)^{d^*}]} \quad (2.4)$$

where $\rho(|(X, Y) - (x, y)|)$ is the rainfall spatial correlation (based on the Pearson correlation coefficient) at the spatial lag between (X, Y) and (x, y) , D and d^* are calibrated parameters by the optimization of M_2 to the observed rainfall spatial autocorrelation (an example is shown in Figure 2.4d). We rearranged the synthetic rainfall field according to M_2 with the same ranking, which means the largest depth of the synthetic rainfall field was set to the place of the largest value of M_2 , and so on. The calibrated parameters we used for the five study sites are listed in Table 2.2.

Table 2.2: Parameters for the Stochastic Rainfall Spatial Autocorrelation Generator

Catchment	HRU Number	D	d_1^*	d_2^*
Holzcamp	51	10	0.3	–
Niedertrebra	429	80	1.5 (for $D \leq 40$)	1.5 (for $D > 40$)
Inkofen	1429	360	1.5 (for $D \leq 140$)	1.5 (for $D > 140$)
Heinde	438	195	1.5 (for $D \leq 30$)	1.5 (for $D > 30$)
Erfurt-Moebisburg	422	110	1.5 (for $D \leq 50$)	1.5 (for $D > 50$)

Figure 2.4 uses the Ilm River at the gauging station of Niedertrebra as an example of this stochastic simulation process. For each catchment (among five select catchments), we randomly generated five hundred realizations within the same range of spatial rainfall variability of scenario 1 and compared their effects on the tail heaviness.

In summary, I used data-based synthetic rainfall fields to investigate a wider range of variability than the one that has been observed. The framework allows for isolating and analyzing separately the roles played by spatial rainfall variability occurring within the same rainfall event or by changes of spatial variability across events and the potential effects of spatial rainfall autocorrelation, thus enabling an understanding of the effect of each of these components.

2. Impacts of Spatial Heterogeneity in Rainfall

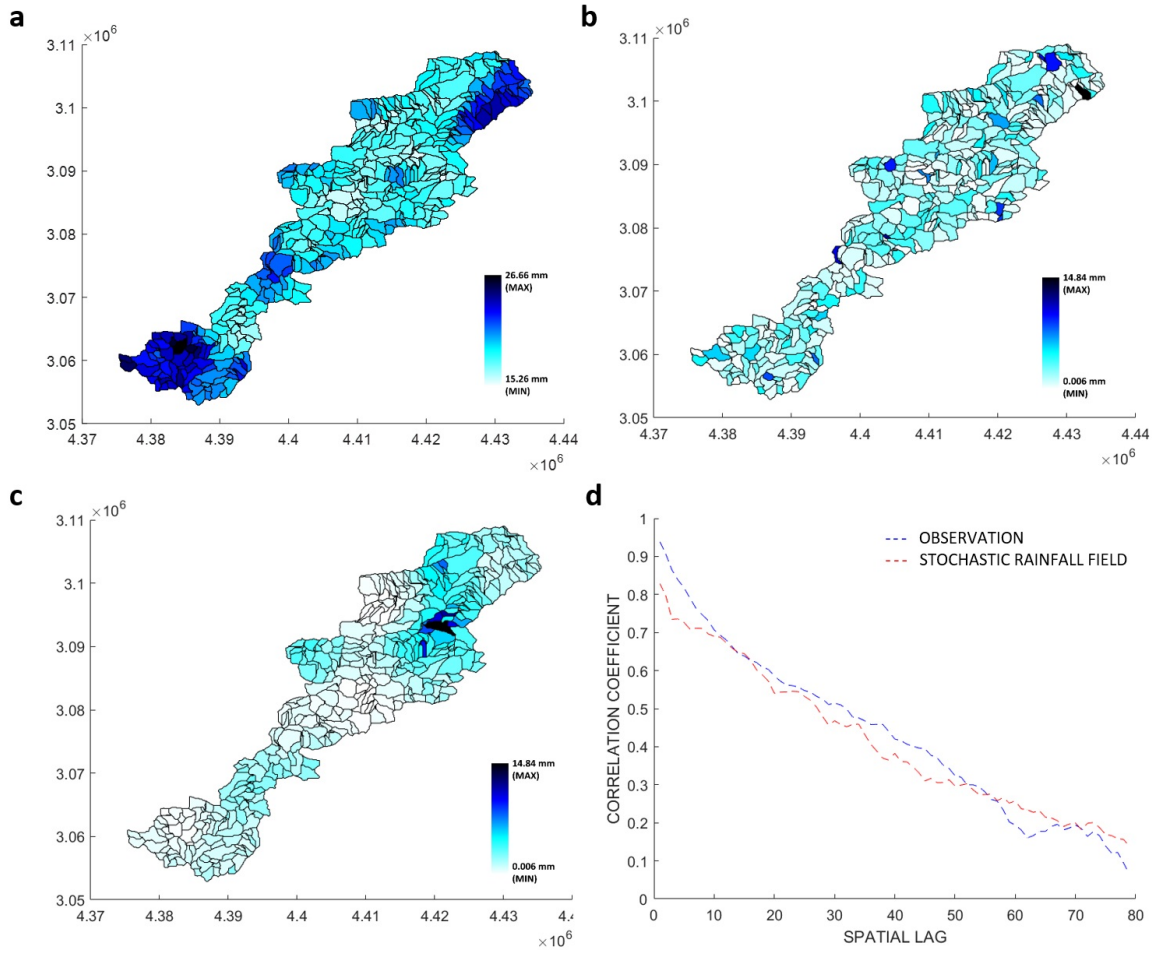


Figure 2.4: An example of data-based stochastic spatial autocorrelation simulation of the rainfall field (a) Average spatial pattern of the integrated rainfall fields of observations from 1980-2002 (summer) for Niedertrebra. (b) An example of a synthetic rainfall field generated by an imposed Gamma distribution (mean=2 mm, shape=1) with a random structure. (c) Stochastic rainfall field which is transformed from panel b and encompasses the observed spatial autocorrelation. (d) Correlograms of the observed rainfall field (panel a) and the simulated rainfall field (panel c). A terrain map of Niedertrebra based on a digital elevation model is displayed in the supporting information Figure 2.14.

2.3 Streamflow simulation

I used a well-established probabilistic model of water transport at the catchment scale (Rodriguez-Iturbe and Valdes, 1979; Rinaldo and Rodriguez-Iturbe, 1996; D’Odorico and Rigon, 2003; Rinaldo et al., 2006; Nictina et al., 2008; Rigon et al., 2016) to simulate streamflow at the outlet of catchments forced by spatially variable rainfall fields. The model accounts for three key components (see Figure 2.5): (1) the soil water balance in hillslopes, (2) the probability distributions of transit times in the hillslopes of catchment units, here assumed to be stationary in time, and (3) the response time distribution in channels derived from a geomorphological analysis of the river network.

The first component represents the water balance among rainfall inputs, evapotranspiration, and leaching to subsurface flow in each catchment unit:

$$\frac{dV(t)}{dt} = [I(t) - ET(t) - L(t)]A \quad (2.5)$$

where V is the water storage in the root zone [mm^3], I is the infiltrated rainfall depth [mm], ET is the evapotranspiration [mm], L is the leaching from the storage to subsurface flow [mm], A is the area of the catchment unit [mm^2], and t is the daily time interval [day]. The water storage V is then used to calculate soil moisture (SM) as:

$$SM(t) = \frac{V(t)}{mZA} \quad (2.6)$$

where m is the soil porosity [-] and Z is the depth of the root zone [mm].

In each catchment unit at time t , hydrological processes were further distinguished as the responses in unsaturated ($A_{us}(t)$) and saturated areas ($A_s(t)$). We estimated the fraction of saturated area to total area as a power-law function of soil moisture (Kirkby, 1975). In unsaturated areas, all the rainfall infiltrates (i.e., $I(t) = P(t) \cdot A_{us}(t)/A$, where P is the rainfall depth [mm]), whereas in saturated areas, no infiltration occurs and rainfall becomes surface runoff, and evapotranspiration equals its potential value estimated according to Thornthwaite (1948). Evapotranspiration in unsaturated areas is instead considered as a linear function of the soil moisture between the wilting point (set equal to 0.05; Ács et al., 2010) and a critical point when the water available for this process is non-limiting. Leaching of water to subsurface flow paths takes place in both unsaturated and saturated areas and is computed as a power-law function of soil moisture multiplied by a coefficient that is equal to the hydraulic conductivity (Clapp and Hornberger, 1978).

Transit times in the hillslopes of catchment units (i.e., the second key component of the model) were assumed to be exponentially distributed. Distinct distributions were considered for the surface ($f_{sup,h}$) and subsurface ($f_{sub,h}$) flow paths. Their respective parameters were determined through a power law relationship with the size of the catchment unit (Boyd, 1978; D’Odorico and Rigon, 2003; Nictina et al., 2008). All water particles in the source hillslopes were assumed to flow towards the lower node of each catchment unit.

The third component, namely the response time distributions in river networks (f_n), was instead derived from the width function of the catchment unit which explicitly considered the length of the pathway from the lower node of the catchment unit to the catchment outlet and the integration of the simplified parabolic model of de Saint Venant equation with suitable boundary conditions (Rinaldo et al., 1991; Rinaldo and Rodriguez-Iturbe, 1996).

The response time distribution of a catchment unit was finally obtained by convoluting f_n with either $f_{sup,h}$ or $f_{sub,h}$ for surface and subsurface flow, respectively. Streamflow generated from a catchment unit was finally simulated as the convolution of the runoff (i.e., the outcome of the first component) and the response time distribution (i.e., the outcome of the second and third components). Both superficial flow and subsurface flow were simulated at the outlet from each catchment unit and the integral of surface and subsurface contributions from all catchment units formed the total streamflow at the catchment outlet.

The model was calibrated on the observed streamflow series at the catchment outlet by using spatially uniform rainfall inputs. We used the shuffled complex evolution algorithm by applying a statistical parameter optimization tool of Python, SPOTPY (Houska et al., 2015), to optimize parameters.

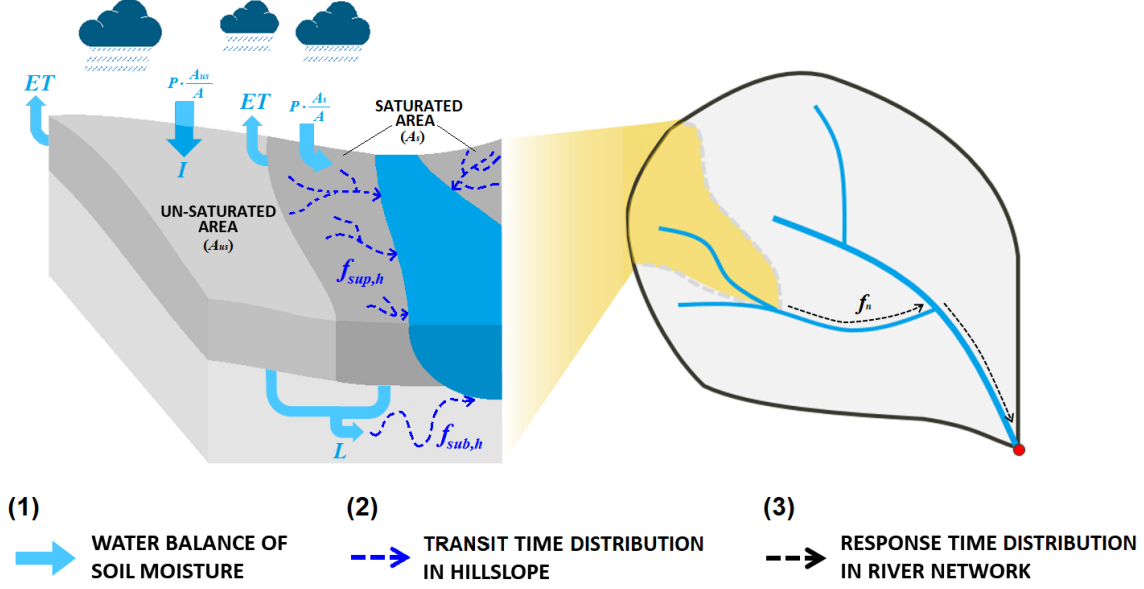


Figure 2.5: Schematic diagram of the adopted hydrological model Three key components were accounted for in the model: (1) the soil water balance in hillslopes, with main fluxes depicted through blue solid arrows, (2) the transit time distributions of surface and subsurface flows in hillslopes, graphically represented through blue dashed arrows, and (3) the response time distribution of each catchment unit in the river network, displayed by means of black dashed arrows.

2.4 Alteration of tail heaviness induced by spatial rainfall variability

I quantified the relative heaviness of the upper tail of streamflow distributions resulting from spatially variable rainfall inputs with respect to the heaviness of the distribution obtained with uniform rainfall. We used the slope S of the upper tail of the exceedance probability distribution of re-scaled daily streamflow (i.e., streamflow divided by its long-term mean value) represented in a double logarithmic plot (see supporting information Figure 2.15) as an index of tail behavior, similarly to Nerantzaki and Papalexioiu (2019) and Mushtaq et al. (2022). This is defined as:

$$S = \left| \frac{\log(0.01) - \log(0.1)}{\log(x_{0.01}) - \log(x_{0.1})} \right| \quad (2.7)$$

where $x_{0.01}$ and $x_{0.1}$ are the values of the normalized streamflow corresponding to the exceedance probabilities of 0.01 and 0.1.

I normalized S with the slope of the reference case, S_0 , here represented by the streamflow distribution obtained from the uniform rainfall case. Since the tail heaviness is proportional to the inverse of the tail slope, the index of relative tail heaviness H is:

$$H = \frac{S_0}{S} \quad (2.8)$$

Accordingly, $H > 1$ indicates a streamflow distribution with a tail that is heavier than the tail of the distribution obtained with uniform rainfall, whereas $H < 1$ labels a distribution with

a lighter tail than the distribution obtained with uniform rainfall (see supporting information Figure 2.15).

2.5 Identification of heavy-tailed streamflow distributions

Power law distributions are a type of heavy-tailed distributions (i.e., distributions with tails heavier than exponential ones) with special relevance as they occur in a wide range of natural and man-made phenomena (Newman, 2005). In particular, power law distributed variables can assume values very far from their mean. For the case of streamflow, this translates into an unneglectable chance of the occurrence of very large flows and, therefore, the possible occurrence of extreme floods.

For this reason, I investigated the possible emergence of power law distribution (i.e., distributions which are heavy-tailed in an absolute sense) as a result of spatially variable rainfall inputs. To this purpose, we applied a robust statistical framework (Clauset et al., 2009) to identify power law distributions in empirical data. The range of the tail was determined as the streamflow above a lower boundary value, which made the probability distribution of the streamflow above the boundary and the best-fit power-law model as similar as possible. Kolmogorov-Smirnov statistic was applied to quantify the distance between two probability distributions (i.e., the cumulative distribution function of data and the fitted power-law model). The associated p-value was used to evaluate whether the power law is a plausible fit for the data.

2.6 Comparable data analyses

The relationship between spatial rainfall variability and tail heaviness of the streamflow distribution was also investigated based on observed rainfall-runoff events from 175 catchments across Germany. In this case, I applied the same index of average spatial rainfall variability (i.e., CV_{in}), index of relative tail heaviness (i.e., H), and suitability of power law distributions for the data. An interpolated gridded rainfall dataset was used to identify the rainfall fields.

Rainfall-runoff events were separated from the continuous recorded series using the approach proposed by Tarasova et al. (2018). The spatial variability of rainfall for each runoff event (i.e., CV_{space}) was computed based on the spatial rainfall pattern of the total event rainfall (Tarasova et al., 2020b).

To highlight the streamflow responses from different spatial rainfall variability, we sorted the rainfall-runoff events based on their CV_{space} value. Events were then binned into groups of 100 members from low to high spatial rainfall variability, which formed a bin with a certain value of spatial rainfall variability equal to the mean CV_{space} of the 100 events (i.e., CV_{in}). Events with the largest spatial rainfall variability were considered as a separate category if their number was equal to or above 50; otherwise, they were included in the prior category.

The streamflow distribution and its tail heaviness were calculated by using the peak values of the 100 runoff events included in each bin. Due to different numbers of events for catchments in our data set, 5 to 17 bins (median: 12 bins) were determined in one catchment for analyzing the relationship between spatial rainfall variability and tail heaviness of streamflow. The

number of events per bin was decided as a tradeoff between the need of obtaining reliable estimates of tail heaviness (which improve with the number of events per bin) and a suitable representation of the relationship between spatial rainfall variability and tail heaviness (which instead requires a certain number of bins).

Linear regression between the tail heaviness of streamflow distributions and the corresponding spatial rainfall variability was performed, and the Pearson correlation coefficient and the Wald Test with t-distribution were computed by means of the `scipy.stats.linregress` tool of SciPy v1.7.1.

2.7 Chapter results and discussion

2.7.1 Spatial rainfall scenario 1: fixed variability without autocorrelation

An illustration of the effects of increasing spatial rainfall variability (CV_{in} , i.e., in a fixed variability scenario) is shown in Figure 2.6 for the Ilm River at the gauging station of Niedertrebra. The red shaded area identifies cases with heavier tails than in the case of uniform rainfall. The black dots and tendency line express the mean tail heaviness among hundred realizations performed for each degree of spatial rainfall variability. They indicate that the tail heaviness of the streamflow distribution increases only beyond a certain threshold of spatial rainfall variability, which in this case is equal to $CV_{in} = 4.7$. To quantify this threshold, we developed and applied an approach that considers the statistical significance of the linear regression slope by the Wald test with t-distribution of the test statistic and the Pearson correlation coefficient between the two variables (i.e., CV_{in} and H). It is detailed in the supporting information S2.

Red dots in Figure 2.6 indicate realizations for which the streamflow distributions are plausibly represented by means of power laws. Large amounts of red dots are visible for high spatial rainfall variability. This indicates that, besides increasing the tail heaviness of streamflow distributions, highly heterogeneous rainfall also gives rise to distributions that are heavy in an absolute sense, which have a non-neglectable probability of the occurrence of extreme flow events.

We analyzed the effects of a wide range of spatial rainfall variability in five catchments with various sizes and shapes to investigate whether these catchment characteristics affect their resilience to increasing spatial rainfall variability. Figure 2.7a shows the effects of increasing spatial rainfall variability on the tail heaviness of streamflow distributions in the Delme River basin at Holzkamp (82 km^2), Ilm River basin at Niedertrebra (887 km^2), and Amper River basin at Inkofen (2814 km^2). Results are plotted together to compare the responses in catchments with different sizes. To highlight the effects of catchment size, these catchments have similar characteristics to other potentially influencing factors, such as catchment shapes (Moussa, 2003; Sassolas-Serrayet et al., 2018; Saharia et al., 2021). All of them are elongated catchments with elongation ratio $R_e < 0.5$ and have small drainage density (i.e., $D_d < 1$).

All three catchments show increasing tail heaviness of streamflow distributions as a result of increasing spatial rainfall variability. However, the minimum value of spatial rainfall variability for which an effect arises varies, indicating different degrees of resilience in small and large catchments. The effect becomes visible at a lower rainfall variability in the smallest catchment

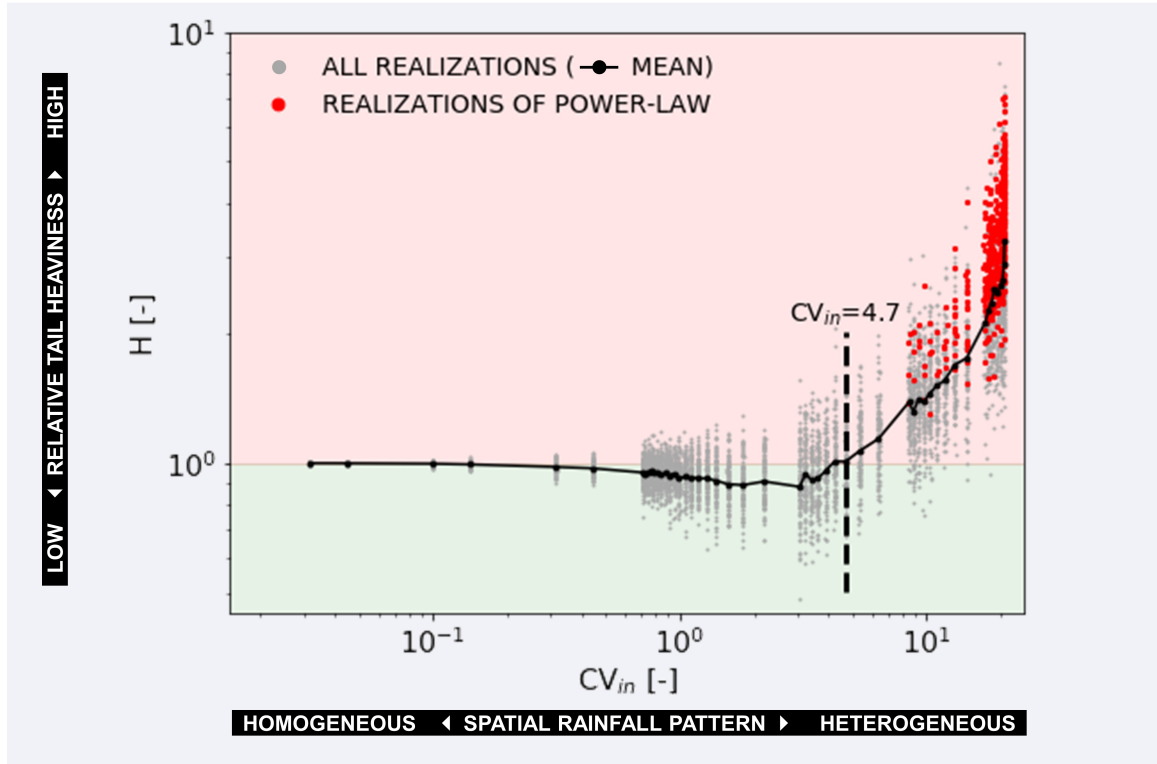


Figure 2.6: Effects of increasing spatial rainfall variability (CV_{in}) on the tail heaviness of the streamflow distribution of the Ilm River at the gauging station of Niedertrebra The red shaded area indicates heavier-tailed streamflow distributions with respect to the uniform rainfall case, whereas the green shaded area indicates lighter-tailed streamflow distributions. For each value of CV_{in} , hundred different realizations of spatial rainfall patterns on the catchment characterized by the same spatial rainfall variability are displayed with gray dots. The mean values of the tail heaviness among hundred results for each CV_{in} are marked in black and linked with a tendency line. The black-dash line indicates the identified threshold of spatial rainfall variability beyond which the tail heaviness of the streamflow-distribution increases (see supporting information S2). Red dots indicate realizations for which the streamflow distributions are not simply relatively heavier than in the uniform rainfall case but heavy in an absolute sense as they are plausibly fit by power law distributions.

(82 km²). The tail heaviness is higher in this case than for the other two catchments for all values of rainfall variability. This suggests the least resilience to spatial rainfall variability in the small catchment. The effect instead arises for similar values of spatial rainfall variability for medium and large catchments. However, the tail heaviness increases quicker in the former, thus suggesting higher resilience of large catchments to increasing spatial variability of rainfall.

Figure 2.7b shows the cumulative distribution functions (CDF) of the runoff routing distance of the three catchments. The small catchment has the narrowest spread of the distribution (i.e., all catchment units are near the outlet compared to the other two larger catchments). Figure 2.7c shows the CDFs of the runoff routing distance normalized with respect to the longest distance in each river network, which allows for a better evaluation of the relative distributions of the distance in the catchments. The black dashed line is the 1:1 line. Both small and medium catchments' CDFs are closed to uniform distributions (i.e., the S-Curves are closed to straight lines), whereas the large catchment's CDF is asymmetrical (i.e., the area below the 1:1 line is different from the area above). When the runoff routing

2. Impacts of Spatial Heterogeneity in Rainfall

distance distribution is uniform, the routing effects from the fast and delayed response are equal and balance each other, whereas this is not happening for asymmetrical distributions.

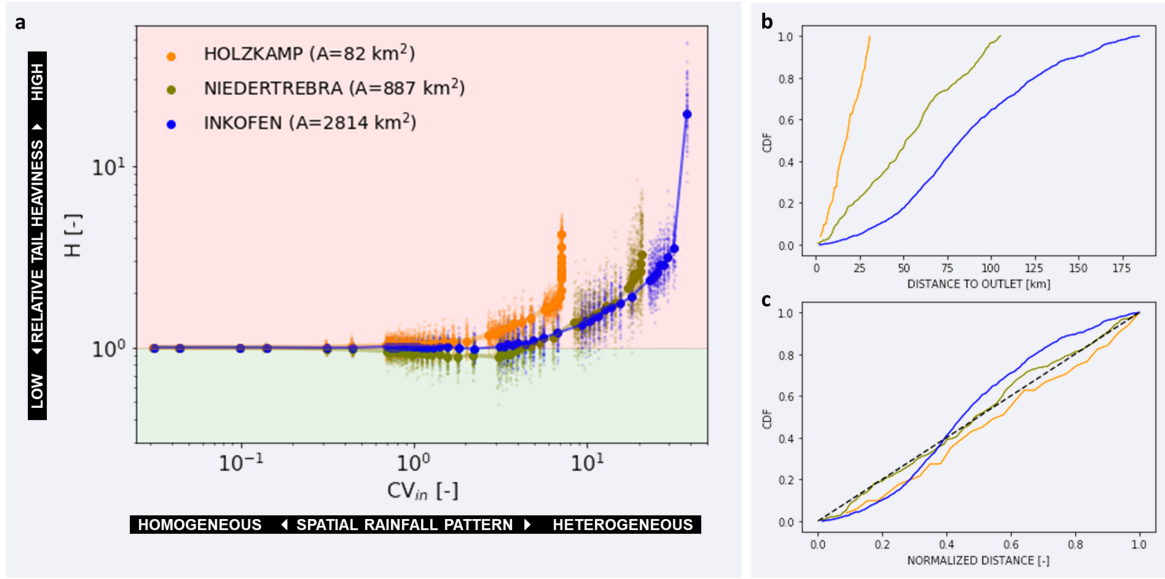


Figure 2.7: Effects of the catchment size (a) Effects of increasing spatial rainfall variability (CV_{in}) on the tail heaviness of streamflow distributions of catchments with increasing area. The red shaded area indicates heavier-tailed streamflow distributions with respect to the uniform rainfall case, whereas the green shaded area indicates lighter-tailed streamflow distributions. For each value of CV_{in} hundred different realizations of spatial rainfall patterns on the catchment characterized by the same spatial rainfall variability are displayed with small dots. The mean values of the tail heaviness among hundred results for each CV_{in} are displayed with large dots and linked with a tendency line. (b) The cumulative distribution function of runoff routing distance which is based on the DEM of each catchment. (c) The cumulative distribution function of runoff routing distance normalized with respect to the longest distance in each river network. The dash-black line is the 1:1 line which provides insights into the asymmetry of the distribution.

In Figure 2.8 I investigate the effects of increasing spatial rainfall variability on the tail heaviness of streamflow distributions in catchments with different shapes. The Ilm River basin at Niedertrebra is more elongated ($R_e = 0.45$), the Innerste River basin at Heinde is less elongated ($R_e = 0.78$), and the Unstrut River basin at Erfurt-Moebisburg is circular ($R_e = 0.90$). Figure 2.8c displays their shapes. These catchments were selected because they have similar sizes (i.e., Niedertrebra: 887 km^2 , Heinde: 898 km^2 , and Erfurt-Moebisburg: 847 km^2) and drainage densities (i.e., $D_d < 1$) to highlight the effect of catchment shape. Figure 2.8a displays the effect of increasing spatial rainfall variability on the tail heaviness of streamflow distributions in these three catchments. For the more elongated catchment, the tail heaviness increases for spatial rainfall variability growing above a certain threshold; for the less elongated and circular catchments, the tail heaviness of streamflow distributions is nearly unaffected over a wide range of increasing spatial rainfall variability. Although the elongated catchments generally have longer lengths of the main streams than circular catchments under the same area (Sassolas-Serrayet et al., 2018), Figure 2.8b shows that both the less elongated and circular catchments have clearly asymmetric runoff routing distance distributions, with more catchment units located far upstream the outlet compared to the most elongated one.

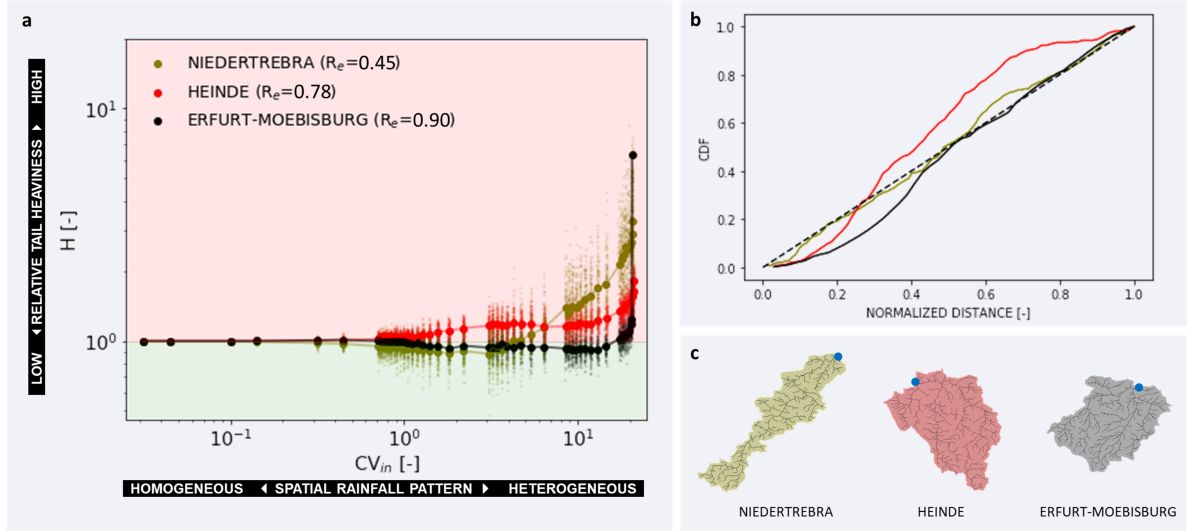


Figure 2.8: Effects of the catchment shape The Ilm River basin at Niedertrebra is the more elongated catchment with $R_e = 0.45$ (green), the Innerste River basin at Heinde is the less elongated catchment with $R_e = 0.78$ (red), and the Unstrut River basin at Erfurt-Moebisburg is the circular catchment with $R_e = 0.90$ (black). (a) Effects of increasing spatial rainfall variability (CV_{in}) on the tail heaviness of streamflow distributions of catchments with different shapes. The red shaded area indicates heavier-tailed streamflow distributions with respect to the uniform rainfall case, whereas the green shaded area indicates lighter-tailed streamflow distributions. For each value of CV_{in} , hundred different realizations of spatial rainfall patterns on the catchment characterized by the same spatial rainfall variability are displayed with small dots. The mean values of the tail heaviness among hundred results for each CV_{in} are displayed with large dots and linked with a tendency line. (b) Cumulative distribution functions of runoff routing distance normalized with respect to the longest distance in each river network. The dash-black line is the 1:1 line (starting from the minimum and ending at the maximum of the S-Curve) which provides insights into the asymmetry of the distribution. (c) Visualization of the river networks and catchment shapes of the three catchments with the related color coding. Blue dots indicate the outlets of catchments.

2.7.2 Spatial rainfall scenario 2: varying variability without autocorrelation

Recent studies point out the increasing trends of extreme events and rainfall heterogeneity due to worldwide climate change (Myhre et al., 2019; Li et al., 2020). Few researchers have addressed the issue of changing spatial rainfall variability in a catchment and its effects. Figure 2.9 shows the response of the tail heaviness of streamflow distributions to the increasing coefficient of variation of the spatial variability of rainfall across events (CV_{cross}). The red shaded area identifies cases with heavier tails than in the case of uniform rainfall. Results for catchments with different sizes and shapes are displayed in Figures 2.9a and b, respectively. In Figure 2.9a, all three tendency lines show positive slopes δ which are significantly different from zero (Wald test with t-distribution, $p < 0.05$). This indicates that the varying degree of spatial rainfall promotes an increase in the tail heaviness of streamflow distributions. The slope δ is significant ($p < 0.05$) higher for the small ($\delta = 0.71$) than the medium ($\delta = 0.20$) and large ($\delta = 0.09$) catchments. However, the consistent change of δ is not found in Figure 2.9b. In fact, the positive slope of the tendency line is significant only for the most elongated catchment ($p < 0.05$). This means that the varying degree has no effect on the streamflow distributions in both less elongated and circular catchments. I recall that these results are consistent with the previous one, which suggested that smaller/elongated catchments are less resilient to increasing spatial heterogeneity in rainfall. I also suggest that the impact of

2. Impacts of Spatial Heterogeneity in Rainfall

changing spatial variability of rainfall across events seems to be much less than the impact due to increasing spatial rainfall variability within the events.

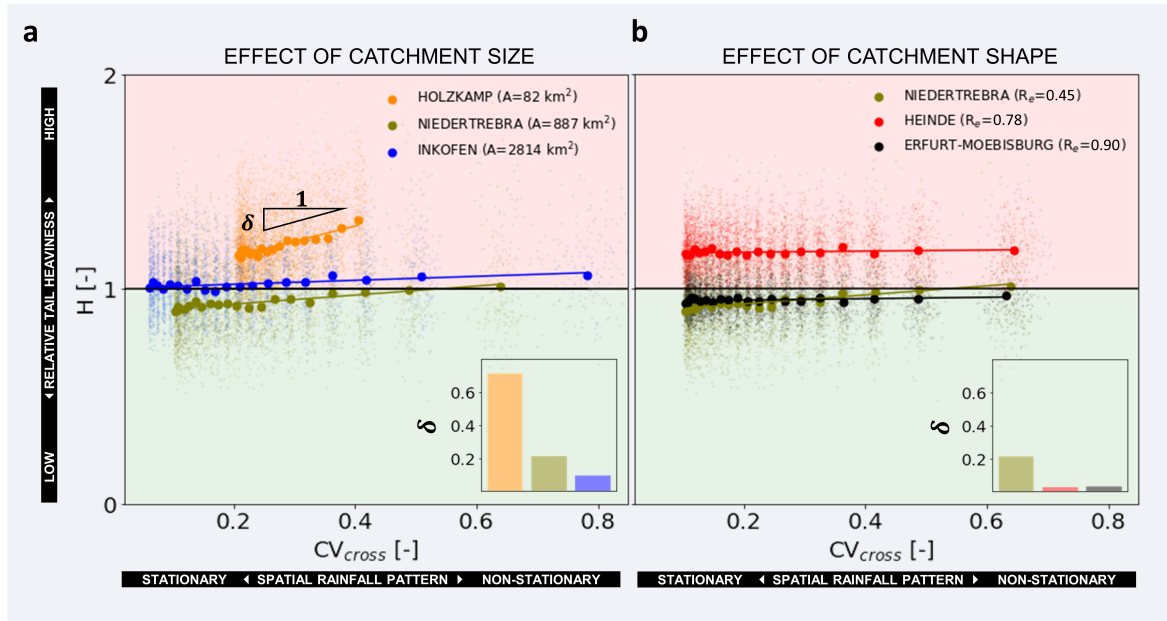


Figure 2.9: Effects of varying rainfall variability (CV_{cross}) The red shaded area indicates heavier-tailed streamflow distributions with respect to the uniform rainfall case, whereas the green shaded area indicates lighter-tailed streamflow distributions. For each value of CV_{cross} , hundred different realizations of spatial rainfall patterns on the catchment characterized by the same spatial rainfall variability are displayed with small dots. The mean values of the tail heaviness among hundred results for each CV_{cross} are displayed with large dots and linked with a tendency line. The slopes of the tendency lines are compared in the bar chart at the lower right corner of each panel. (a) Comparison of catchments with different sizes. (b) Comparison of catchments with different shapes.

2.7.3 Spatial rainfall scenario 3: fixed variability with autocorrelation

Figure 2.10 shows the effects of spatial autocorrelation of rainfall on streamflow distribution tail heaviness by comparing the results with the cases without spatial autocorrelation (i.e., random structures in space). Blue dots indicate the cases encompassing data-based stochastic spatial autocorrelation, whereas gray dots indicate those without. For all five select catchments, cases with autocorrelation respond similarly to those without with respect to the rising threshold and rising rate. However, more profound heavier tails emerge in the autocorrelation cases for high spatial rainfall variability (beyond the threshold). In particular, tail heaviness seems to congregate in the variations of heavier tails (in comparison with the cases with random structures), while the extreme cases do not exceed the expected range (inferred by the cases with random structures). These findings show that the spatial autocorrelation of rainfall does not change the general catchment response but indeed increases the probability of heavy tails for the higher variability cases. It should be noted that the small catchment (82 km^2 , panel a) shows fewer differences between the cases with and without spatial autocorrelation, in line with Zocatelli et al. (2011), who concluded that rainfall spatial organization had a limited impact on floods in smaller catchments. In addition, the circular catchment ($R_e = 0.9$, panel e) shows a significant congregation of the variations of heavier tails.

2.7.4 Corroboration through real rainfall and streamflow data

I analyzed (1) the relation between spatial variability of rainfall and tail heaviness of streamflow distributions, (2) the effects of the catchment size, and (3) catchment shape on this relation in real data to corroborate the simulation results obtained for the five select catchments and prove their transferability. The results in this section are based on the analysis of data for 175 catchments across Germany.

The increase of tail heaviness with increasing rainfall variability is seen in the data and the catchment resilience is evaluated. Figure 2.11a shows the slope and the correlation coefficient of the regression between the streamflow-distribution tail heaviness (H) and the spatial rainfall variability (CV_{in}) for every single catchment. The slope value indicates the tendency of this empirical relation, whereas the correlation coefficient evaluates its reliability. Red circles show catchments with slopes significantly ($p < 0.05$) different from zero, while gray crosses show those that are not significantly different. To evaluate the catchment resilience, we use the approach outlined in supporting information S2 for identifying the effective threshold (i.e., the threshold of CV_{in} beyond which an effect of spatial rainfall variability begins to be visible) and mark these catchments in purple color if their identified thresholds are beyond zero. In general, the majority of catchments (102 out of 175) display positive relations between spatial rainfall variability and tail heaviness. Although there are some (73 out of 175) display negative relations, all of these cases display weak tendencies (i.e., closed-to-zero slopes). Moreover, the percentage of significant cases is much higher in positive cases (37 out of 102, i.e., 36 %) than in negative (9 out of 73, i.e., 12 %). These results confirm what we found in simulations, i.e., the increasing spatial variability of rainfall determines heavier streamflow-distribution tails in most cases. Concerning the analysis of the effective threshold, there is a certain number of catchments (39 catchments shown in purple) are identified. This indicates that the catchment's resilience to spatial variability of rainfall is also revealed in the data. In total, there are 48 catchments for which we identified either significant-positive slopes for the entire range of spatial rainfall variability (shown in red) or for the spatial rainfall variability above an effective threshold (shown in purple)

The results of data analysis also show that large catchments tend to be more resilient to spatial variability of rainfall regarding the emergence of streamflow tail heaviness. Figure 2.11b shows the impact of catchment size on the positive tendency (shown by the value of the regression slope) for those 39 catchments (shown in purple in figure 2.11a) which displayed the catchment resilience (i.e., an effective threshold is identified). In general, the lower the slope is (i.e., weaker tendency), the more resilient the catchment is, and vice versa. The dependence of the regression slope on catchment size shown in figure 2.11b displays a significantly ($p < 0.05$) negative relationship with the shaded area representing the 95 % confidence interval of the tendency line. Besides, I select two extreme groups among these 39 catchments: smallest- and largest-area groups, characterized by areas below 25 (approx. 350 km^2) and beyond 75 (approx. 3800 km^2) percentiles of these 39 catchments. By computing the regression slopes of these two groups we indeed find a statistically significant difference ($p < 0.05$ using the Mann-Whitney U test via the `scipy.stats.mannwhitneyu` tool of SciPy v1.8.0). Catchments in the smallest-area group display a higher mean of regression slopes (approx. 3.96) than in the largest-area group (approx. 1.37). These findings confirm that large catchments tend to be

2. Impacts of Spatial Heterogeneity in Rainfall

more resilient (i.e., smaller regression slopes) to spatial variability of rainfall. The circle size (in figure 2.11b) indicates the elongation ratio of each catchment.

Similarly, the results of data analysis also show that circular catchments tend to be more resilient to spatial variability of rainfall regarding the emergence of streamflow tail heaviness than elongated catchments. I analyzed the impact of catchment shape on the relation between spatial variability of rainfall and tail heaviness for those 39 catchments (shown in purple in figure 2.11a) which displayed the catchment resilience (i.e., an effective threshold is identified). Figure 2.11c shows the dependence of the regression slope on the elongation ratio of catchments. It displays a significant ($p < 0.05$) negative relationship whereas the shaded area represents the 95 % confidence interval of the tendency line. As previously, we compute the regression slopes of two extreme groups: elongated and oval/circular groups, characterized by R_e below 0.7 and beyond 0.8 (Strahler, 1964; Withanage et al., 2014). I indeed find a statistically significant difference ($p < 0.05$ using the Mann-Whitney U test via the `scipy.stats.mannwhitneyu` tool of SciPy v1.8.0) in their regression slopes. Catchments in the elongated group display a higher mean of regression slopes (approx. 3.72) as compared to the oval/circular group (approx. 1.33). These findings confirm that circular catchments tend to be more resilient (i.e., smaller regression slopes) to spatial variability of rainfall. The circle size (in figure 2.11c) indicates the area of each catchment.

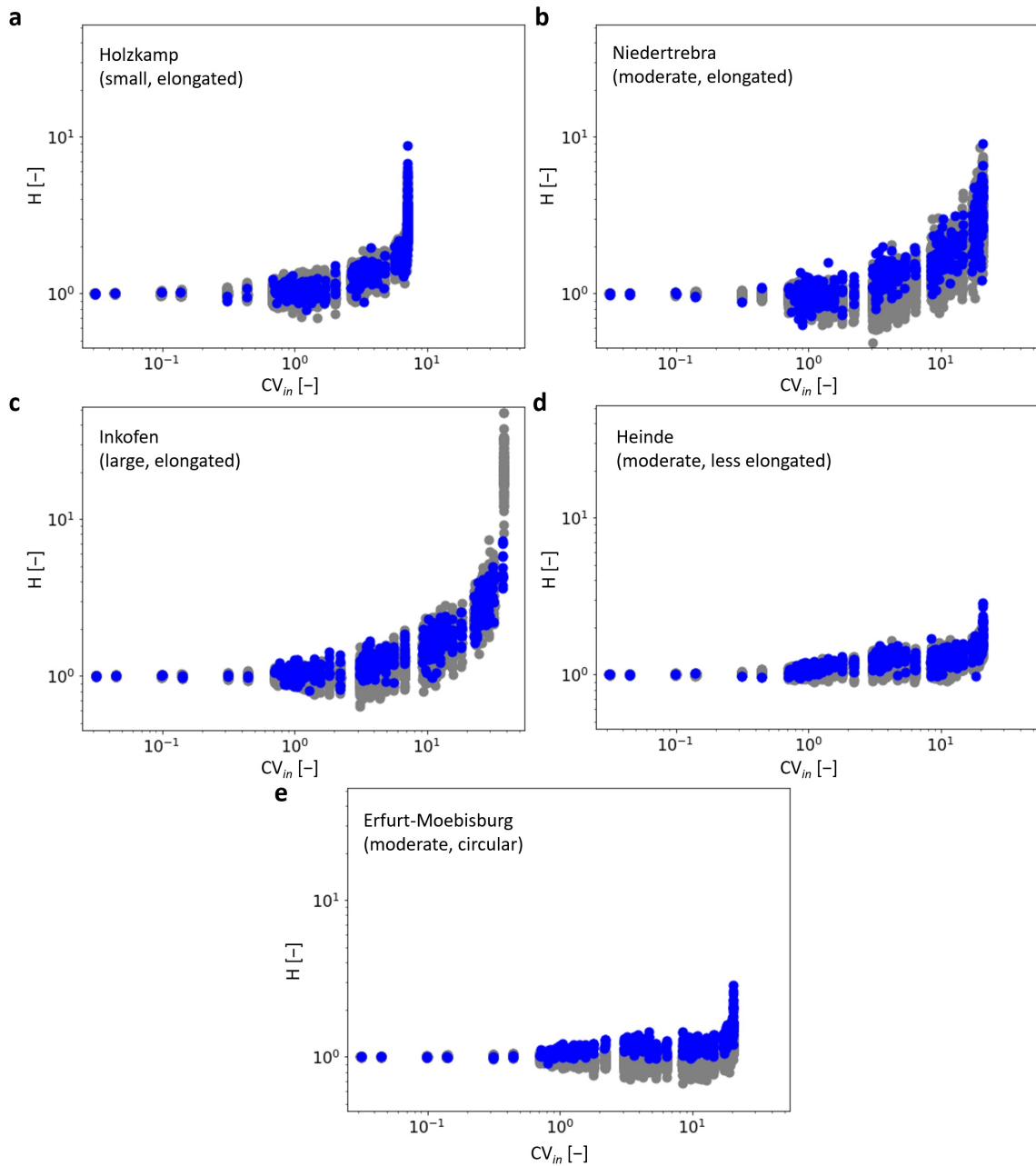


Figure 2.10: Effects of spatial autocorrelation of rainfall The relative tail heaviness (H) responses to the increasing (fixed) spatial rainfall variability (CV_{in}) under synthetic rainfall fields with/without spatial autocorrelation (blue/gray) in (a) catchment Holzkamp, (b) catchment Niedertrebra, (c) catchment Inkofen, (d) catchment Heinde, and (e) catchment Erfurt-Moebisburg.

2. Impacts of Spatial Heterogeneity in Rainfall

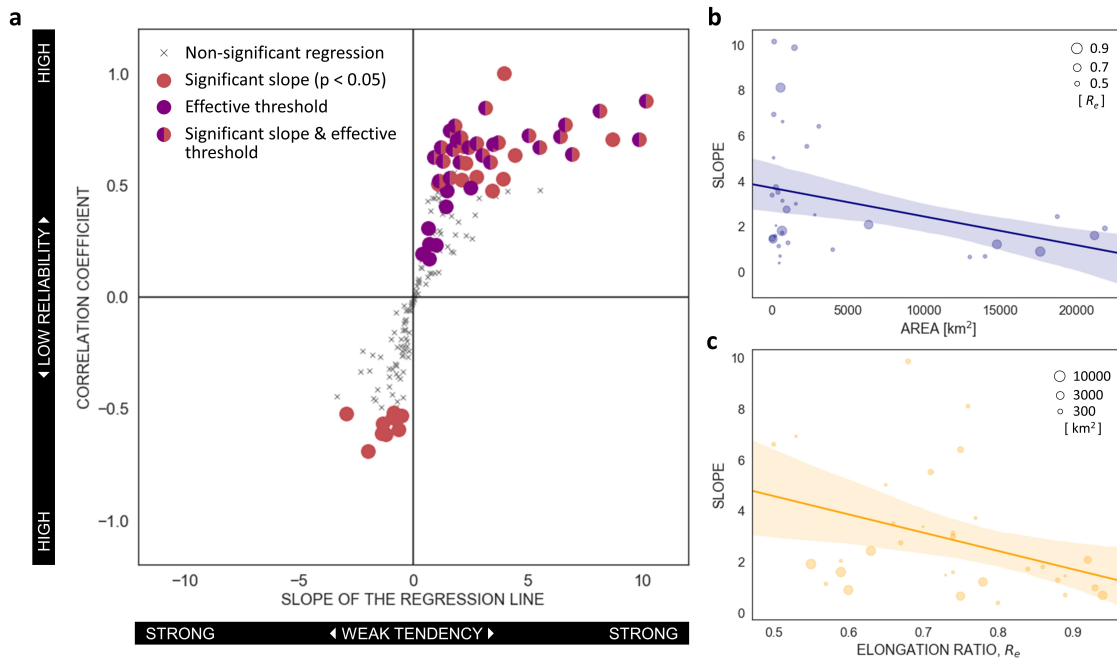


Figure 2.11: Data-based analysis of the relation between streamflow-distribution tail heaviness and spatial rainfall variability with the impacts of catchment size and shape (a) Results for 175 catchments in Germany are displayed. Each circle shows the regression slope and the correlation coefficient between the streamflow-distribution tail heaviness and the spatial rainfall variability for a single catchment. The slope value indicates the tendency of this empirical relation, whereas the correlation coefficient evaluates its reliability. Red circles indicate the catchments whose slopes were significantly greater/less than zero at a 0.05 significance level; gray crosses indicate the catchments whose slopes were not significantly different from zero; purple circles indicate catchments for which a minimum threshold of spatial rainfall variability for seeing an effect on tail heaviness was identified. (b) The relation between regression slopes and catchment sizes for 39 catchments which displays the catchment resilience (i.e., a minimum threshold was identified and shown in purple in panel a). The shaded area represents the 95 % confidence interval of the tendency line, and the circle size indicates the elongation ratio of each catchment. (c) The relation between regression slopes and elongation ratios in 39 catchments which displays the catchment resilience (i.e., a minimum threshold was identified and shown in purple in panel a). The shaded area represents the 95 % confidence interval of the tendency line, and the circle size indicates the area size of each catchment.

2.7.5 Discussion

I found that for the smaller spatial rainfall variability (the left-hand side of the threshold in Figure 2.6), both heavier and lighter tails than in the case of uniform rainfall are produced. This is an outcome of more intensive rainfall occurring either in downstream or upstream parts of the catchment for a single considered stochastic realization. The range spanned by values of the tail heaviness increases along with the increase of spatial rainfall variability, suggesting that the specific rainfall patterns have an increasing influence on the observed tail heaviness. In fact, the effect of heterogeneous rainfall patterns can be modulated by soil moisture dynamics in each catchment unit (Kling and Gupta, 2009). Viglione et al. (2010) suggested that the spatial heterogeneity of rainfall was more impactful in regions affected by saturation excess than infiltration-excess phenomena. In addition, Zhao et al. (2013) found that the spatial variability of soil moisture was linked to the spatial rainfall variability in both wet and dry conditions. In general, spatially variable rainfall produces heterogeneity in the soil moisture and thus generates more opportunities for partial saturation excess (i.e., when rainfall fall on saturated catchment units) than in the uniform rainfall case. This may increase the contribution from fast flows and cause streamflow distributions with heavier tails, as suggested by Merz and Blöschl (2003) and Bates and Aryal (2014). However, other attenuating effects, such as runoff routing through the river network, may average out the effects of these heterogeneities on the hydrograph (Merz and Blöschl, 2009; Pokhrel and Gupta, 2011). Smith et al. (2004) studied the effects of spatial rainfall variability on the hydrograph from observed data and suggested that catchments dampen the impacts of rainfall due to the filtering effect.

For the larger spatial rainfall variability (the right-hand side of the threshold in Figure 2.6), heavier tails than in the case of uniform rainfall and heavy tails (which are reliably fitted as power laws) emerge. Saharia et al. (2021) confirmed the dampening effects under lower rainfall intensities and further suggested that higher rainfall eventually subdues other geographical factors from catchments and generates flashy floods by breaking the dampening effects, which indicates a higher probability of occurrence of high-magnitude flow in a short period (Saharia et al., 2017).

On average, I evaluated a certain degree of resilience of catchments to increasingly variable spatial rainfall by identifying distinct responses of tail heaviness to the increase of spatial rainfall variability below and beyond a certain threshold. This is likely due to the increasing dominant effect of surface flow generated from partially saturated areas in the condition of high spatial heterogeneity. We used here the term ‘catchment resilience’ to indicate the capability of a catchment to buffer increasingly variable spatial rainfall patterns with limited changes in the tail heaviness of its streamflow distribution, which is an indicator of the probability of occurrence of extreme flow events. This property may have important implications for the assessment of the peril of floods under climate change, which is predicted to increase the spatial heterogeneity of rainfall both locally and globally (Donat et al., 2016; Pendergrass, 2018; Donat et al., 2019; Li et al., 2019; Myhre et al., 2019; Papalexiou and Montanari, 2019).

It was found that the size of a catchment affects its resilience. Smaller catchments tend to exhibit greater hydrological responses due to higher runoff coefficients (Joel et al., 2002; Cerdan et al., 2004). Meanwhile, the observed hydrographs at small catchments often resemble the rainfall intensity variation (Dunne and Black, 1970; Dingman, 2002), while larger catchments

2. Impacts of Spatial Heterogeneity in Rainfall

tend to more strongly smooth the spatial rainfall patterns (Obled et al., 1994; Sivapalan et al., 2003). Some studies argued that small catchments are subject to higher flood risk than large catchments due to the possible occurrence of larger extension of saturated areas (e.g., Darras et al., 2015; Rogger et al., 2012; Zhu et al., 2018). Rainfall is generally more spatially uniform in small than in large catchments, though intense convective rain cells with very high spatial variability exist and are the main cause of flooding in small catchments (Morin and Yakir, 2014). The analysis of the hydrological responses to a given spatial rainfall variability between small and large catchments remains challenging. Gupta and Waymire (1998) proposed a power-law scaling of peak discharge resulting from a single rainfall event (which is one of the main contributors to the heavy tail of a streamflow distribution) with a drainage area. They showed that the dominant factor of the scaling was the rainfall variability at smaller catchments and the river network structure with its flow dynamics at larger catchments. As we discussed in the previous section, the runoff routing may balance the impacts of partial saturation excess from catchment units caused by the spatially heterogeneous rainfall. Although partial saturation excess may increase fast flow (Viglione et al., 2010), the increase would be contrasted by a delayed response if it occurs at upstream catchment units (i.e., more remote regions from the outlet). This capability of a catchment is linked to the runoff routing distance distribution of its river network. In general, the longer the distance, the more significant the delay would be. Previous studies also suggested that the impacts of channel processes on the hydrological responses increase while the catchment scales increase (Botter and Rinaldo, 2003; Li and Sivapalan, 2011), which may therefore decrease the importance of spatial and temporal rainfall variability (Woltemade and Potter, 1994; Menabde and Sivapalan, 2001; Marchi et al., 2010). For our case, Figure 2.7b clearly shows that the large catchment has a longer routing distance, so it responds more slowly. Moreover, the normalized CDF (Figure 2.7c) of the runoff routing distance of the large catchment displays an S-Curve shape which indicates less probability when the distance is shorter and more probability when it is longer (compared to the catchments with uniform distribution of routing distance). This shows a stronger overall routing effect contributing from the distanced pathways to partial saturation excess, which may be an additional reason for the large catchment responses being slower than the medium catchment in Figure 2.4a. In agreement with previous studies, we show stronger attenuation of spatially heterogeneous rainfall from river networks of larger than smaller catchments.

It is unavoidable to have a different number of modeling units for small and large catchments or the same number of units across catchments while the unit sizes are different in the practice of a distributed (e.g HRU-based or grid-based) modeling approach. We recognize that, for the former case, the number of the units may limit the extreme value of the generated spatial variability and affect its accuracy, which influences the representative of the results (especially for the smaller catchments); whereas for the latter case, the large differences in the unit sizes across catchments will also strongly bias the estimation of spatial variability. In this study, we used the same topological approach to determine the HRU for catchments with different sizes to control the unit sizes on the same scale (see Table 2.1) while allowing the numbers to be different. Our results suggest that the generated variability is great enough to observe its impacts on tail heaviness across catchments with different sizes (e.g., Figure 2.7), and the synthetic results are confirmed by an analysis of observations on a large set of catchments.

However, we are aware that the smallest catchment we used in the synthetic analysis is 98 km^2 , and a fair number of modeling units should be confirmed for the smaller catchments in future studies.

I found that another geomorphological factor of a catchment that may affect the catchment resilience is its shape. Our findings show that the most elongated catchment, which has the longest routing distance of the mainstream, displays the least resilience to spatial rainfall variability (Figure 2.8a). Elongated catchments are frequently found to display approximately symmetric runoff routing distance distributions (Collischonn et al., 2017; Roy et al., 2022). Figure 2.8b shows that the elongated catchment has an apparent symmetric runoff routing distance distribution, while the other two have asymmetric distributions. This finding confirms the previous discussion and suggests more catchment resilience to spatial rainfall variability may exist in circular catchments than elongated ones due to markedly asymmetric runoff routing distance distributions. Moreover, Post and Jakeman (1996) found that the draining of subsurface flow is easier in circular catchments. This may lead to lower storage and less occurrence of saturated excess. Moussa (2003) showed that more elongated catchments have higher peaks and shorter recession times in their hydrographs than circular catchments. Roy et al. (2022) found that elongated catchments tend to have shorter response times for partial contributions in catchments due to topological characteristics. In line with these studies, our findings extend the knowledge of the effects of catchment shape on hydrological response and further show that elongated catchments tend to have less catchment resilience to spatially variable rainfall.

When I further implemented the variations of the spatial rainfall variability in time, we found it affects the emergence of heavier tails lightly. From the numerical point of view, this may be attributed to the averaging between high and low spatial variabilities based on the randomness of the change, which does not result in a significant enhancement of rainfall through this process. We acknowledge that this work presents a preliminary step toward the effects of the dynamics of spatial rainfall variability on floods, and further investigations of the changing rate, temporal correlation of the changing magnitude, and the linkages to soil moisture dynamics are preferable.

Another process we implemented in this work is spatial rainfall autocorrelation. This process enhances the emergence of heavier tails, which is likely due to the reduction of the attenuating effect of river networks because of the aggregation of the extreme rainfall cells. As a result, the effects of saturated excess on the hydrological responses overlay each other for they occurred in the closer regions rather than offsetting each other through the opposite routing distances. In fact, Zoccatelli et al. (2011) showed that the bimodal spatial distribution of rainfall is rare in the real case. In particular, the degree of the spatial rainfall concentration has a significant impact on the flood peak, while the location of the rainfall centroid mainly affects the hydrological response time for cases with extremely heavy rainfall. Along the same line, I found that the effects of spatial autocorrelation are more profound under conditions of high spatial rainfall variability, which generate extreme heavy rainfall locally. In addition, we have suggested that the attenuating effects of river networks are more significant in larger/circular catchments; in agreement with this, the reduction of the attenuating effect arising from spatial rainfall autocorrelation is also more evident in these catchments, which is shown by the results.

I included analyses of observations on a large set of real catchments and indeed validated our findings from the synthetic work. The results of data analyses show that the tail heaviness increases significantly when rainfall spatial variability increases in the majority of catchments. Meanwhile, most larger or more circular catchments show stronger resilience in this process. It is worthwhile to mention that the findings based on the synthetic works, in which we disentangled the climate processes (i.e., spatially variable rainfall, varying variability in time, and autocorrelation in space), are confirmed by observations in which all of these processes are involved. However, we acknowledge that there may still be several physical processes that are effective, and further studies are needed to clarify the issues. For example, we have shown that the smaller catchments have less resilience to spatial rainfall variability. In agreement with our findings, the hydrological responses in urban or small rural catchments are found to be very sensitive to spatial rainfall variability (Peleg et al., 2017), yet a sub-daily/hourly temporal scale may be more appropriate for cases of these very small catchments due to the quick response time of their hydrological processes.

I recall here the ongoing debate surrounding the impacts of spatially heterogeneous rainfall on runoff generation. Previous studies often indicated a significant effect of spatial rainfall variability on hydrological responses (Kim et al., 2008; Zoccatelli et al., 2010; Looper and Vieux, 2012; Mei et al., 2014), whereas others claimed relatively little impact (Brath et al., 2004; Nicótina et al., 2008; Adams et al., 2012). Our findings highlight the existence of a threshold that distinguishes between ranges of spatial rainfall variability, which may modulate or not affect the tail behavior of streamflow distributions. The identified threshold also indicates the attenuating effects in catchments (Smith et al., 2004; Merz and Blöschl, 2009; Pokhrel and Gupta, 2011; Saharia et al., 2021). In agreement with previous studies (which are however mostly analyzing single events) (Dunne and Black, 1970; Obled et al., 1994; Sivapalan et al., 2003; Darras et al., 2015) I show that larger or more circular catchments tend to be more resilient to spatial rainfall variability based on the data analysis of observed rainfall and discharge in a large set of catchments in Germany and corroborate the simulation results.

2.8 Chapter conclusions

In this study, I evaluated the impacts of spatial rainfall variability on the increasing hazard of high flows. I synthetically generated spatially variable rainfall in one fundamental scenario and two sub-scenarios. Scenario 1: rainfall with fixed spatial variability (i.e., constant spatial rainfall variability across events and random spatial structures). Scenario 2: Spatially variable rainfall with varying variability in time (while the spatial structure is random). Scenario 3: Spatially variable rainfall with autocorrelation in space (while the spatial rainfall variability is fixed). A continuous probabilistic model of water transport at the catchment scale was used to simulate the hydrological response. The effects of spatial rainfall variability (in all scenarios) on the tail heaviness of streamflow distributions were investigated in five select catchments under different catchment sizes (i.e., small, medium, large) and catchment shapes (i.e., more elongated, less elongated, circular). Furthermore, we analyzed the relation between spatial rainfall variability and tail heaviness of streamflow distributions by using recorded data from

175 river catchments to validate the simulation results and test their transferability. The key conclusions of the study are:

1. Increasing spatial variability of rainfall determines enhanced hazard of high flows (i.e., heavier tails of flow distributions) only beyond a certain increase threshold. Both the value of this threshold and the growth rate of tail heaviness beyond the threshold indicate the resilience of catchments to spatially variable rainfall.
2. Small or elongated catchments show less resilience to increasing spatial variability of rainfall compared to large or circular catchments. High asymmetry of the distribution of runoff routing distances along river networks identified for large and circular catchments (with more contributions from upstream than downstream catchment units) is likely to provide more resilience to increasing spatial variability of rainfall.
3. In line with the previous results, smaller or more elongated catchments are more influenced by the increasing variations of the spatial rainfall variability across events. However, this process seems to influence less the tail of streamflow distributions than the spatial variability of rainfall during the event.
4. Spatial rainfall autocorrelation does not change the overall response of the tail heaviness to spatial rainfall variability. However, it indeed enhances the emergence of heavy tails under the high spatial rainfall variabilities, which is especially evident in larger/circular catchments.
5. Analyses of daily rainfall-runoff records for a large set of catchments in Germany agree with the simulation results, showing positive correlations between spatial rainfall variability and streamflow-distribution tail heaviness in the majority of the case studies. Data analyses also confirm that small/elongated catchments are less resilient to increasing spatial rainfall variability, as previously found by means of synthetic simulations.

This work pinpoints the role played by the spatial heterogeneity of rainfall in controlling the emergence of heavy-tailed distributions of streamflow. It thus establishes a link between expected alterations of rainfall caused by the ongoing climate change, such as more localized rain, and the resulting modifications in the frequency of large event magnitudes, with implications for the hazard of high flows in different river basins. Based on the findings of this work, we show the catchment geomorphological characteristics may dampen or enhance the climate impacts on the emergence of floods and a better understanding of these processes is much more valuable for catchment flood management in climate change. In addition, more sophisticated processes and approaches (e.g., spatiotemporal autocorrelation of rainfall, sub-daily studies for urban or small catchments, effects of rainfall storm pathways) and their linkages to the rainfall-runoff dynamics are helpful for accurately evaluating the impacts of climate variability on flood hazards.

2.9 Supporting information

This supporting information contains three supplementary methods, five figures, and three tables.

S1. Identification of suitable probability distribution for REGNIE rainfall fields

I computed the chi-square statistic (χ^2) for the goodness-of-fit test to identify the suitable probability distribution for REGNIE rainfall fields in 5-tested catchments. REGNIE is a daily data set that is a 1-km^2 interpolated rainfall field estimated from point observations through multiple regression provided by the German Weather Service (Rauthe et al., 2013). Ten probability distributions (see figure 2.12) which were often applied in climatology or environmental research were tested (e.g., Allard and Soubeyrand, 2012; Ayalew et al., 2014; Ben-Gai et al., 1998; Clauset et al., 2009; Gaitan et al., 2007; Hu et al., 2019; Li et al., 2014; Li and Shi, 2019; Ng et al., 2019; Tabari, 2020; Ye et al., 2018). To evaluate the goodness-of-fit between distributions on the same scale we normalized the χ^2 of each distribution by the total χ^2 of all tested distributions for each rainfall day (i.e., re-scale each χ^2 into $[0, 1]$). To identify the general expression of rainfall-field distributions from all days in the time series we computed the summation of all the normalized χ^2 (i.e., all rainfall days) as the total normalized χ^2 which showed the overall size of the discrepancies between the rainfall fields and the tested probability distribution. Finally, the most suitable probability distribution was determined by the lowest summation of the normalized χ^2 (i.e., the distribution has the least discrepancies to the rainfall fields). The computing procedure for each catchment is listed as follows:

1. Compute χ^2 for 10-tested distributions day by day (rainfall days).
2. Normalized each χ^2 by the total χ^2 for 10-tested distributions of each day.
3. Sum up the normalized χ^2 of all days for each distribution.
4. Select the distribution which has the least summation in step 3.

I estimated the suitable probability distribution for the rainfall fields of the summer data of the daily time series from 1980 to 2002 in 5-tested catchments. Figure 2.12 shows the total normalized χ^2 of 5-tested catchments for 10 distributions. Gamma, Exponential-Normal, Generalized Extreme Value (GEV), and Skew-Normal distributions appear well goodness-of-fit to the rainfall fields. We finally selected the Gamma distribution as the most suitable one for our simulation because it is a 2-parameter distribution while the others are 3-parameter distribution, which has less uncertainty.

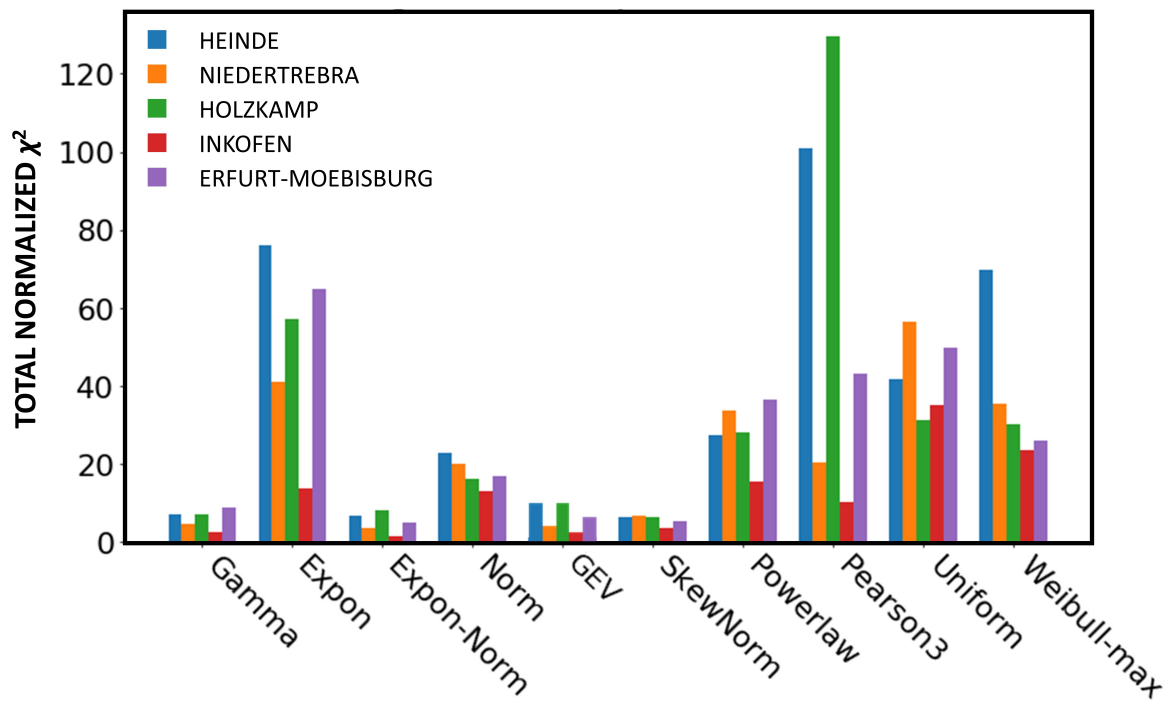


Figure 2.12: Overall goodness-of-fit of probability distributions for the rainfall fields in REGNIE data set The total normalized chi-square statistic (χ^2) is the summation of normalized χ^2 (which is normalized by the total χ^2 of 10-tested distributions for each rainfall day) of all rainfall days from the summer data of the daily time series from 1980 to 2002 in five-tested catchments.

S2. Identification of an effective threshold of the increasing spatial rainfall variability

I proposed statistical-cascade criteria to test if an effective threshold exists in the results of the scatter plot of spatial rainfall variability (i.e., CV_{in}) and the streamflow-distribution tail heaviness (i.e., H). In this criteria, we separately analyzed the trend between CV_{in} and H below and beyond a certain CV_{in} value by computing their linear regressions. All the values of CV_{in} from the results are tested as potential thresholds (CV_{thre}). We call the sequence below a potential threshold as sequence α whereas the sequence beyond it is sequence β . The cascade criteria are structured as follows:

- Criteria 1. Significance Test: the slope of the tendency line of α is not significant ($p \geq 0.05$) different from zero while the slope of β is tested to be significantly ($p \leq 0.05$) larger than zero.
- Criteria 2. Optimization of Correlation: we select the best value of CV_{thre} among all the qualified CV_{thre} from criteria 1, which is determined by maximizing the optimization matrix: $(1 - r_\alpha) + r_\beta$, where r_α is the correlation coefficient of α and r_β is the correlation coefficient of β . We ensure that CV_{in} and H correlated to each other as strongly as possible in sequence β (i.e., beyond the threshold) while they correlated to each other as weakly as possible in sequence α (i.e., below the threshold).
- Criteria 3. Increasing Trend: the maximum value of H in sequence α has to be smaller than the maximum value of H in sequence β , by which we ensure the general increasing trend of H toward CV_{in} .

The unique effective threshold is identified as the CV_{thre} if and only if it is qualified by criteria 1 \rightarrow criteria 2 \rightarrow criteria 3. The Pearson correlation coefficient and the Wald Test with t-distribution were computed by means of the `scipy.stats.linregress` tool of SciPy v1.7.1. for the linear regression.

S3. Other supplementary figures and tables

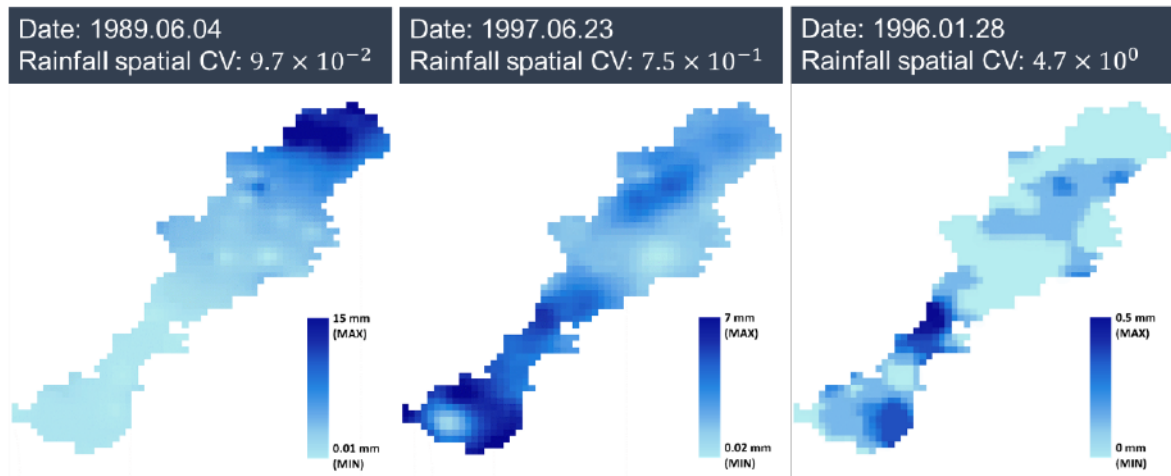


Figure 2.13: Examples of observed rainfall fields with the increasing spatial CV of the Ilm River at the gauging station of Niedertrebra

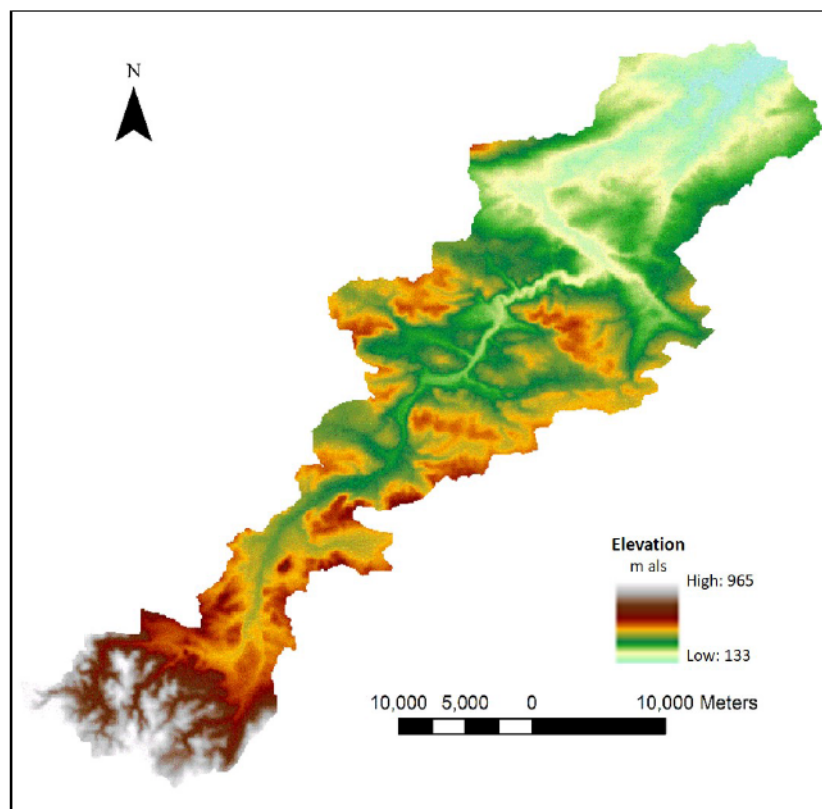


Figure 2.14: The terrain map of Niedertrebra based on a digital elevation model

2. Impacts of Spatial Heterogeneity in Rainfall

Table 2.3: Parameters for fixed variability scenarios (1, 3): shape parameter (k)

#	k [-]	#	k [-]	#	k [-]	#	k [-]
1	1000	14	1.3	27	0.09	40	0.005
2	500	15	1.2	28	0.08	41	0.004
3	100	16	1.1	29	0.07	42	0.003
4	50	17	1	30	0.06	43	0.002
5	10	18	0.9	31	0.05	44	0.001
6	5	19	0.8	32	0.04	45	0.0009
7	2	20	0.7	33	0.03	46	0.0008
8	1.9	21	0.6	34	0.02	47	0.0007
9	1.8	22	0.5	35	0.01	48	0.0006
10	1.7	23	0.4	36	0.009	49	0.0005
11	1.6	24	0.3	37	0.008	50	0.0004
12	1.5	25	0.2	38	0.007	51	0.0003
13	1.4	26	0.1	39	0.006	52	0.0001

Table 2.4: Parameters for the varying variability scenario (2): shape parameter (k), central shape parameter (k_0), and random fluctuation range ($\pm b$)

#	k_0 [-]	$\pm b$	k [-]	#	k_0 [-]	$\pm b$	k [-]
1	10^{-1}	0.05×10^{-1}	(0.095, 0.105)	11	10^{-1}	0.55×10^{-1}	(0.045, 0.155)
2	10^{-1}	0.10×10^{-1}	(0.090, 0.110)	12	10^{-1}	0.60×10^{-1}	(0.040, 0.160)
3	10^{-1}	0.15×10^{-1}	(0.085, 0.115)	13	10^{-1}	0.65×10^{-1}	(0.035, 0.165)
4	10^{-1}	0.20×10^{-1}	(0.080, 0.120)	14	10^{-1}	0.70×10^{-1}	(0.030, 0.170)
5	10^{-1}	0.25×10^{-1}	(0.075, 0.125)	15	10^{-1}	0.75×10^{-1}	(0.025, 0.175)
6	10^{-1}	0.30×10^{-1}	(0.070, 0.130)	16	10^{-1}	0.80×10^{-1}	(0.020, 0.180)
7	10^{-1}	0.35×10^{-1}	(0.065, 0.135)	17	10^{-1}	0.85×10^{-1}	(0.015, 0.185)
8	10^{-1}	0.40×10^{-1}	(0.060, 0.140)	18	10^{-1}	0.90×10^{-1}	(0.010, 0.190)
9	10^{-1}	0.45×10^{-1}	(0.055, 0.145)	19	10^{-1}	0.95×10^{-1}	(0.005, 0.195)
10	10^{-1}	0.50×10^{-1}	(0.050, 0.150)	20	10^{-1}	1.00×10^{-1}	(0.000, 0.200)

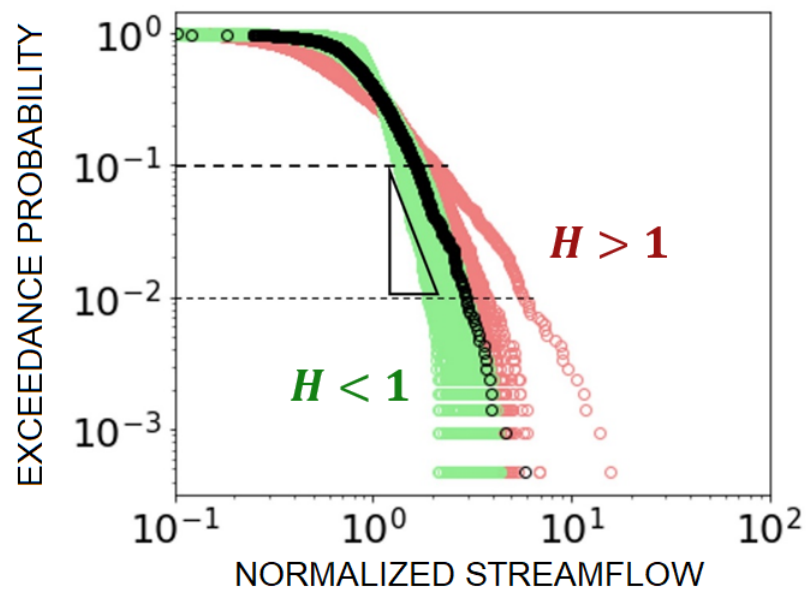


Figure 2.15: Definition of the index of relative tail heaviness The exceedance probability distribution of the normalized streamflow (i.e., divided by the long-term mean daily flow) is displayed in a double logarithmic plot. Black dots represent the case with uniform rainfall. Exceedance probability distributions resulting from spatially variable rainfall are marked with either red or green dots depending on their values of the index of relative tail heaviness H . All the cases marked in red have $H > 1$, which signifies heavier tails than in the case of uniform rainfall; all the cases marked in green have $H < 1$, which indicates lighter tails than in the case of uniform rainfall.

2.10 Data availability

For providing the discharge data for Germany, we are grateful to the Bavarian State Office of Environment (LfU, <https://www.gkd.bayern.de/de/fluesse/abfluss>) and the Global Runoff Data Centre (GRDC) prepared by the Federal Institute for Hydrology (BfG, <http://www.bafg.de/GRDC>). Climatic data can be obtained from the German Weather Service (DWD; <ftp://ftp-cdc.dwd.de/pub/CDC/>). The digital elevation model can be retrieved from Shuttle Radar Topography Mission (SRTM; <http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1>). Land use data are available from the Copernicus monitoring system (<http://land.copernicus.eu/pan-european/corine-landcover>). Characteristics of separated rainfall-runoff events for every streamflow gauge used in the analysis can be found in Data Set S1 of Tarasova et al., 2018 (<https://doi.org/10.1029/2018WR022588>).

3

Insights from Common Streamflow Dynamics

It is crucial to accurately predict the probability of extreme flood occurrences, as they are often substantially linked to flood hazards. These hazards persistently pose threats to livelihoods (McDermott, 2022; Rentschler et al., 2022) and result in severe economic losses (Bevere and Remondi, 2022) on a global scale. To assess the flood hazards of extreme events, flood frequency analysis is the central and commonly used tool, which is usually achieved by parametrically fitting a selected probability distribution on flow maxima, e.g., the annual maximum flood (Villarini and Smith, 2010), or peak-over-threshold series (Pan et al., 2022). Selecting a suitable distribution that can properly describe (or predict) the extreme events is, however, often challenging due to the notable uncertainties caused by the lack of data in the maxima approach (Papalexiou and Koutsoyiannis, 2013; Hu et al., 2023). The upper-tailed behavior (which we will refer to as 'tail behavior' throughout the paper for simplicity) of the underlying distribution critically determines the accuracy of the extreme events. If a catchment has the potential for heavy-tailed flood behavior but this characteristic is not accounted for in the selection of probability distributions, the probability of extreme floods may be significantly underestimated (Merz et al., 2022). This can lead to disastrous floods and severe damages (Merz et al., 2021). Therefore, correctly identifying the tail behavior of flood distributions is crucial for avoiding potential underestimation of extreme floods.

The tail heaviness of an empirical distribution is typically estimated through graphical or statistical methods, although both methods have limitations. Graphical methods, such as log-log plots (Beirlant et al., 2004), generalized Hill ratio plots (Resnick, 2007; El Adlouni et al., 2008), and mean excess functions (Embrechts et al., 1997; Nerantzaki and Papalexiou, 2019), are less objective and efficient for large-scale analyses (Cooke et al., 2014). In contrast, statistical methods, such as parametric metrics that fit distributions to the observed data (Papalexiou et al., 2013; Seckin et al., 2011; Smith et al., 2018; Villarini and Smith, 2010) and non-parametric metrics like the upper tail ratio (Lu et al., 2017; Smith et al., 2018; Villarini et al., 2011; Wang et al., 2022), Gini index (Eliazar and Sokolov, 2010; Rajah et al., 2014),

and obesity index (Cooke and Nieboer, 2011; Sartori and Schiavo, 2015), provide a more objective insight into tail behavior. However, obtaining reliable estimates from these methods requires long data records (Papalexiou and Koutsoyiannis, 2013), which can be challenging globally (Lins, 2008) and may cause bias when comparing data across sites with different record lengths (Cunderlik and Burn, 2002; Wietzke et al., 2020). To reduce uncertainty, especially in estimating extremes, it is recommended to analyze ordinary dynamics instead of focusing solely on maximum values (Marani and Ignaccolo, 2015; Mushtaq et al., 2022), and to investigate the underlying factors that contribute to extreme events (Wilson and Toumi, 2005; Tarasova et al., 2020a; Merz et al., 2022).

Floods are conventionally thought to be triggered by rainfall, and numerous studies have contributed to an improved understanding of rainfall extremes (e.g., Koutsoyiannis, 2004a,b; Martinez-Villalobos and Neelin, 2021; Koutsoyiannis, 2022). However, several studies have clarified that rainfall extremes do not necessarily translate into flood extremes (e.g., McCuen and Smith, 2008; Pall et al., 2011; Hall et al., 2014; Archfield et al., 2016; Rossi et al., 2016; Zhang et al., 2016; Hodgkins et al., 2017; Sharma et al., 2018). For instance, McCuen and Smith (2008) showed that skewed rainfall distributions do not always produce skewed flood distributions. They proposed that catchment responses and storage dynamics contribute to the generation of flood extremes. This view was supported by Sharma et al. (2018), who argued that despite a significant increase in rainfall extremes, a corresponding increase in flood extremes was not observed. The thorough review of Merz et al. (2022) concluded that while rainfall plays a primary role in generating runoff, the emergence of flood extremes is largely determined by catchment responses and water balance. Given these premises, an appropriate approach for describing runoff and its extremes should be rooted in the dynamics of soil moisture and rainfall-runoff processes within catchments.

This study aims to investigate whether a suitable descriptor of the tail behavior of flood distributions exists by exploring the intrinsic hydrological dynamics of the flow regime. Currently, widely-used metrics for tail behavior estimation of flood distributions do not incorporate such a physical description, to the best of our knowledge. Using this descriptor as a proxy for estimating heavy-tailed flood behavior, rather than relying solely on statistical analysis of extreme events, this study aims to bridge this gap and improve the accuracy and reliability of tail behavior estimation for flood distributions. I begin the analysis with a mechanistic description of hydrological processes. I subsequently distinguish between the key processes generating heavy and nonheavy tailed behavior of flood distributions and propose a physical descriptor for heavy-tailed flood behavior which is based on common streamflow dynamics. I verify its ability to identify heavy-tailed flood behavior and its robustness in datasets with decreasing lengths through numerous case studies across Germany, encompassing various climate and physiographic characteristics. This confirms the practical transferability and stability of the descriptor.

3.1 Asymptotic analysis of common streamflow dynamics

I describe key hydrologic processes occurring at the catchment scale and the resulting probability distributions of streamflow and floods by means of the PHysically-based Extreme Value (PHEV)

distribution of river flows (Basso et al., 2021). This framework is grounded on a well-established mathematical description of precipitation, soil moisture, and runoff generation in river basins (Laio et al., 2001; Porporato et al., 2004; Botter et al., 2007b, 2009). Rainfall is described as a marked Poisson process with frequency λ_p [T^{-1}] and exponentially distributed depths with average α [L]. Soil moisture increases due to rainfall infiltration and decreases due to evapotranspiration. The latter is represented by a linear function of soil moisture between the wilting point and an upper critical value expressing the water holding capacity of the root zone. Runoff pulses occur with frequency $\lambda < \lambda_p$ when the soil moisture exceeds the critical value. These pulses replenish a single catchment storage, which drains according to a nonlinear storage-discharge relation. The related hydrograph recession is described via a power law function with exponent a [-] and coefficient K [L^{1-a}/T^{2-a}] (Brutsaert and Nieber, 1977), which allows for mimicking the joint effect of different flow components (Basso et al., 2015). Such a description of runoff generation and streamflow dynamics was successfully tested in a variety of hydro-climatic and physiographic conditions (Arai et al., 2020; Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Mejía et al., 2014; Müller et al., 2014; Müller et al., 2021; Pumo et al., 2014; Santos et al., 2018; Schaeffli et al., 2013).

PHEV provides a set of consistent expressions (Basso et al., 2021) for the probability distributions of daily streamflow, ordinary peak flows (i.e., local flow peaks occurring as a result of streamflow-producing rainfall events; sensu Miniussi et al., 2020), and floods (i.e., flow maxima in a certain timeframe; Basso et al., 2016). The probability distribution of daily streamflow q can be expressed as (Botter et al., 2009):

$$p(q) = C_1 \cdot q^{-a} \left(e^{\frac{-1}{\alpha K(2-a)} \cdot q^{2-a}} \right) \left(e^{\frac{\lambda}{K(1-a)} \cdot q^{1-a}} \right) \quad (3.1)$$

where C_1 is a normalization constant. The probability distribution of ordinary peak flows and flow maxima can be expressed as $p_j(q)$ and $p_M(q)$, respectively (Basso et al., 2016):

$$p_j(q) = C_2 \cdot q^{1-a} \cdot e^{-\frac{q^{2-a}}{\alpha K(2-a)}} \cdot e^{\frac{q^{1-a}}{K(1-a)}} \quad (3.2)$$

$$p_M(q) = p_j(q) \cdot \lambda \tau \cdot e^{-\lambda \tau \cdot D_j(q)} \quad (3.3)$$

where $D_j(q) = \int_q^\infty p_j(q) dq$, τ [day] is the duration of the considered time frame, C_2 is a normalization constant.

Notably, the mathematical expression of flow distributions provided by the PHEV framework are composed of a power law and two stretched exponential distributions, although it's important to note that PHEV doesn't assume a specific probability distribution for streamflow representation. The use of stretched exponential distributions introduces greater flexibility in capturing tail behavior compared to the exponential distribution. Depending on its parameter values, the stretched exponential distribution can display either light-tailed or heavy-tailed behavior, whereas the exponential distribution consistently exhibits a light-tailed behavior. In fact, recent studies (Basso et al., 2016; 2021; 2023) have substantiated and documented PHEV's efficacy in representing high flow behaviors.

Taking the limit of Equation 3.1 for $q \rightarrow +\infty$ provides indications on the tail behavior of the flow distribution (Basso et al., 2015). This is determined by the three terms in the

3. Insights from Common Streamflow Dynamics

equation, namely, one power law and two exponential functions, which behave differently depending on the value of the hydrograph recession exponent a (Equation 3.4; notice that $a > 1$ in most river basins; Biswal and Kumar, 2014; Tashie et al., 2020b).

$$\lim_{q \rightarrow +\infty} p(q) = \lim_{q \rightarrow +\infty} \left\{ C_1 \cdot \underbrace{\left(q^{-a} \right)}_{\mapsto 0} \underbrace{\left(e^{\frac{-1}{\alpha K(2-a)} \cdot q^{2-a}} \right)}_{\mapsto 0} \underbrace{\left(e^{\frac{\lambda}{K(1-a)} \cdot q^{1-a}} \right)}_{\mapsto e^0=1} \right\}_{\substack{1 < a < 2 \\ a > 2}} \quad (3.4)$$

When $1 < a < 2$, the last term on the right-hand side converges to a constant value of one as q increases, thereby no more influencing how the distribution decreases toward zero. The first two terms instead decrease toward zero, affecting how the probability decreases for increasing values of q . The tail behavior is in this case determined by both a power law and a stretched exponential function, indicating that the probability decreases faster than a stretched exponential but slower than a power law. When $a > 2$, both the stretched exponential terms converge to a constant value of one as q increases, and thus no more influence how the probability decreases toward zero. In this case the tail of the distribution is solely determined by the power law function. Despite being aware that several definitions of heavy-tailed distribution exist (El Adlouni et al., 2008; Vázquez et al., 2006), in the remainder of the manuscript we refer to distributions which exhibit a power law tail as heavy-tailed.

From the above derivations, the hydrograph recession exponent emerges as a key index of the tail behavior of streamflow distributions, which shall be heavy-tailed for values of $a > 2$. I apply the same analyses to infer the tail behavior of the probability distributions of ordinary peak flows and floods by taking the limit of $q \rightarrow +\infty$ for both Equations 3.2 and 3.3. Because $\lim_{q \rightarrow +\infty} D_j(q) = \int_{\infty}^{\infty} p_j(q) dq = 0$, the Equations 3.2 and 3.3 can be transformed into: (set $C_3 = \lambda \tau C_2$)

$$\lim_{q \rightarrow +\infty} p_j(q) = \lim_{q \rightarrow +\infty} \left\{ C_2 \cdot \underbrace{\left(q^{1-a} \right)}_{\mapsto 0} \underbrace{\left(e^{\frac{-1}{\alpha K(2-a)} \cdot q^{2-a}} \right)}_{\mapsto e^0=1} \right\}_{\substack{1 < a < 2 \\ a > 2}} \quad (3.5)$$

$$\lim_{q \rightarrow +\infty} p_M(q) = \lim_{q \rightarrow +\infty} \left\{ C_3 \cdot \underbrace{\left(q^{1-a} \right)}_{\mapsto 0} \underbrace{\left(e^{\frac{-1}{\alpha K(2-a)} \cdot q^{2-a}} \right)}_{\mapsto e^0=1} \right\}_{\substack{1 < a < 2 \\ a > 2}} \quad (3.6)$$

Notably, I observe that the same critical value of the recession exponent equal to 2 separate the absence and presence of heavy-tailed behavior also in these cases. Therefore, I propose the hydrograph recession exponent a as a suitable indicator of heavy-tailed flood behavior, based on the description of hydrological processes embedded in the physically-based extreme value model. I test its capability to correctly predict such behavior in the following Sections of this chapter.

3.2 Dataset description

To test the proposed index of heavy-tailed flood behavior (i.e., the hydrograph recession exponent a), I use daily streamflow records of 98 gauges across Germany (supporting information S2). The analyzed river basins encompass a variety of climate and physiographic settings (Tarasova et al., 2020a). Their areas range from 110 to 23,843 km^2 with a median value of 1,195 km^2 . The length of the streamflow records range from 35 to 63 years with a median value of 58 years (inbetween 1951 – 2013). I perform all analyses on a seasonal basis (winter: December–February, spring: March–May, summer: June–August, fall: September–November) to account for the seasonality of the hydrograph recessions (Tashie et al., 2020b) and flood distributions (Durrans et al., 2003). I term the analysis of a given river gauge during a season a case study. I select gauges for which processes driving streamflow dynamics are reasonably consistent with the adopted theoretical framework. Hence, I discard gauges affected by large dams, reservoirs (Lehner et al., 2011) and anthropogenic flow disturbances (based on visual examination; Tarasova et al., 2018). Case studies with strong snowfall (during a season), for which the average daily temperature is below zero degrees during precipitation events for over 50 % of a season, are also discarded (i.e., only the affected season is removed from the analyses). This results in an overall number of 386 case studies, including 97 case studies in spring, 96 in summer, 98 in autumn and 95 in the winter season.

3.3 Hydrograph recession analysis

The proposed index is derived from hydrograph recession analysis. The hydrograph recession is typically described by a power law relationship between the rate of change of streamflow in time, dq/dt , and the magnitude of streamflow q (Brutsaert and Nieber, 1977). Recent studies have suggested estimating this power law relationship for individual recession events rather than aggregating them, enhancing the representation of observed recession behavior (Biswal and Marani, 2010; Basso et al., 2015; Karlsen et al., 2019; Jachens et al., 2020; Tashie et al., 2020a; Biswal, 2021). In line with these studies, I calculate the recession exponent for each individual event and then take the median exponent across all events as the representative value for a given case study. In particular, a power law is used to represent hydrograph recessions of a single event i , $dq/dt = -K_i \cdot q^{a_i}$, where t denotes the unit time, K_i and a_i denote the estimated coefficient and exponent of hydrograph recessions for event i , respectively. The median value of all the a_i is the estimated value of a considered in this study and here used to represent the average nonlinearity of catchment response. Hydrograph recessions are composed of ordinary peak flows and the following streamflow values decreasing for a minimum duration of five days. The proposed index of heavy-tailed flood behavior can thus be estimated based on commonly available daily discharge observations.

3.4 Benchmarking heavy-tailed case studies: recession exponents vs. empirical power law

To validate the identification of tail behavior obtained by means of the proposed index, I benchmark it against data by fitting a power law distribution to the empirical data distribution. A case study is considered to be heavy-tailed according to the observations if the fitted power law reliably describes the tail behavior of the data distribution. This is evaluated by means of a state-of-the-art method proposed by Clauset et al. (2009). The exponent b of the empirical power law is first computed by fitting a power law to the upper tail of the data distribution. An optimized lower boundary is determined by considering the best fit according to the Kolmogorov-Smirnov (KS) statistic, one of the most common measures of the distance between two non-normal distributions. The method then assesses whether the fitted power law reliably represents the observed data by using statistical tests, such as the Kolmogorov-Smirnov statistic and a Monte Carlo procedure, to verify that the residual errors between the data and the power law distribution fall within the range of fluctuations expected from random sampling. If the residual errors are found to be within the range of fluctuations expected from random sampling, the power law is deemed a reliable representation of the empirical data distribution (supporting information S1). I use the python package `plfit` 1.0.3 to implement these computations and refer to Clauset et al. (2009) for further details concerning the approach.

I analyze three types of empirical data, namely daily streamflow, ordinary peaks, and monthly maxima, and obtain estimates of the fitted exponent b for each case. We use these results to validate the capabilities of the proposed index to infer heavy-tailed flood behavior from common hydrological dynamics, i.e., from the analysis of hydrograph recessions. I acknowledge that the benchmark I use, i.e., the empirical power law, may be influenced by fitting uncertainty due to data scarcity in some cases (i.e., especially when we analyse maxima; I indeed considered monthly maxima (Fischer and Schumann, 2016; Malamud and Turcotte, 2006) instead of the seasonal maxima previously used in the literature (e.g., Basso et al., 2021) to extend the sample size). The parallel analyses for cases with larger sample size (i.e., daily streamflow and ordinary peaks) provide more robust validation and support the interpretation of results for maxima.

3.5 Chapter results and discussion

3.5.1 Identification of heavy-tailed case studies

I examine if power law distributions fitted to the empirical distributions of daily streamflow, ordinary peaks, and monthly maxima well describe the observed data for case studies identified as having heavy-tailed behavior (i.e., $a > 2$) according to the proposed index. First, I identify the case studies with either heavy- ($a > 2$) or nonheavy ($a < 2$) -tailed behavior based on the proposed index. Then, I utilize the KS statistic κ to measure the distance between the frequency distributions of observations and a power law distribution (specifically, on the tail of the distribution). This assessment gauges the effectiveness of the fitted power law distribution in characterizing the dataset (with $\kappa \in [0, \infty]$, where $\kappa = 0$ represents the utmost reliability).

The KS test is a common nonparametric method suitable for non-normal distributions. Low values of the KS statistic κ indicate that the empirical data are likely to be drawn from a power law. Figures. 1a-1c show that the histograms of the number of case studies for decreasing values of the KS statistic are significantly skewed (i.e., the skewness is significantly different from zero) toward lower values of κ for all cases of daily streamflows, ordinary peak flows, and monthly flow maxima with $a > 2$ (red histograms), whereas this is not true for cases with $a < 2$ (green histograms) (i.e., the skewness is not significantly different from zero in these cases). Statistical significance of the skewness was evaluated through the Jarque–Bera test at a significance level of 0.05. The result indicates that data from case studies which are identified with heavy-tailed behavior according to the proposed index ($a > 2$, red) are indeed more likely to come from power law distributions.

I further estimate the accuracy of the proposed index based on the fraction of case studies that are identified as heavy-tailed by the proposed index among all cases that are heavy-tailed according to the available observations. To define the latter, I set a threshold value of κ : the power law is a reliable representation of the data for cases with κ below the threshold. Mathematically, the accuracy can be expressed as $P(a > 2 \mid \kappa < \kappa_r) = N_c(a > 2 \mid \kappa < \kappa_r) / N_c(\kappa < \kappa_r)$, where κ_r is the imposed threshold of κ , $N_c(\kappa < \kappa_r)$ is the number of case studies whose $\kappa < \kappa_r$, and $N_c(a > 2 \mid \kappa < \kappa_r)$ is the number of case studies with $a > 2$ among the $N_c(\kappa < \kappa_r)$ case studies. Higher accuracy essentially means that a higher fraction of heavy-tailed cases is correctly identified by means of the proposed index. To achieve this, I systematically reduce the threshold of KS statistic κ_r (imposing a more stringent criterion for incorporating cases in the computation of conditional probability of accuracy) along the x-axis in Figure 1, progressing from left to right. It's important to note that as the κ_r threshold becomes smaller, the reliability of describing the data using power law distributions increases (as denoted by the second axis legend of Figure 1).

Figures 1d-1f display the accuracy of the proposed index as a function of the reliability threshold κ_r . In all three cases (daily streamflows, ordinary peak flows, and monthly flow maxima), the accuracy values increase with the reliability level of the power law distribution fitted on observed data. This means that the proposed index shows high accuracy for case studies where the empirical distributions of observed data are more consistent with power laws. In other words the proposed index, which is estimated from common streamflow dynamics as the hydrograph recession exponent, accurately identifies heavy-tailed behavior of streamflow and flood distributions displayed by the available observations.

3.5.2 Evaluation of tail heaviness

I further employ the goodness-of-fit testing procedure proposed by Clauset et al. (2009) (supporting information S1) to identify case studies for which the representation of daily streamflow, ordinary peak flows, and monthly maxima by means of power law distributions is convincingly supported by the available data. We refer to these case studies as 'confirmed heavy-tailed cases' (Figure 2, black dots). Conversely, I term the remaining ones as 'uncertain cases' (Figure 2, gray). The latter label denotes the fact that it cannot be determined with certainty whether the distributions underlying the available observations in these cases are or not power laws due to scarcity of data.

3. Insights from Common Streamflow Dynamics

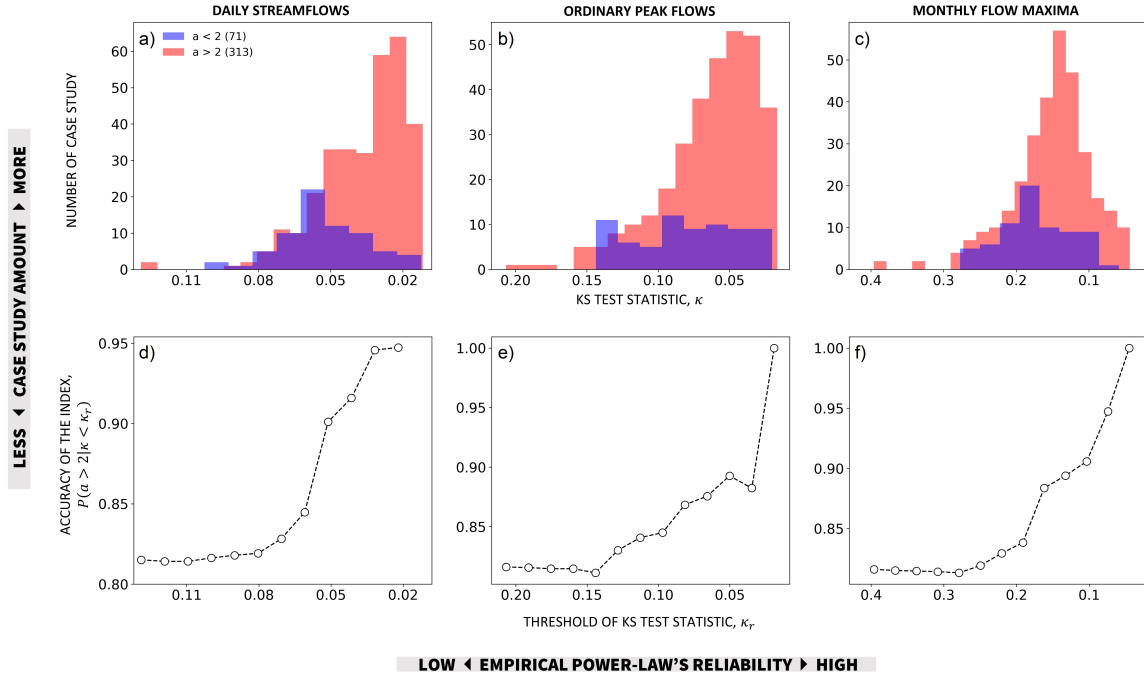


Figure 3.1: Accuracy of the proposed index (a)-(c) Number of analyzed case studies as a function of the KS statistic κ of empirically fitted power law distributions (the latter is a measure of how reliable the power law is as a model for the given data: the lower κ , the more reliable the power law model). Case studies are identified as having either heavy- ($a > 2$, red histograms) or nonheavy ($a < 2$, green histograms) –tailed behavior based on the hydrograph recession exponent a estimated from daily flow records, which is proposed as an index of heavy-tailed streamflow and flood behavior. (d)-(f) Accuracy of the proposed index as a function of decreasing thresholds of κ_r (i.e., increasing reliability of empirical power laws). The values of the KS statistic κ are derived from records of (a, d) daily streamflows, (b, e) ordinary peak flows, and (c, f) monthly flow maxima.

Figure 2 shows the empirical power law exponent b as a function of the proposed index of heavy-tailed behavior a . Red markers display the median values of a and b (squares), the interquartile intervals of b (vertical bars), and the binning ranges of a (horizontal bars, equal number of case studies in each bin), highlighting the correlation between the empirical power law exponent b and the hydrograph recession exponent a for confirmed heavy-tailed cases (black dots) in all three cases (i.e., daily streamflows, ordinary peak flows, and monthly flow maxima). I test the correlation by calculating their distance (Székely et al., 2007) and Spearman (Spearman, 1904) correlations, which are valid for both linear and nonlinear associations between random variables. I find that a and b are significantly correlated at a significance level of 0.05 in all three cases with distance (Spearman) correlation coefficients of 0.45, 0.44, and 0.81 (0.42, 0.46, and 0.60) for daily streamflows, ordinary peak flows, and monthly flow maxima. The high values of the correlation coefficients for monthly flow maxima are likely affected by the existence of two clusters in Figure 2c. Nonetheless, the existence of a statistically significant correlation between the empirical power law exponent and the proposed index, obtained for panels a, b, and c, confirms that the proposed index not only can be used to identify heavy-tailed flood behavior (as Figure 1 shows) but also to evaluate the degree of the tail heaviness of the underlying distributions.

Figure 2c is of particular interest because it shows an example of the typical limitations of methods that rely solely on observations to determine the tail behavior of the distribution of maxima (e.g., Papalexiou and Koutsoyiannis, 2013) and, at the same time, highlights the power of the proposed index. Large values of the recession exponent a , in agreement with corresponding large values of b , are found for all confirmed heavy-tailed cases (black dots in Figure 2c) where the power law provides a plausible representation of the empirical distribution of monthly maxima. For uncertain cases (gray dots in Figure 2c) the values of the empirical power law exponents are unreliable (according to the applied method; Clauset et al., 2009) since it cannot be determined with certainty whether the empirical distributions are or not power laws due to data scarcity. Conversely, the hydrograph recession exponent is calculated from daily streamflow data. I can therefore identify cases with heavy-tailed behavior and evaluate their tail heaviness based on the values of a . This estimate is deemed robust, provided that the predictions of the proposed index are confirmed by observations in cases (panels a and b) where data size is not a limitation (i.e., for daily streamflow and ordinary peak flows).

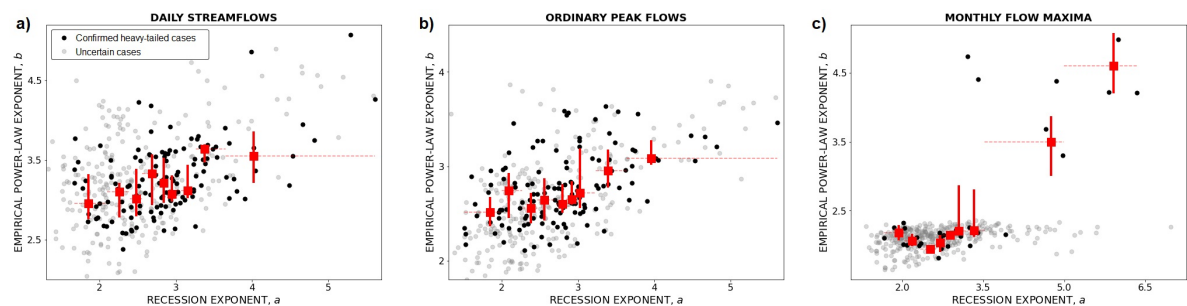


Figure 3.2: Empirical power law exponent b as a function of the proposed index of heavy-tailed behavior a Case studies are classified into groups of confirmed heavy-tailed (black dots) and uncertain (gray dots) cases on the basis of the hypothesis test (supporting information S1; Clauset et al., 2009). The former denotes cases for which a power law provides a reliable description of the empirical data distribution, while the latter denotes cases whose data cannot convincingly support such a distribution. Red markers highlight the correlation between the empirical power law exponent b and the hydrograph recession exponent a for confirmed heavy-tailed cases in the case of (a) daily streamflows ($n = 121$ case studies), (b) ordinary peak flows ($n = 116$), and (c) monthly flow maxima ($n = 34$). Red markers display the median values of a and b (squares), the interquartile intervals of b (vertical bars), and the binning ranges of a (horizontal bars, equal number of case studies in each bin).

3.5.3 Effectiveness in short-length record datasets

In Figure 3 I test the stability of the categorization of case studies into heavy/nonheavy-tailed flood behavior provided by the proposed index (i.e., the hydrograph recession exponent a) for decreasing data lengths. We compare results for the proposed index against two other frequently used metrics of heavy tails in hydrological studies: (1) the upper tail ratio (UTR) (Lu et al., 2017; Smith et al., 2018; Villarini et al., 2011) and (2) the shape parameter ξ of the GEV distribution (Morrison and Smith, 2002; Papalexiou et al., 2013; Villarini and Smith, 2010). The UTR is defined as the ratio of the flood of record to the 0.9 quantile of floods (Smith et al., 2018) here represented by monthly flow maxima, while ξ is estimated by fitting a GEV distribution on the sample of monthly maxima using the python package OpenTURNS 1.16 (Baudin et al., 2017). For all three indices (a , UTR, and ξ), we estimate their values

3. Insights from Common Streamflow Dynamics

for data lengths decreasing from 35 (i.e., the shortest entire record length in the dataset) to 2 years. I acknowledge that estimating parameters of extreme value distributions from such short records is not recommended. However, the exercise highlights the perks of the proposed index that, as it will be shown, is able to provide robust results also when short data series only are available. For each case study, I obtain 30 samples with the assigned test length from the entire data series using resampling without substitution. For each test length, I calculate the median values of the indices estimated from these samples, and plot them in Figure 3 together with their variability across case studies (vertical shaded bars and lines in Figure 3 show the 0.25-0.75 and 0.05-0.95 quantile ranges of the index estimates across case studies).

To evaluate the consistency of the categorization of tail behavior across different data lengths we proceed as follows. For each case we first compute the hydrograph recession exponent and GEV shape parameter from the entire data record and denote them with an asterisk superscript (i.e., a^* or ξ^*). Heavy-tailed cases are defined as having $a^* > 2$ or $\xi^* > 0$ (Godrèche et al., 2015), while non-heavy-tailed cases have values below these thresholds. To visualize heavy-tailed and non-heavy-tailed behaviors, I mark them in Figure 3 in red and green colors, respectively, based on the reference values obtained from the entire data record. I then recalculate the indices from shorter samples and evaluate whether their values are consistent with the above categorization. For the UTR, we cannot implement this approach because there is no specific threshold for the identification of heavy/nonheavy tails. I therefore directly compare the stability of the UTR's values across data lengths (a larger value indicates a heavier tail).

The proposed index provides consistent categorization of heavy/nonheavy-tailed flood behavior across varying data lengths (Figure 3a). The index estimates remain above 2 for most heavy-tailed cases (red) and below 2 for most nonheavy-tailed cases (green) (as defined according to the reference value a^* computed using the entire data record) when the data length decreases. The index estimates demonstrate the consistency throughout the test data length, as evidenced by the narrow range of variation in the median values of the estimates. For heavy-tailed cases, the median values ranged from 2.64 to 2.92, while for nonheavy-tailed cases, they ranged from 1.84 to 2.0. Additionally, the coefficient of variation for the estimates remained relatively constant, ranging from 0.29 to 0.33 for both heavy and nonheavy-tailed cases. This indicates that the variability of the results (vertical shaded bars and lines in Figure 3) is mostly due to pooling together different case studies belonging to the same category (heavy or nonheavy-tailed), and does not increase as a result of decreasing length of the available data.

In contrast, the upper tail ratio shows pronounced instability for decreasing data lengths (Figure 3b). The median value of the index estimates varies between 1.32 and 2.36, with a coefficient of variation ranging from 0.15 to 0.64. These values indicate uncertain assessments based on the UTR and its tendency to underestimate the tail heaviness as the data length decreases.

Figure 3c illustrates the categorization of tail behavior using GEV shape parameter estimates. The results indicate that ξ estimates are stable with longer data series, yet their variability increases — leading to both underestimation and overestimation of tail heaviness — when data length is short. To ensure a stable categorization of flood tail behavior using this

method data series spanning more than 10 years (for seasonal analyses and monthly maxima, i.e., sample sizes of around 30 values) are needed, in line with the findings of previous studies (Cai and Hames, 2010; Németh et al., 2019). The median values of ξ range from 0.39 to 0.52 for heavy-tailed cases and remain at 0 for nonheavy-tailed cases. Furthermore, the coefficient of variation demonstrates relatively higher variation across different test data lengths, ranging from 0.37 to 1.03 for heavy-tailed cases.

Figure 3d presents a summary of the consistency in identifying tail behavior (either heavy or nonheavy) compared to the identification based on the complete data record (i.e., fraction of cases for which categorization based on shorter data series provides the same result obtained with the complete data record). This assessment is conducted for both the methods of recession exponents and GEV shape parameters (unfortunately, this approach is inapplicable to the UTR due to the absence of a specific threshold for distinguishing heavy/nonheavy tails). The comparison underscores that discrepancies in consistency between the two indices (ξ and a) are predominantly noticeable when analyzing data series shorter than 10 years in this study. Conversely, for data series longer than 10 years, both indices exhibit comparable consistency and display an ascending trend, with the performance of the GEV shape parameters slightly higher than the one of the recession exponents.

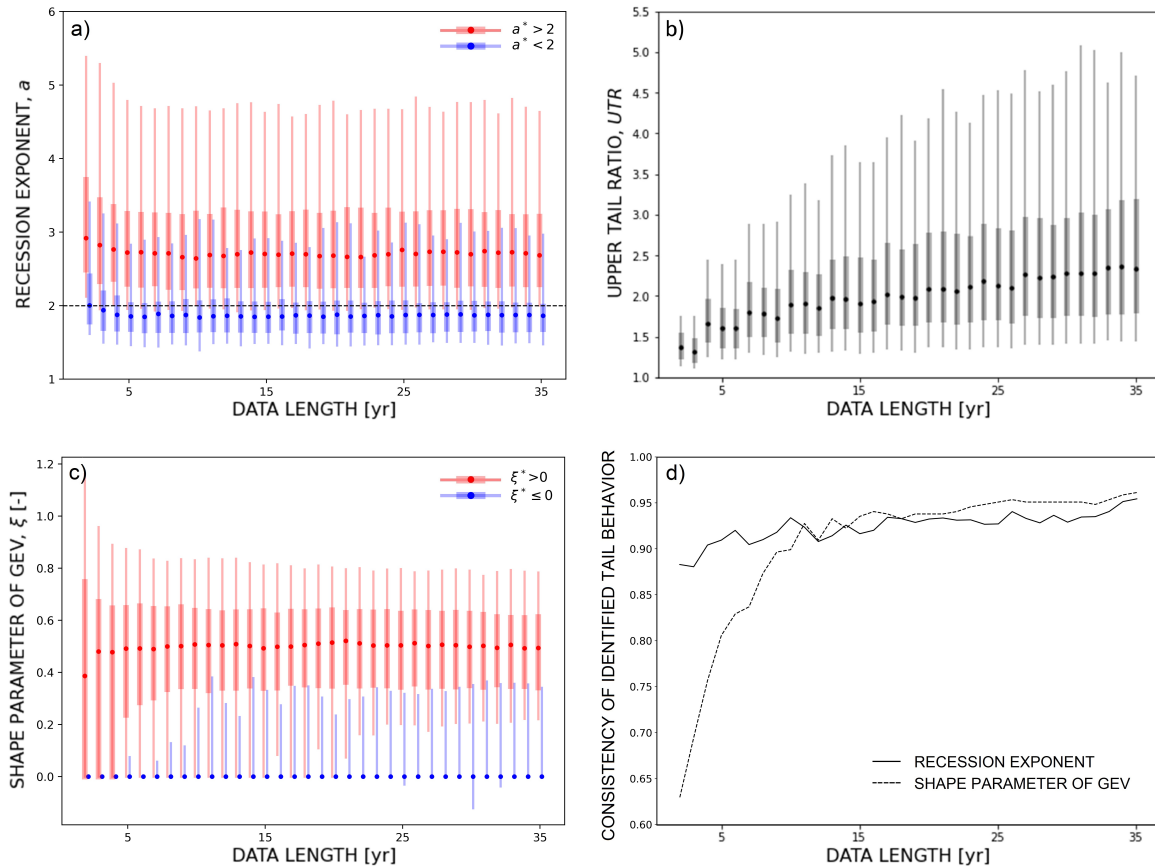


Figure 3.3: Stability of the categorization of case studies into heavy/nonheavy-tailed flood behavior for decreasing data lengths Estimates of three different indices of tail behavior as a function of data length. (a) Hydrograph recession exponent a (i.e., the proposed index of this study). Two frequently used metrics of heavy tails in hydrological studies: (b) the upper tail ratio UTR, and (c) the shape parameter ξ of the GEV distribution. Dots display the median values of the estimates for 386 case studies; vertical shaded bars and lines show the 0.25-0.75 and 0.05-0.95 quantile ranges of the estimates, respectively. The entire data record was used for computing the reference values of the hydrograph recession exponent a^* and the GEV shape parameter ξ^* and categorizing each case study as either having (red) or not (green) heavy-tailed behavior. (d) Consistency of identified tail behavior (either heavy or nonheavy) as a function of available data length for the indices recession exponent and shape parameter of GEV.

3.5.4 Discussion

Assessment of flood tail behavior is challenging due to high levels of uncertainty arising from the scarcity of data on floods, which are by definition rare events. This issue is particularly prominent when maxima are used in the analysis as in the annual maximum approach. Despite the widespread use of this method, its limitations for what concerns the reliability of flood tail estimates are well recognized. Very large sample sizes are indeed essential for obtaining accurate prediction of tail behavior (Papalexiou and Koutsoyiannis, 2013).

To address the challenge of obtaining reliable estimates, alternative methods have been proposed. A frequently used approach is the peak-over-threshold analysis, which uses the information content of a larger sample of data (Lang et al., 1999; Pan et al., 2022). Previous studies have demonstrated that this method leads to lower uncertainty in estimating high floods (Kumar et al., 2020). Volpi et al. (2019) also showed the advantage of using all the available observations (i.e., not only the peaks over a certain threshold) for estimating extreme events. In summary, all these methods suggest that discharge values other than maxima can provide information about the characteristics of extreme events. Specifically, incorporating information from less extreme (but more numerous) observations can reduce the uncertainty in the estimation of extreme events and lead to improved accuracy. Furthermore, non-asymptotic methods suggest that extremes are realizations of the underlying ordinary events (Marani and Ignaccolo, 2015; Lombardo et al., 2019), which can thus be used to assess rare events. These methods have significantly improved the estimation of extreme values with lower uncertainty (Marra et al., 2018; Miniussi and Marani, 2020; Mushtaq et al., 2022; Hu et al., 2023).

Similarly to the latter approaches, the index introduced in this study (i.e., the hydrograph recession exponent) leverages information on ordinary discharge dynamics to infer the tail behavior of flood distributions. This approach entails some advantages: firstly, it effectively extracts information from a larger amount of available streamflow data. Secondly, estimating the hydrograph recession exponent requires significantly less data than conventional approaches that involve fitting probability distributions to hydrological samples, while providing reliable results. But most importantly, the proposed index offers a mechanistic approach to understand the emergence of heavy-tailed flood behavior, thus providing a process-based alternative to methods that solely rely on statistical analysis of observations. The importance of understanding intrinsic watershed dynamics which promote the occurrence of extreme events and contributing factors that lead to heavy-tailed flood behavior (Tarasova et al., 2020a) was recently highlighted in a comprehensive review by Merz et al. (2022). Identify reliable proxies for inferring such behavior (sensu Wilson and Toumi, 2005) is as well important. The proposed index, which represents such a proxy grounded on intrinsic hydrologic dynamics of the river basin, is thus especially useful in the very common cases when the tail of the flood distribution cannot be known from limited available observations.

The hydrograph recession exponent (which is the identified index of heavy-tailed flood behavior) essentially represents the nonlinearity of the storage-discharge response in catchments (Wittenberg, 1999; Biswal and Marani, 2010). A higher degree of nonlinearity leads to higher peak flows and heavier tail of the streamflow distribution (Basso et al., 2015). In agreement with these findings, former simulation-based and field studies have shown that high nonlinearity of the catchment hydrological response linked to an increase of the runoff contributing area

3. Insights from Common Streamflow Dynamics

results in a marked increase of the slope of flood frequency curves (Fiorentino et al., 2007; Rogger, Pirkl et al., 2012), which may be indicative of a heavy-tailed flood behavior. Gioia et al. (2012) also demonstrated that a nonlinear catchment response can convert light-tailed rainfall inputs into flood distributions with heavy tails, further confirming the role of nonlinear storage-discharge responses in producing heavy-tailed flood behavior.

Merz et al. (2022) established, based on a comprehensive review, that the nonlinearity of the catchment response is a plausible contributor to the emergence of heavy-tailed flood behavior. Additionally, Basso et al. (2023) demonstrated that the hydrograph recession exponent aids in predicting the propensity of rivers for generating extreme floods. In line with these studies, our research further highlights that the hydrograph recession exponent, which provides a description of catchment nonlinearity obtained from common streamflow dynamics, is capable of robustly identifying heavy-tailed flood behavior.

The findings in Figure 2 showcase the drawbacks of relying on purely statistical data analyses (which supply the empirical power law exponents b) to identify flood tail behaviors and the advantages of adopting the mechanistic approach proposed in this study (which yield the hydrograph recession exponent a). The gray markers in Figure 2 indicate uncertainty in determining whether the distribution has a power law tail, which is shown to be more prevalent when the sample size is reduced, based on statistical analyses according to the Clauset (2009) method (69 %, 70 %, and 91 % of the case studies for daily streamflow, ordinary peaks, and monthly maxima, respectively). The proposed index finds a solution to these limitations through a mathematical description of hydrological processes. Such an index is shown to perform well in cases where statistical methods may be limited due to a lack of data, as confirmed by the significant correlations between the recession exponent and the reliably empirical power law exponent in all three panels (represented by black dots in Figure 2). Even in cases where the statistical method is unable to confirm the underlying distribution (e.g., monthly maxima in panel c), the proposed index can still provide robust estimates of tail heaviness based on the values of recession exponents. This is supported by the analyses of daily streamflows and ordinary peaks, where sample size is not a limitation and the predictions of the proposed index are confirmed by observations. Overall, the proposed index offers a promising solution for accurately characterizing the tail behavior of flood distributions, especially when traditional statistical methods may be limited due to a lack of data.

Data scarcity is a major challenge for reliable flood hazard assessment, mainly because of relatively short hydrological data records worldwide (Lins, 2008). The availability of a robust index of heavy-tailed flood behavior that work even with short data records is desirable. I test three indices, namely the recession exponent (the proposed index), the upper tail ratio (UTR), and the shape parameter of the Generalized Extreme Value (GEV) distribution (ξ), for categorizing tail behavior for decreasing data lengths. The results (Figure 3a) show that the recession exponent provides stable estimates and categorizes cases consistently into heavy or non-heavy tails for decreasing data lengths. Furthermore, the slight variation in the estimates of the recession exponent for each test data length implies that variation in estimates primarily arises from case study heterogeneity rather than decreasing data length. Conversely, UTR significantly underestimates both the tail heaviness and the variation across cases for decreasing data lengths (Figure 3b). In agreement with previous studies, underestimation of tail heaviness

occurred using UTR when the sample size was small (Smith et al., 2018; Wietzke et al., 2020). Meanwhile, the categorization of tail behavior was stable for cases with dataset longer than 10 years using the GEV shape parameter. However, high uncertainty in the variation of estimates across cases is observed when available data is relatively short as also highlighted by previous studies (e.g., Wietzke et al., 2020) (Figure 3c). Implied by this observation is that the estimates are biased by the short analyzed data and a longer data record is desirable for a more reliable fitting of a GEV on data (Papalexiou and Koutsoyiannis, 2013). In summary, both the recession exponent and the GEV shape parameter exhibit greater stability across data lengths than the UTR, which is highly dependent on the available amount of data. When comparing the first two indices (recession exponent and GEV shape parameter) (Figure 3d), the recession exponent demonstrates a high level of stability across all data lengths, even those shorter than 10 years based on this study's analyses. On the other hand, the GEV shape parameter displays lower stability when the available data are shorter than 10 years, but this stability significantly improves as the data length exceeds 10 years. Beyond the 10-year threshold, both indices show comparable consistency and an upward trend, with GEV shape parameters slightly outperforming recession exponents.

The hydrograph recession exponent allows at least two significant applications as a proxy for heavy-tailed flood behavior. Firstly, it can be directly used to improve comparability across catchments and provide a fair assessment of mapping regional patterns of flood hazards (Merz et al., 2022). Traditionally, assessing flood behavior across catchments using the same record length has been preferred (Cunderlik and Burn, 2002), but this is often not possible due to differences in data availability. The proposed index can robustly estimate heavy-tailed flood behavior from data with different record lengths, overcoming this limitation. Secondly, it can be applied as a preliminary step to correctly identify whether a considered catchment exhibits heavy-tailed flood behavior or not, and to select an appropriate probability distribution to be used in flood frequency analysis. This prior identification of tail behavior is crucial to avoid potential underestimation of flood extremes. (Miniussi et al., 2020; Mushtaq et al., 2022).

3.6 Chapter conclusions

A new index of heavy-tailed flood behavior is identified from a physically-based description of streamflow dynamics. The new index is embodied by the hydrograph recession exponent and can be readily estimated from daily streamflow records. Our findings demonstrate that this index enables the identification of heavy or nonheavy tailed flood behaviors in a large set of case studies across Germany. Importantly, it provides an evaluation of the tail heaviness (i.e., the severity of flood risks) based on analyses of common discharge dynamics, and remarkably, the results remain robust even with limited data records.

The proposed index addresses the main limitations of current approaches, including the lack of physical support and low reliability in cases with limited data records. By extracting more information from available data and manifesting the nonlinearity of catchment response, it represents a reliable method to select suitable underlying distributions for flood frequency analyses and assess the peril of extreme floods in data poor areas.

3.7 Supporting information

S1. Goodness-of-fit tests for the empirical power laws

To test if the empirical power law is a plausible underlying distribution of the observed data, we follow the hypothesis test proposed by Clauset et al. (2009). The null hypothesis is ‘The empirical power law is a plausibly underlying distribution of the observed data.’ Residual errors exist between the empirical power law and the observed data, which can be estimated by the error distance ε_d by means of the Kolmogorov-Smirnov statistic. The Kolmogorov-Smirnov test is selected because it is one of the most common measures for non-normal data. The core of the hypothesis test is to statistically prove that the errors between the data and the power law (i.e., ε_d) are rational fluctuation of sampling randomness rather than being drawn from an incorrect underlying distribution. To determine the rationality of the sampling randomness, a Monte Carlo procedure is introduced: (1) a large number of groups n of synthetic data (with the same size as the observed data) are randomly generated from the empirical power law; (2) the error distance ε_{s_i} of each synthetic group to the empirical power law is calculated for $i = 1, 2, \dots, n$; (3) the frequency of $\varepsilon_s > \varepsilon_d$ defines the p-value of the hypothesis test, which indicates the probability that the residual errors between the empirical power law and the observed data locates within the range of sampling randomness fluctuations; and (4) the rationality is determined by $p > 0.1$ using this package.

When $p \leq 0.1$, the null hypothesis is rejected; that is, the observed data are not plausibly drawn from the empirical power law. On the contrary, the empirical power law is considered a plausible distribution for the observed data for their residual errors are statistically rational fluctuation of sampling randomness when $p > 0.1$. Notice that a greater p-value is better in this case because the aim is to verify the null hypothesis rather than to indicate it is unlikely to be correct, as others often considered. Thus $p > 0.1$ is a more rigorous setting than $p > 0.05$ in this case.

The setting of $n = 1000$ is used as an adequate (great enough) number of iterations in this framework to distinguish underlying distributions that are commonly mixed (as suggested by Clauset et al. (2009)).

The hypothesis test of the empirical power law including all the above procedures can be implemented via the function `test_pl` in the python package `plfit 1.0.3` (<https://pypi.org/project/plfit/>).

It is worth mentioning that, statistically, we cannot say those who do not pass the hypothesis test ‘are not’ power law distributions. There are at least two potential reasons for this result: (1) they are indeed not power law functions, or (2) the underlying distribution cannot be concluded due to the high uncertainty in the empirical data with small sample sizes. We thus use the term ‘uncertain cases’ to indicate this awareness in the main manuscript.

S2. A reference map of gauges across Germany used in chapter 3

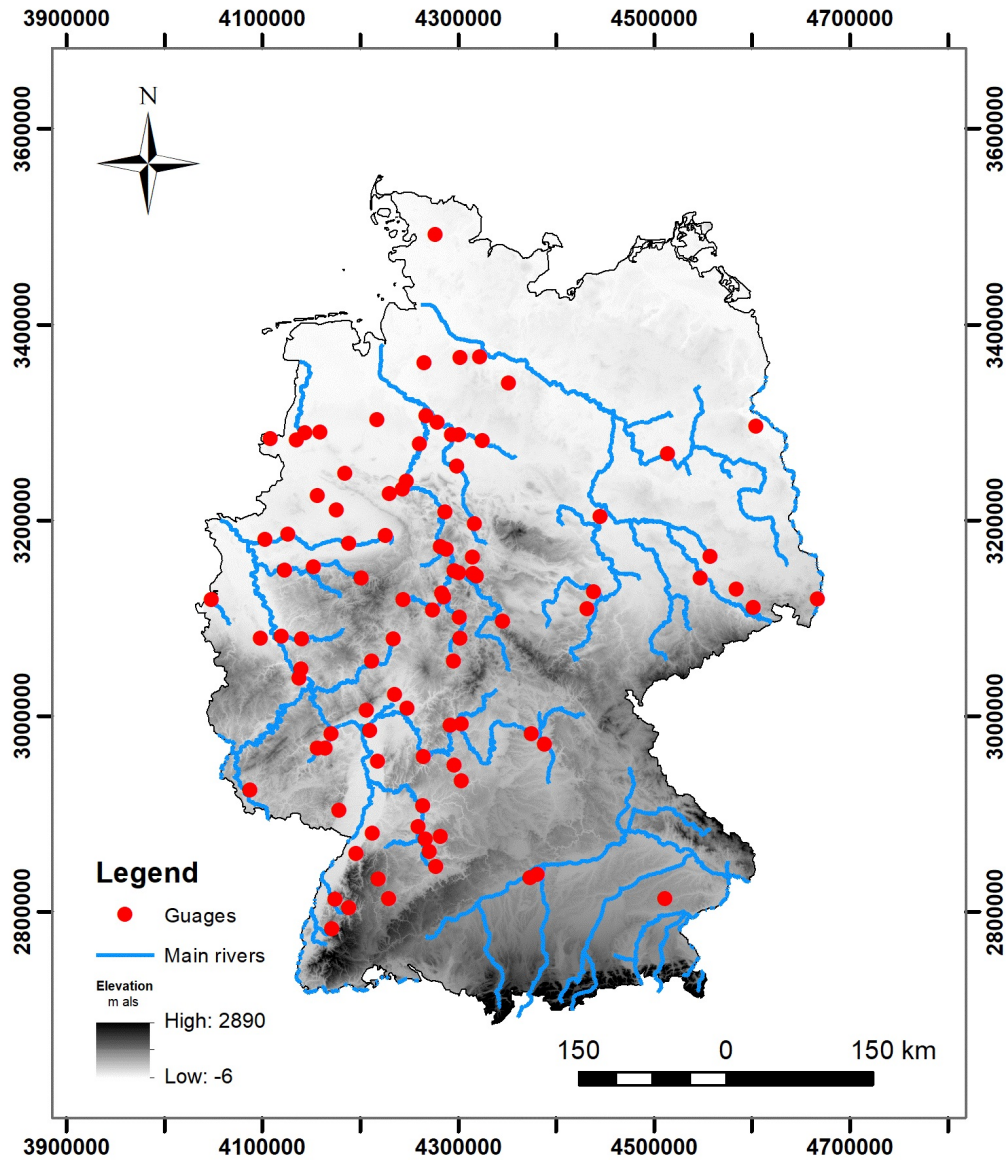


Figure 3.4: A reference map of gauges across Germany used in chapter 3. These river basins encompass a variety of climate and physiographic settings without strong impact from dams and snowfall. Their areas range from 110 to 23,843 km^2 with a median value of 1,195 km^2 . The minimum, median, and maximum lengths of the daily streamflow records are 35, 58, and 63 years (in between 1951 - 2013).

3.8 Data availability

For providing the discharge data for Germany, we are grateful to the Bavarian State Office of Environment (LfU, <https://www.gkd.bayern.de/de/fluesse/abfluss>) and the Global Runoff Data Centre (GRDC) prepared by the Federal Institute for Hydrology (BfG, <http://www.bafg.de/GRDC>). Climatic data can be obtained from the German Weather Service (DWD; <ftp://ftp-cdc.dwd.de/pub/CDC/>). The digital elevation model can be retrieved from Shuttle Radar Topography Mission (SRTM; <https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/>).

4

Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

Utilizing the newly proposed index introduced in the previous chapter, an extensive and reliable exploration of heavy-tailed flood behavior is feasible across a diverse range of geographic conditions. The objective of this chapter is to advance our understanding of heavy-tailed behavior by examining its influences under various natural conditions and unraveling the underlying mechanisms.

To organize current knowledge on the drivers and underlying mechanisms of heavy-tailed flood distributions, Merz et al. (2022) conducted an extensive review of current studies and summarized their findings into nine hypotheses. Notably, they pointed out that while one might intuitively assume that heavy-tailed flood distributions are inherited from heavy-tailed rainfall distributions, the evidence does not always support this hypothesis. For instance, a study by McCuen and Smith (2008) revealed that cases with skewed rainfall distributions, implying longer and heavier tails, do not necessarily translate into skewed flood distributions. This finding is supported by similar results from Sharma et al. (2018), who discovered that although there has been a significant increase in rainfall extremes, a corresponding increase in flood extremes is not observed. Indeed, Gaume (2006) pointed out that the asymptotic behavior of flood distributions is primarily controlled by rainfall distributions only for situations with very large return periods.

In the review of Merz et al. (2022), it becomes evident that multiple hydro-physiographic characteristics interact within a complex system, collectively shaping flood tail behavior. Specifically, the interplay between characteristic flood generation (Bernardara et al., 2008; Thorarinsdottir et al., 2018), the presence of mixed flood types (Morrison and Smith, 2002; Villarini and Smith, 2010), the tail heaviness of rainfall distributions (Gaume, 2006), catchment aridity (Molnar et al., 2006; Merz and Blöschl, 2009; Guo et al., 2014), and catchment area (Pallard et al., 2009; Villarini and Smith, 2010) are proposed as contributing factors to the

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

nonlinearity of catchment responses. This nonlinearity is increasingly recognized as a plausible driver of heavy-tailed flood behavior (Fiorentino et al., 2007; Struthers and Sivapalan, 2007; Gioia et al., 2008; Rogger, Pirkl et al., 2012; Basso et al., 2015; Merz et al., 2022; Basso et al., 2023; Wang et al., 2023b).

The nonlinearity of catchment hydrological responses manifests in the hydrograph recession behavior, commonly described by a power law function (Brutsaert and Nieber, 1977; Biswal and Marani, 2010; Tashie et al., 2020b):

$$\frac{dq}{dt} = -B \cdot q^a \quad (4.1)$$

Here, q represents streamflow, t denotes time, and B and a are empirical constants referred to as the recession coefficient and exponent, respectively. Particularly, the recession exponent a is used to express linear to nonlinear responses. Higher a values indicate streamflow behavior with quicker rise for a peak and faster decay during high flow, while slower decay and more stability during low flow (Tashie et al., 2019). Given that a higher recession exponent reflects significant nonlinearity in catchment responses, it has been proposed as an indicator of the emergence of heavy-tailed flood distributions (Basso et al., 2015; Wang et al., 2023b).

In chapter 3, I introduced hydrograph recession exponents as a newly proposed indicator for heavy-tailed flood behavior. This indicator allows for an inference of heavy-tailed flood distributions based on physical mechanisms (i.e., typical hydrological processes within common streamflow dynamics). Importantly, it has shown its capacity to provide robust estimates for both short and long data records. Unlike traditional methods for identifying heavy tails, such as the upper tail ratio (Villarini et al., 2011; Lu et al., 2017) and the shape parameter of a fitted Generalized Extreme Value (GEV) distribution (Morrison and Smith, 2002; Villarini and Smith, 2010), which are sensitive to sample sizes (Wietzke et al., 2020), recession exponents offer more consistent estimates of flood tail heaviness across various data lengths. This characteristic makes it a valuable tool for analyzing regions with diverse gauge data records.

The aim of this following work is to construct a geography of flood tail behavior based on the inferred heavy-tailed flood ‘hotspots’, recognized by this indicator, thus ensuring comparability of analyses across different data lengths. Given that longer and comparable record lengths are desirable for analyzing heavy-tailed distributions using conventional methods (Cunderlik and Burn, 2002; Papalexiou and Koutsoyiannis, 2013), and considering the global variation in available hydrological data lengths (Lins, 2008), this work contributes to filling the research gap by providing a reliable estimation of heavy-tailed flood behavior across a wide range of geography (Merz et al., 2022). Specifically, the objectives are twofold: (1) to validate the effectiveness of recession exponents in identifying heavy-tailed flood behavior through an extensive analysis, and (2) to investigate the underlying factors related to diverse physiographical settings, taking into account spatial patterns, seasonality, and catchment scale characteristics, and how they influence catchment nonlinearity, leading to the emergence of heavy-tailed floods.

I organize the structure of this chapter as follows: Section 4.1 describes the study areas and the hydrological data based on an extensive dataset composed of four countries, Sections 4.2 to 4.4 describe the methods of estimation and validation of hydrograph recession exponents in identifying heavy-tailed flood behavior in the dataset, the framework of the analyses of

spatial patterns of inferred heavy-tailed flood behavior, the framework of the analyses of seasonal dynamics of inferred heavy-tailed flood behavior, and statistical tests. In Section 4.5, I present the validation results of our heavy-tailed flood behavior index, along with analyses of the relationships between flood tail behavior and geographical spatial characteristics, seasonal patterns, and catchment scales in these comparable countries. Particularly, physical interpretations of the results and remarks from the literature are discussed in Section 4.5.5. The main conclusions are summarized in Section 4.6.

4.1 Study areas and data

This study uses four distinct datasets, each serving a specific objective. I begin our analysis by examining case studies from Germany, providing an initial investigation into our research aims. The inclusion of case studies from the United Kingdom (UK) and Norway allows us to explore the applicability of the index in regions with both similar and contrasting physiographic settings compared to Germany. Lastly, the comprehensive dataset of case studies from the United States (US), known for its diverse range of physiographical settings, enables us to validate and consolidate the transferability of our findings.

Germany has a temperate oceanic climate, with mild temperatures and relatively evenly distributed precipitation throughout the year. The western parts of the country are influenced by the North Atlantic Drift, resulting in milder winters compared to the eastern parts, and the Alps play a significant role in the local climate of the south. The country's elevation ranges from sea level to 2,962 *m*, with lowlands in the north, uplands in the central, and mountain ranges in the south. The dominant soil types in Germany are podzols and brown earths, which can result in higher runoff and flash floods in areas with steep slopes and sparse vegetation.

The UK also has a temperate oceanic climate, with diverse topography ranging from upland areas in the north and west to lowland areas in the south and east. The elevation ranges from sea level to 1,345 *m*, with mostly lowland terrain dominated by limestone, shale, and sandstone. The dominant soil types are clay soils, which can result in higher runoff and flood risk in urban areas and other places with limited vegetation cover.

Norway serves as a contrasting climate and terrain compared to Germany. It has a subarctic climate, and the country's terrain is mostly mountainous, with high altitude areas covered by snow and ice for much of the year. The elevation ranges from sea level to 2,469 *m*, with mostly mountainous terrain and significant snow effects in the winter season. The dominant soil types are cryosols and podzols, which are characterized by low temperatures and low water-holding capacity.

The US has a diverse range of climate and terrain conditions, with various climate types ranging from tropical to polar. The average temperature range varies widely depending on the region. The elevation ranges from sea level to 6,190 *m*, with a diverse range of terrain types such as plains, plateaus, mountains, and coastal regions. The country includes arid desert regions, high mountain ranges, and extensive river systems. The dominant soil types vary widely, from aridisols in the deserts to mollisols in the Great Plains.

I collected daily continuous streamflow records from 575 gauges across the four study regions to conduct our analyses. The corresponding drainage areas range from 4 to 40,504 *km*².

Our analysis was performed on a seasonal basis, considering spring (March-May), summer (June-August), autumn (September-November), and winter (December-February) to account for the seasonality of hydrograph recessions (Tashie et al., 2020b) and flood distributions (Durrans et al., 2003). Each analysis conducted on a specific river gauge during a season was treated as a case study.

I excluded gauges located downstream of large dams in all four regions (Lehner et al., 2011; Wang et al., 2022). Consistent with previous studies (e.g., Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Basso et al., 2021; Basso et al., 2023), I chose case studies in Germany, the UK, and the US characterized by limited snowfall, which minimizes the potential transfer of water across seasons due to strong snow accumulation and melting. Specifically, this condition is defined as having an average daily temperature below zero degrees Celsius during precipitation events for over 50 % of a season (Basso et al., 2021). However, recognizing that recession exponents can inherently capture both linear and nonlinear catchment responses, I intentionally included case studies in Norway, which are characterized by a dominant runoff generation process driven by snow dynamics. This deliberate inclusion provides a counter-verification, allowing us to explore the capability of the recession exponent as a measure of flood tail behavior in regions primarily characterized by snowmelt-driven flood generation processes. In summary, this analysis encompasses regions dominated by both rainfall-driven and snowmelt-driven floods, providing an extensive examination of these factors. These procedures resulted in a total of 1997 case studies, distributed as follows: 540 in spring, 520 in summer, 543 in autumn, and 394 in winter (refer to supporting information Table 4.2 for detailed information of each region).

I incorporated the Köppen climate classification to recognize the spatial distribution of diverse hydroclimatic characteristics. This is obtained from the work presented by Beck et al. (2018), offering high-resolution (1-*km*) maps depicting present-day conditions (1980-2016).

4.2 Inferring heavy tails of flood distributions from common streamflow dynamics

I adopt a framework of the Physically-based Extreme Value (PHEV) distribution of river flows, introduced by Basso et al. (2021). This framework offers a mechanistic-stochastic characterization of both the magnitude and probability of flows, underpinned by essential hydrological processes like precipitation, infiltration, evapotranspiration, soil moisture, and runoff generation within river basins, as previously described in well-established mathematical description (Laio et al., 2001; Porporato et al., 2004; Botter et al., 2007b, 2009).

Within the PHEV framework, I obtain consistent expressions for the probability distributions of various flow metrics, including daily streamflow, ordinary peak flows (local flow peaks resulting from streamflow-producing rainfall events), and floods (flow maxima within a specified timeframe) (Basso et al., 2016). The description of runoff generation and streamflow dynamics provided by this framework has been successfully tested across a diverse range of hydroclimatic and physiographic conditions through a number of studies (Arai et al., 2020; Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Mejía et al.,

2014; Müller et al., 2014; Müller et al., 2021; Pumo et al., 2014; Santos et al., 2018; Schaeffli et al., 2013).

By taking the limit of the PHEV framework, insights into the tail behavior of the flow distribution are obtained (Basso et al., 2015). Wang et al. (2023) showed that the tail of the distribution is exclusively governed by a power law function (indicating heavy tails) when the hydrograph recession exponent exceeds two, signifying discernible nonlinearity of catchment responses. Conversely, the tail appears as nonheavy when the recession exponent is below two, suggesting linearity of catchment responses. As a result, the hydrograph recession exponent has been proposed as a suitable indicator of heavy-tailed flood behavior, based on the analysis of common discharge dynamics.

The hydrograph recession exponent a for each case study can be estimated as the median value of the exponents from power law functions fitted to the pairs of dq/dt and q of individual hydrograph recessions (Jachens et al., 2020; Biswal, 2021). I identify the hydrograph recessions as the ordinary peak flows and the subsequent daily streamflow decay, constrained by a minimum duration of five days.

4.3 Validation of hydrograph recession exponents as an index of heavy-tailed flood behavior

To validate the identification of heavy-tailed flood behavior obtained through estimated recession exponents, I fit a power law distribution to the empirical data distribution and evaluate the reliability of empirical power laws, serving as the benchmark of heavy-tailed flood behavior presented by data.

A case study is considered to exhibit heavy-tailed behavior if the empirical power law effectively describes the tail behavior of the data distribution. I used the Kolmogorov-Smirnov (KS) statistic (κ), a common measure of distance between non-normal distributions, to preliminarily assess empirical power law distribution reliability ($\kappa \in [0, \infty]$, with $\kappa = 0$ denoting highest reliability). To set the reference point of plausible empirical power laws, I employed a method introduced by Clauset et al. (2009), a state-of-the-art approach for fitting empirical power laws. In such an approach, the empirical power law exponent b is computed by fitting a power law to the upper tail of the data distribution. An optimized lower boundary is established by selecting the best fit based on the KS statistic. Subsequently, a Monte Carlo procedure is employed to determine if the fitted power law reliably represents the observed data (based on the KS statistic). This procedure aims to verify whether the residual errors between the data and the power law distribution fall within the range of fluctuations expected from random sampling. If the residual errors lie within this range, the power law is considered a dependable (plausible) representation of the empirical data distribution. We conduct these computations using the Python package `plfit` 1.0.3. I calculate empirical power law exponents b for each case and assess the consistency of identifying heavy-tailed behavior using both a and b .

I conduct our approach using three distinct empirical data distributions: daily streamflow, ordinary peaks, and monthly maxima. These multiple analyses strengthen our validation process and enhance the evaluation of our results. It's worth noting that our chosen benchmark,

the empirical power law, may be influenced by fitting uncertainty due to data scarcity in certain cases, particularly when analyzing maxima. To mitigate this, we consider monthly maxima (Fischer and Schumann, 2016; Malamud and Turcotte, 2006) instead of the seasonal maxima previously used in the literature (e.g., Basso et al., 2021) in order to expand the sample size. Parallel analyses for cases with larger sample sizes (i.e., daily streamflow and ordinary peaks) provide more robust validation and lend support to the interpretation of results for maxima.

In this section, Welch’s t-test (at 0.05 significance level) is also used to determine significant differences ($p < 0.05$) or lack thereof ($p > 0.05$) in mean values between distributions. Such a statistical test was selected due to its robustness in handling skewed distributions, unequal variances, and different sample sizes in the analyzed data (Welch, 1947; Derrick et al., 2016).

4.4 Analyses of spatial and seasonal patterns of inferred flood tail behavior

I construct a geographical representation of inferred heavy-tailed flood behavior by utilizing estimated recession exponents derived from common streamflow dynamics across study countries and for each season. This representation serves as an evaluation of the propensity of heavy-tailed flood behavior across various regions and seasons. I simplify the seasonal results by identifying the dominant tail behavior, which refers to the majority of seasons exhibiting either heavy-tailed or nonheavy-tailed behavior, as the representative inferred flood tail behavior in the analysis of spatial pattern.

To determine the dominant hydroclimatic characteristic of each catchment, we overlay the Köppen climate map (Beck et al., 2018) with the river gauge and catchment boundary data. The most prevalent climate within the catchment (determined by overlapping areas within the boundary, or by the river gauge location if the boundary data is absent) is assigned as the representative feature.

To analyze seasonal patterns, I initially investigate the coherence of inferred flood tail behavior across seasons, focusing on consistency between heavy- or nonheavy-tailed behavior. Catchments with valid recession exponents from only one season are omitted from this analysis. As a result, the selection comprises 98 out of 98 catchments in Germany, 81 out of 82 in the UK, 79 out of 82 in Norway, and 290 out of 313 in the US. I also employ the Wilcoxon signed-rank test (Wilcoxon, 1945), a non-parametric statistical hypothesis test, at a significance level of 0.05 in this section. This test assesses whether the median of recession exponents (within a climate group on a seasonal basis) is above two, below two, or shows no significant difference from two (Figure 4.7).

4.5 Chapter results and discussion

4.5.1 Effectiveness of identifying heavy-tailed flood behavior using common discharge dynamics

Figure 4.1 shows the frequency histograms of KS statistics κ for two groups of cases: red histograms denote cases with recession exponents a above two, and blue histograms denote

those below two. The mean κ is significantly smaller ($p < 0.05$) for the former group (red histograms) compared to the latter one (blue histograms) for the case studies from Germany, the US, and the UK. This result confirms that power law distributions (characterized by heavy-tailed behavior) better represent the empirical data in case studies with recession exponents above two. In the Norwegian case studies, no significant difference was instead identified between the two groups. This is likely due to the absolute values of the recession exponent in this context, which is lower than in the other three countries and mostly comprised between 1 and 2, thus indicating a prevalence of nonheavy-tailed behaviors to date.

To quantify the accuracy provided by the identification of heavy-tailed flood behavior through recession exponents, I set decreasing thresholds for κ , which correspond to increasing reliability of power laws as descriptions of the empirical data. The accuracy of our index (i.e., the recession exponent) can therefore be calculated as $P(a > 2 \mid \kappa < \kappa_r) = N_c(a > 2 \mid \kappa < \kappa_r) / N_c(\kappa < \kappa_r)$, where κ_r is the threshold, $N_c(\kappa < \kappa_r)$ is the number of case studies with $\kappa < \kappa_r$, and $N_c(a > 2 \mid \kappa < \kappa_r)$ is the number of case studies with $a > 2$ among the $N_c(\kappa < \kappa_r)$ case studies. I found that the accuracy is clearly correlated to the reliability level requested for the empirical power laws (represented by κ_r) for case studies in Germany, the US, and the UK. This confirms that the recession exponent provides higher accuracy in detecting heavy-tailed behaviors when the empirical distributions of observed data can be represented by power laws with more certainty, thus underscoring the consistency between identifying heavy-tailed cases by using the proposed index and the observations. The accuracy increases in the same way also for case studies in Norway, but it always remains below 0.5. I will elucidate below reasons and implications of this finding after considering the results presented in Figure 4.2.

In Figure 4.2, I explore the correlation between the values of empirical power law exponents b and the values of recession exponents a for case studies confirmed to exhibit heavy-tailed behavior. This is achieved by utilizing the goodness-of-fit testing procedure of Clauset et al. (2009) to categorize case studies into ‘confirmed power-law-tailed case studies’ and ‘uncertain case studies.’ The former are depicted as black dots, while the latter are depicted as gray dots. The presence of a sizable number of uncertain case studies indicates the difficulty of establishing with certainty whether the underlying distribution of empirical data is or not a power law. This difficulty is often due to limited data availability, although the possibility that they indeed do not follow power laws cannot be excluded.

To highlight the correlation, I binned the confirmed power-law-tailed case studies and used red markers showing the median values of a and b (squares), the interquartile intervals of b (vertical bars), and the binning ranges of a . In each country, the composition of each bin encompasses one-seventh of the total number of case studies, except for Norway, where this fraction is adjusted to one-fifth due to the limited number of confirmed power-law-tailed cases. I calculated Spearman correlations r_s (Spearman, 1904) to test the correlation between a and b , which is valid for both linear and nonlinear associations between random variables. I found that a and b are significantly correlated at a significance level of 0.05 in Germany, the US, and the UK. In these three countries, a larger number of uncertain case studies emerge in the analysis of flow maxima compared to the analysis of daily streamflow and ordinary peak flow (respectively for daily streamflow, ordinary peaks, and flow maxima: 265, 270, and 352 out of 386 case studies in Germany; 258, 280, and 306 out of 325 case studies in the UK; and

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

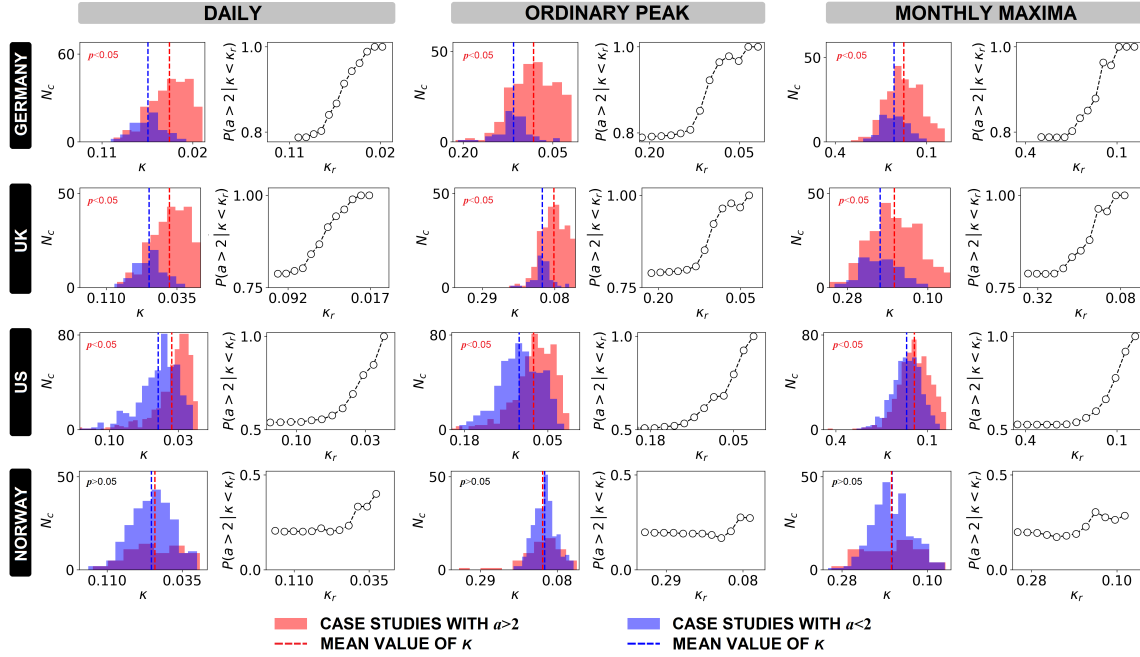


Figure 4.1: Effectiveness of identifying heavy-tailed flood behavior using hydrograph recession exponents Case studies in each country are categorized into two groups: red histograms representing recession exponents a above two and blue histograms representing recession exponents a below two. In all three analyses (daily streamflow, ordinary peak flow, and monthly maxima), every case study is subjected to empirical power law fitting, resulting in a representative power law for the dataset, measured by the KS statistic κ (where $\kappa \in [0, \infty]$, with $\kappa = 0$ signifies maximum reliability). The histograms portray the count of case studies N_c analyzed as a function of κ for two distinct groups. Dashed lines on the histogram plots indicate the means of the histograms. The means of two groups ($a > 2$ and $a < 2$) are subjected to Welch’s t-test at a significance level of 0.05 to determine whether they are significantly different ($p < 0.05$) or not ($p > 0.05$). The line chart shows the accuracy of using the recession exponent to identify heavy-tailed behavior (denoted as $P(a > 2 | \kappa < \kappa_r) = N_c(a > 2 | \kappa < \kappa_r) / N_c(\kappa < \kappa_r)$) as the κ_r threshold decreases (i.e., as the reliability of empirical power laws increases). The results for Germany were initially presented in Chapter 3 but are reproduced in this chapter for the purpose of comparison with case studies from other countries.

589, 624, and 836 out of 980 case studies in the US). Since the same case studies have already been confirmed to exhibit power-law-tailed distributions in their daily streamflow and ordinary peak flow data, the increase of uncertain case studies in the analysis of flow maxima suggests that the greater level of uncertainty is due to limited data availability rather than indicating a rise in the number of non-power-law-tailed case studies.

In Norway, however, the majority of case studies across all three analyses (i.e., daily streamflow, ordinary peaks, and flow maxima) are identified as uncertain (respectively 291, 289, and 300 out of 306 case studies). These results align with the fact that the values of the recession exponent for the Norwegian case studies predominantly fall between 1 and 2 (Figure 4.2), indicating that to date catchment responses are relatively closer to being linear in Norway compared to the other countries, and implying the prevalence of nonheavy-tailed flood behavior. This also explains the pattern presented in the Norway panel of Figure 4.1. Given that the case studies generally have recession exponents below two, the number of case studies with recession exponents above two are not enough to distinguish between the two distributions of κ .

Overall, the effectiveness of recession exponents in distinguishing heavy- and nonheavy-tailed flood behavior has been substantiated (see also Wang et al., 2023b). This differentiation hinges on a critical threshold: the value two. In datasets showcasing diverse physiographical characteristics, the interpretation is consistent. Areas with higher recession exponents (above two), indicating discernible nonlinearity in catchment responses, tend to exhibit heavy-tailed flood behavior. Conversely, regions with lower recession exponents (below two), reflecting relatively linear responses in catchments, are more likely to signify nonheavy-tailed flood behavior.

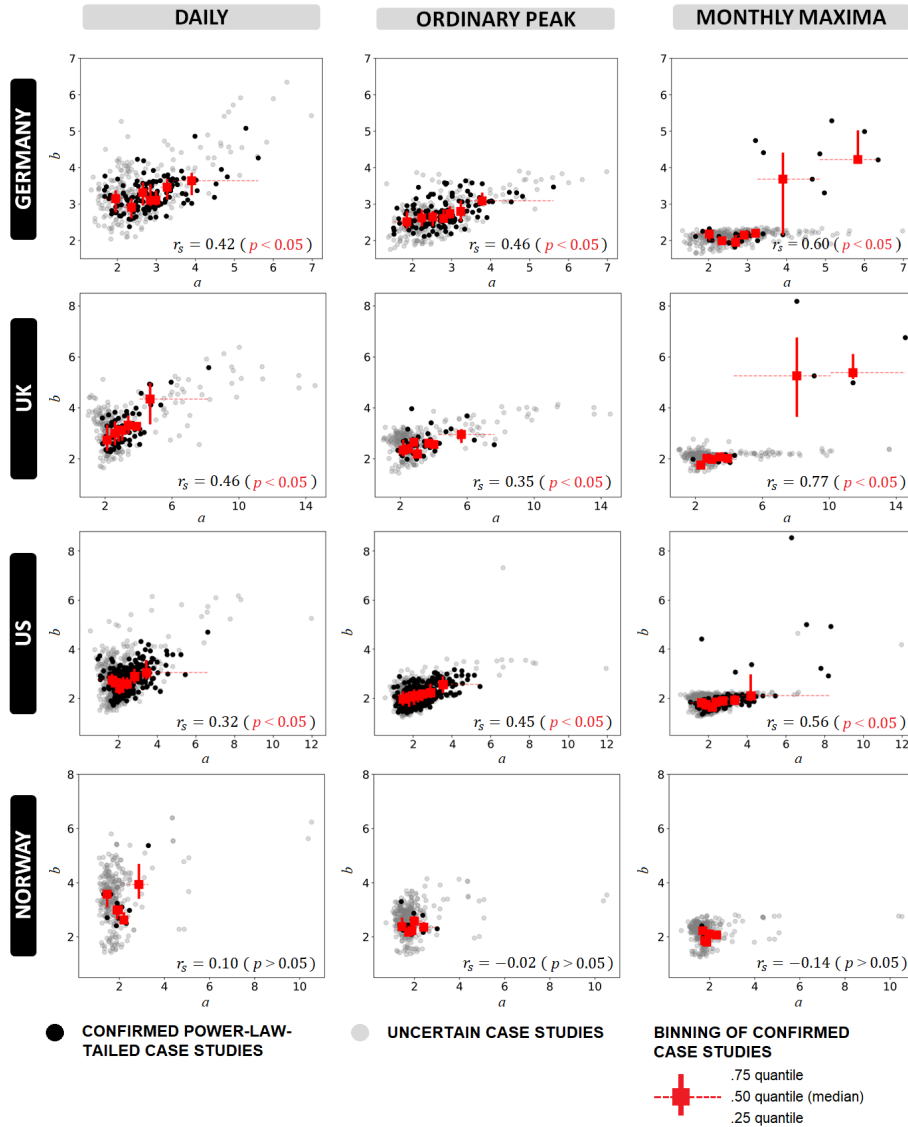


Figure 4.2: Empirical power law exponent b as a function of the hydrograph recession exponent a (physically-based index of heavy-tailed flood behavior) Case studies are classified into groups of confirmed power-law-tailed (black dots) and uncertain (gray dots) case studies on the basis of the goodness-of-fit test (Clauset et al., 2009). The former group shows statistical confirmation that the data's distribution tail can be accurately characterized by a power law, indicating heavy-tailed behavior. Conversely, the latter group indicates our inability to statistically affirm whether the data follows a power law distribution or not. For the confirmed power-law-tailed case studies, the correlation between the empirical power law exponent b and the hydrograph recession exponent a is underlined by red markers. This correlation is quantified using the Spearman correlation coefficient r_s at a significance level of 0.05. The squares represent the median values of a and b , vertical bars indicate the interquartile intervals of b , and horizontal dashed bars indicate the binning ranges of a . In each country, the composition of each bin encompasses one-seventh of the total number of case studies, except for Norway, where this fraction is adjusted to one-fifth due to the constraint posed by the total number of confirmed power-law-tailed case studies. The count of the confirmed power-law-tailed case studies in the analyses of daily streamflows, ordinary peak flows, and monthly flow maxima are as follows: 121, 116, and 34 out of 386 case studies for Germany, respectively; 67, 45, and 19 out of 325 case studies for the UK, respectively; 391, 356, and 144 out of 980 case studies for the US, respectively; and 15, 17, and 6 out of 306 case studies for Norway, respectively. The results for Germany were initially presented in Chapter 3 but are reproduced in this chapter for the purpose of comparison with case studies from other countries.

4.5.2 Spatial patterns of inferred flood tail behavior

Figure 4.3 displays the spatial distribution of dominant flood tail behavior across seasons, based on the recession exponent values, for Germany, the UK, Norway, and the US, respectively. This dominant behavior represents either heavy or nonheavy tails, depending on what is observed in the majority of seasons. Additionally, Figure 4.4 and Table 4.1 provide quantitative analyses of the propensity of flood tail behavior across different regions.

In Germany (Figure 4.3a), approximately 81 % of catchments are identified as sites with dominant heavy-tailed flood behavior (red dots), indicating a prevalence of such behavior. This result agrees with the findings of Mushtaq et al. (2022), which reported that a distribution with a relatively heavier tail (i.e., the log-normal) best represent ordinary peak flows in the majority of German basins considered in their study. The inferred heavy-tailed sites are spread across Germany. They dominate in the eastern part, while there are mixed patterns of heavy- and nonheavy-tailed behavior in the western part. The climate conditions are primarily humid continental (Dfb) and temperate oceanic (Cfb). Humid continental climate is prominent in the east, while temperate oceanic climate generally covers the west.

In the UK (Figure 4.3b), four climate types are present, with temperate oceanic climate (Cfb) being the dominant one. The terrain of this country in comparison to the other three countries is relatively homogeneous, with no high mountains. According to our findings, heavy-tailed flood behavior is prevalent in the UK, with a prevalence of 77 %, especially in the eastern and southern coastal regions. Huntingford et al. (2014) reported a case in which a rapid succession of vigorous Atlantic low-pressure systems crossed much of the UK, resulting in repeated heavy rainfall events. Southeast England was identified as a distinct region characterized by exceptionally high flows, exacerbated by increasingly saturated catchments. These catchment characteristics and hydrological responses align with our findings, which indicate the pronounced heavy tails in such a region.

In Norway (Figure 4.3c), however, nonheavy-tailed flood behavior dominates. Approximately 89 % of sites are inferred to have nonheavy-tailed flood behavior. Norway encompasses nine climate types but is primarily covered by Subarctic climate (Dfc), characterized by low temperatures and reduced evapotranspiration. Hydrological processes are significantly influenced by snow dynamics, which generally determine linear catchment responses as a result of snow accumulation and melting processes (Santos et al., 2018).

In contrast to the aforementioned countries with relatively consistent climate and dominant flood behavior, the US (Figure 4.3d) displays a diverse range of climate types and a balanced propensity toward heavy- and nonheavy-tailed flood behavior. The eastern regions are dominated by humid subtropical climate (Cfa), hot-summer humid continental climate (Dfa), and temperate oceanic climate (Cfb) from south to north. The interior western states feature a cold semi-arid climate (BSk), while mixed patterns are observed in the western mountainous and coastal areas. An overall relatively even distribution of inferred heavy-tailed (52 %) and nonheavy-tailed (48 %) flood behavior prevails in this diverse climate country.

Figure 4.3e provides an example of how the spatial distribution of flood behavior is influenced by regional physioclimatic features. In particular, catchments on the east side of the mountains exhibit pronounced heavy-tailed flood behavior, which aligns with the findings of Smith et al. (2018). This is likely due to the interaction between cold air from the inland polar

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

jet stream and warm ocean currents leads to the formation of Nor'easters, which are synoptic-scale extratropical cyclones in the western North Atlantic Ocean along the US northeast coast. These weather systems often resulted in heavy rain or rain-on-snow events. Conversely, on the west side of mountains, catchments tend to exhibit nonheavy-tailed behavior, potentially due to the leeward rain shadow effect.

In summary, the spatial distributions of inferred flood tail behavior denote that regions with dominant climate types (e.g., Germany, the UK, and Norway) tend to exhibit single or dominant flood tail behavior. Conversely, in regions with diverse climate conditions (e.g., the US), the interplay among regional physioclimatic conditions emerges shows its impacts on the propensity of regional flood behavior.

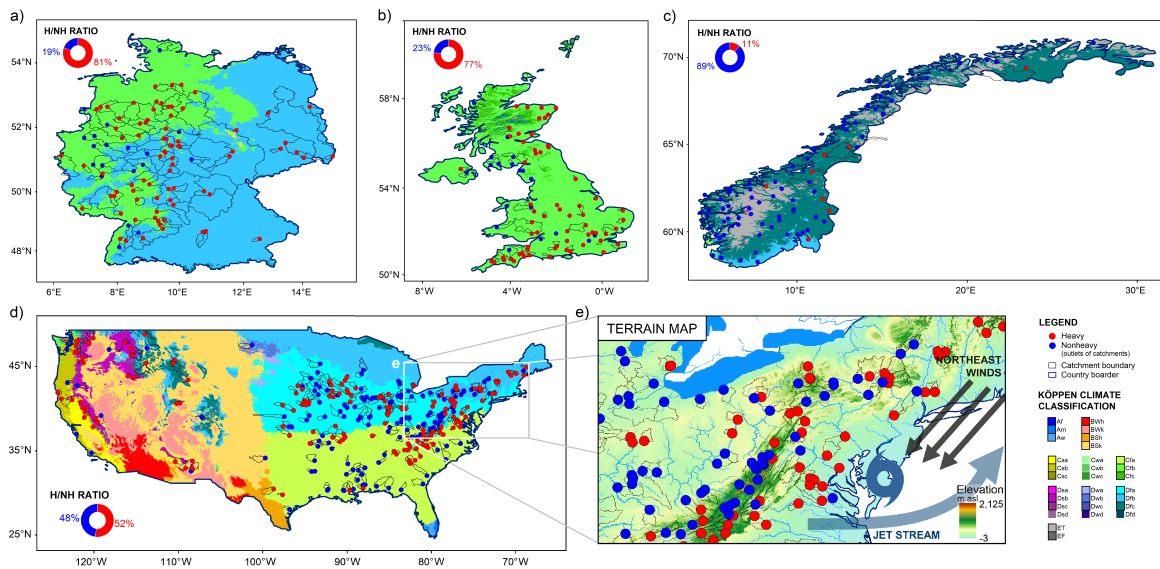


Figure 4.3: Spatial distribution of dominant flood behavior The dominant pattern determines the representative flood tail behavior of catchments across all study countries, whether it is heavy or non-heavy, which is defined by the major pattern recognized across seasons. Tail behavior is inferred by hydrograph recession exponents. The ratio of heavy- to nonheavy-tailed catchments is indicated as H/NH ratio. Köppen climate classification based on Beck et al. (2018), with present climate types outlined by bold dark frames in the legend. (a) Germany, a total of 98 gauges represent catchments ranging from 110 to 23,843 km^2 , with a median area of 1,195 km^2 . (b) The UK, a total of 82 gauges represent catchments ranging from 15 to 9,948 km^2 , with a median area of 283 km^2 . (c) Norway, a total of 82 gauges represent catchments ranging from 4 to 40,504 km^2 , with a median area of 234 km^2 . (Note: some catchment boundaries are absent in the dataset for catchments in the UK and Norway.) (d) The US, a total of 313 gauges represent catchments ranging from 66 to 9,935 km^2 , with a median area of 1769 km^2 . (e) A zoomed-in map illustrates the discernible patterns of flood tail behavior resulting from specific flood generation processes influenced by the interplay between regional terrain and meteorological features. (Note: the cylinder map projection is employed in these maps.)

To obtain quantitative results I examine the predominant flood tail behavior (inferred by recession exponents) of catchments across various climate regions and sort these regions based on the proportion of heavy-tailed catchments from high to low, as illustrated in Figure 4.4. By categorizing climate type regions based on the proportion of heavy-tailed catchments, we establish three groups according to their propensity of flood tail behavior: Heavy-tailed group, indicating regions with over 66.6 % of catchments dominated by heavy tails; Neutral group, encompassing regions with 33.3 % to 66.6 % of catchments dominated by heavy tails, represents a relatively even propensity for both heavy and nonheavy tails in the catchments within these

regions; and Nonheavy-tailed group, representing regions with less than 33.3 % of catchments dominated by heavy tails, denotes the propensity for nonheavy tails. According to the Köppen climate type classification, the overarching hydroclimatic characteristics can be delineated by three hierarchical features: 1. the main group, which encompasses five areas—Tropical, Arid, Temperate, Continental, and Polar; 2. precipitation characteristics; and 3. temperature characteristics. The findings are synthesized in Figure 4.4 and Table 4.1, where the groups of flood tail behavior propensity are juxtaposed with the distinctive traits of each climate region.

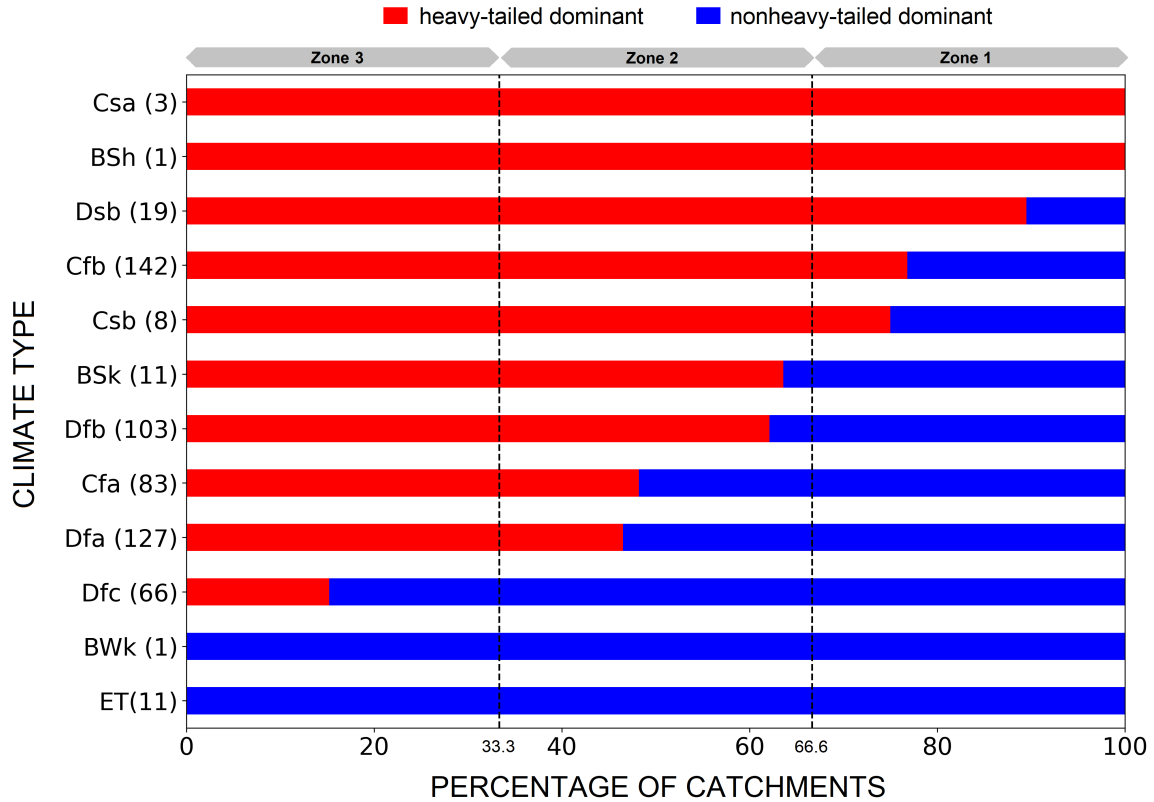


Figure 4.4: Propensity of inferred flood tail behavior in diverse climate regions Catchments are categorized by climate types and grouped by dominant (across seasons) heavy-tailed case percentages. Three groups are defined by heavy-tailed case proportions: Zone 1 ($> 66.6\%$) represents heavy tails, Zone 2 ($33.3\text{--}66.6\%$) is neutral, and Zone 3 ($< 33.3\%$) represents nonheavy tails. The number of catchments in each climate region is indicated in parentheses after the climate type.

Five climate regions are identified as having a higher propensity for heavy tails: mediterranean climate (Csa), hot semi-arid climates (BSh), humid continental climate (Dsb), temperate oceanic climate (Cfb), and cool-summer mediterranean climate (Csb). These regions are characterized by warm to hot temperatures, often accompanied by occasional dry periods (except for Cfb). Based on the definition of Köppen climate classification the occurrence of dry periods is a result of significantly uneven rainfall throughout the year, with at least three times as much rainfall in the wettest month as in the driest month. In semi-arid climates (BSh), there is generally lower annual rainfall (summarized in Table 4.1). Higher temperatures increase the potential evapotranspiration, often enhancing atmospheric moisture content and facilitating convective rainfall. Moreover, the dynamics of evapotranspiration in hillslopes influence the nonlinearity of runoff processes in catchments (Tashie et al., 2019).

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

Table 4.1: Comparison of inferred flood tail behavior propensity with climate characteristics

Propensity of Tail Behavior	Köppen Climate Classification					
	Code	1 st Main Group	2 nd Seasonal Precipitation	3 rd Temperature	Dry Period	Warm-Hot
Heavy (Zone 1)	Csa	Temperate	Dry Summer	Hot Summer	•	•
	BSh	Arid	Semi-Arid	Hot	•	•
	Dsb	Continental	Dry Summer	Warm Summer	•	•
	Cfb	Temperate	No dry season	Warm Summer		•
	Csb	Temperate	Dry Summer	Warm Summer	•	•
Neutral (Zone 2)	BSk	Arid	Semi-Arid	Cold	•	
	Dfb	Continental	No dry season	Warm Summer		•
	Cfa	Temperate	No dry season	Hot Summer		•
	Dfa	Continental	No dry season	Hot Summer		•
Nonheavy (Zone 3)	Dfc	Continental	No dry season	Cold Summer		
	BWk	Arid	Dessert	Cold	•	
	ET	Polar	–	Tundra		

Dry periods can lead to lower catchment soil moisture, facilitating nonlinear runoff generation (Merz and Blöschl, 2009; Viglione et al., 2009). The findings presented here indicate that heavy-tailed flood behavior tends to emerge due to the substantial nonlinearity observed in catchment hydrological processes, which is facilitated by temporally uneven rainfall and higher evapotranspiration variation throughout the year.

I also find that certain regions show a relatively neutral propensity regarding flood tail behavior (either heavy- or nonheavy-tailed) and aggregate them into the second group of Figure 4.4 and Table 4.1. These regions encompass cold semi-arid climates (BSk), humid continental climate (Dfb), humid subtropical climate (Cfa), and humid continental climate (Dfa). While cold semi-arid climates (BSk) experience dryness, they are characterized by very limit precipitation. In the other three regions (Dfb, Cfa, and Dfa), heavy tails may still occur due to higher evapotranspiration, which is driven by high temperatures. However, the relatively even distribution of rainfall throughout the year in these regions may reduce the propensity for heavy tails, resulting in a smoother occurrence of heavy-tailed flood behavior. In summary, the regions in this group still have a certain probability of exhibiting heavy-tailed flood behavior. However, the absence of either a drier state of the catchment (caused by uneven rainfall) or higher temperatures (that ensure sufficient atmospheric moisture for rainfall and strengthened nonlinearity) could constrain the occurrence of such behavior.

In the last group, which includes regions with subpolar climate (Dfc), tundra climate (ET), and cold desert climates (BWk), there is a higher propensity for nonheavy tails, and the two evident factors for heavy tails recognized from previous results are generally lacking. Runoff generation in Dfc and ET is primarily driven by snow dynamics, with snowmelt being the main contributor to runoff. Snowmelt is highly dependent on energy capacity, resulting in hydrological responses that are more likely to exhibit linearity. This favors the occurrence of nonheavy-tailed flood behavior (Thorarinsdottir et al., 2018). Catchments located in the region of BWk exhibit nonheavy-tailed behavior might also be attributed to limited precipitation in desert.

In this study, I do not find substantial influences of the general hierarchical feature (especially the temperate and continental climate classifications) on the propensity of flood tail behavior.

To sum up this section, I have identified the conjunction of dry periods and higher temperatures as crucial meteorological factors significantly contributing to the dynamics of catchment storage, thereby influencing the nonlinearity of hydrological responses. These findings shed light on the interplay between catchments and meteorological conditions in the manifestation of heavy-tailed flood behavior. We acknowledge that these results are based on overarching conditions and do not encompass all climate types, and achieving an equal number of study sites across various climate regions might not always be feasible. Expanding the number of study sites could further enhance our understanding, especially for extreme cases.

4.5.3 Seasonal patterns of inferred flood tail behavior

I analyze the seasonality of flood tail behavior, an aspect of this phenomenon which has been previously suggested but remains poorly understood (Durrans et al., 2003; Basso et al., 2015; Macdonald et al., 2022). Figure 4.5 illustrates the spatial distribution of catchments with consistent tail behavior across seasons (i.e., with either heavy or nonheavy tails across all seasons; black triangles) and those with varying tail behavior across seasons (green dots). Catchments exhibiting inconsistent behavior spread across the whole US and Germany, whereas they are mostly concentrated in the southern parts of the UK and in the central mountainous regions of Norway. The percentages of catchments exhibiting inconsistent flood tail behaviors are respectively 33 %, 33 %, 17 %, and 34 % in the US, Germany, the UK, and Norway. The results indicate that although the majority of catchments tend to exhibit stable heavy-/nonheavy-tailed behavior, still around one-third reveal changing patterns across seasons. Notably, there is a particularly high percentage of consistent patterns (83 %) in the UK, likely due to the relatively uniform climate and terrain conditions across the country characterized by continuous rainfall throughout a year (as shown in Figure 4.3b).

I further investigate the dynamics of heavy- and nonheavy-tailed case studies across seasons in Figure 4.6. Heavy-tailed case studies increase from spring to autumn (approximately corresponding to the growing season in the northern hemisphere) and decrease from autumn to spring (approximately corresponding to the dormant season in the northern hemisphere), as seen in the aggregated patterns across all regions (panel a). This pattern can be attributed to the increasing temperature in the growing season, during which increasing evapotranspiration consumes water storage in the shallow subsurface, escalating the nonlinearity of catchment responses (Tashie et al., 2019). The seasonality of evapotranspiration effects on catchment nonlinearity is supported by the findings of Tarasova et al. (2018), who observed clear seasonal dynamics of catchment average runoff coefficients. These coefficients tend to be higher in wet winters and lower in dry summers. It has been shown that significant variation in runoff coefficients is linked to high nonlinearity of hydrological responses, facilitating heavier-tailed floods. This phenomenon is often observed in dry catchments (Merz and Blöschl, 2009). Other studies confirmed that the nonlinearity of catchment responses favors the emergence of heavy-tailed flood behavior (Gioia et al., 2008; Rogger, Pirkl et al., 2012; Basso et al. 2015),

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

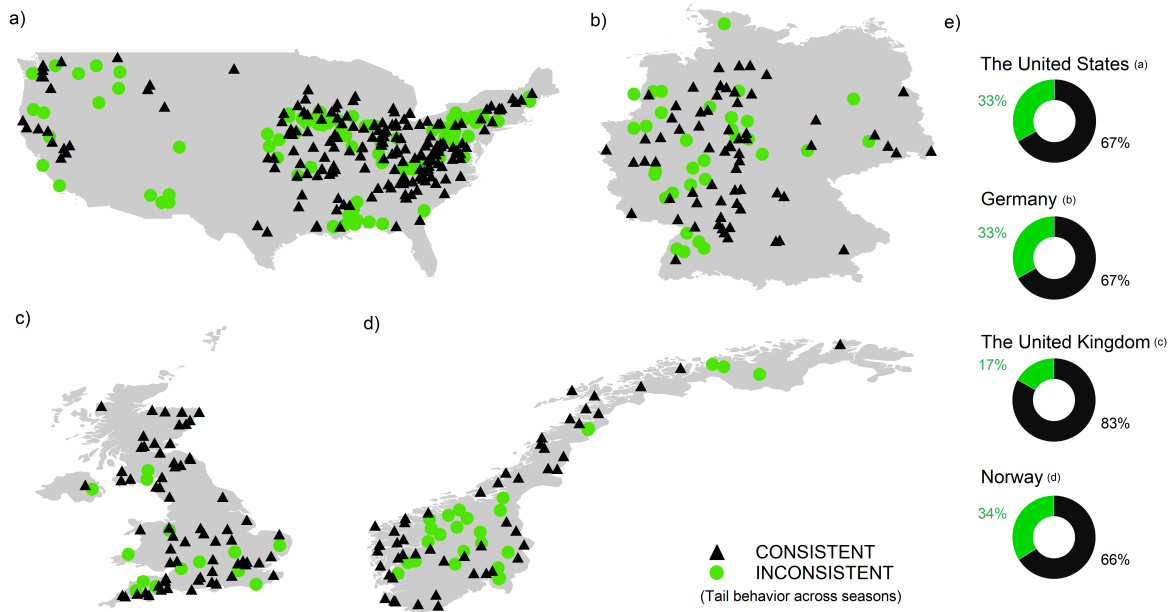


Figure 4.5: Consistency of inferred flood tail behavior across seasons (a) 290 catchments in the US. (b) 98 catchments in Germany. (c) 81 catchments in the UK. (d) 79 catchments in Norway. (e) Percentage of consistent and inconsistent catchments in each country.

and is often expressed by quicker recession during high flow periods and greater stability during low flow periods. Conversely, during the dormant season, nonlinearity decreases due to reduced competition from evapotranspiration and replenished water storage. I underscore that the significant variability in evapotranspiration amplifies the fluctuation of catchment storage conditions, causing soil moisture levels to oscillate between drier and wetter states. This alternation leads to the occurrence of both very small and very large events, which are characteristic of heavy-tailed flood behavior.

This dynamic is particularly pronounced in the US (panel b), where is characterized by a wide range of geography and diverse temperate and continental climates. The number of inferred heavy-tailed cases can increase by 50 % from spring to autumn. In Germany and the UK (panels c and d), heavy-tailed behavior is relatively prevalent and shows no significant distinction from spring to autumn, but still experiences a noticeable decrease in winter, likely due to lower temperatures and evapotranspiration. Norway (panel e) presents different patterns due to varying controls on runoff generation. A slight increase in heavy-tailed cases during the winter is observed, which could be attributed to a relatively higher contribution of rainfall-driven flood events during a season when snowmelt-driven events are less common.

I delve into the seasonal characteristics of this behavior further by combining the regional patterns based on climate classification. In Figure 4.7, the square dots represent the median of each box, marked as red, blue, or black to indicate the significance of its value above 2 (heavy), below 2 (nonheavy), or not significantly different from 2, respectively. The last one (black squares) may imply an equal occurrence of heavy- and nonheavy-tailed cases or a lack of samples to draw conclusions. Based on the patterns of significance across seasons, regions with seasonality (defined as having different tail behavior propensity across seasons according to the significant values of median) are grouped in the white area, while those considered stable in heavy tails are in the red area, and those stable in nonheavy tails are in the blue area. For

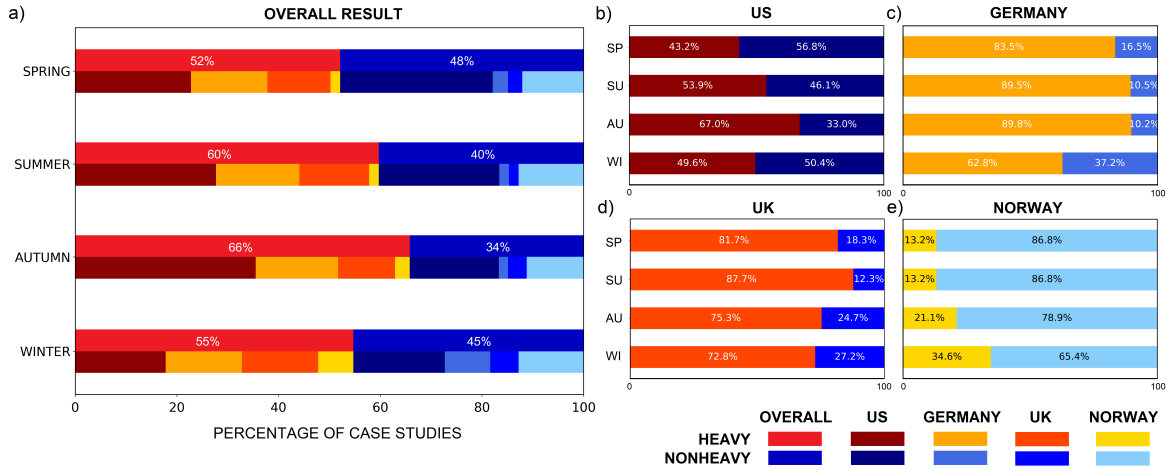


Figure 4.6: Seasonal variations in the percentage of inferred flood tail behavior between heavy and nonheavy case studies (a) The aggregated results encompass all study regions, while the second line provides a breakdown by country. In total, there are 1,997 case studies composed by 540 in spring, 520 in summer, 543 in autumn, and 394 in winter. (b)-(e) Results for each study country (see supporting information Table 4.2 for detailed case numbers across seasons in each country).

regions where statistical significance cannot be concluded for all seasons, we group them based on the absolute values of their medians.

I find that the grouping based on their distinct patterns of seasonality (Figure 4.7) closely aligns with the grouping based on the analysis of dominant patterns throughout the year (Figure 4.4 and Table 4.1). Regions (red area in Figure 4.7 corresponded to the heavy-tailed group in Table 4.1) characterized by uneven rainfall distribution throughout the year, leading to pronounced fluctuations between drier and wetter soil states, combined with higher evapotranspiration rates (indicated by warm to hot temperatures), tend to exhibit a dominance of heavy-tailed behavior in their hydrological responses across all seasons. In areas (white area in Figure 4.7 corresponded to the neutral group in Table 4.1) where rainfall is more evenly distributed annually, the emergence of heavy-tailed behavior is often linked to increased evapotranspiration during the growing seasons, particularly in spring and summer, and is less prominent during dormant seasons. This mechanism, which depends on evapotranspiration dynamics, substantiates the seasonality of flood tail behavior. Regions (blue area in Figure 4.7 corresponded to the nonheavy-tailed group in Table 4.1) where runoff generation is primarily influenced by snow dynamics tend to display linear hydrological responses. This is due to the fact that most runoff in these areas results from snowmelt during the growing seasons, driven by energy availability. These findings support the proposed mechanism of heavy-tailed flood behavior concluded in the spatial analyses and further demonstrate the pivotal effect played by the variation of evapotranspiration and catchment storage on the emergence of heavy-tailed flood behavior.

In summary, while heavy-/nonheavy-tailed behavior is generally consistent across seasons, there is a certain probability for cases to exhibit seasonality. This seasonality of inferred heavy-tailed behavior shows a dynamic pattern of increasing during the growing season and decreasing during the dormant season. Regions with pronounced temperature variations

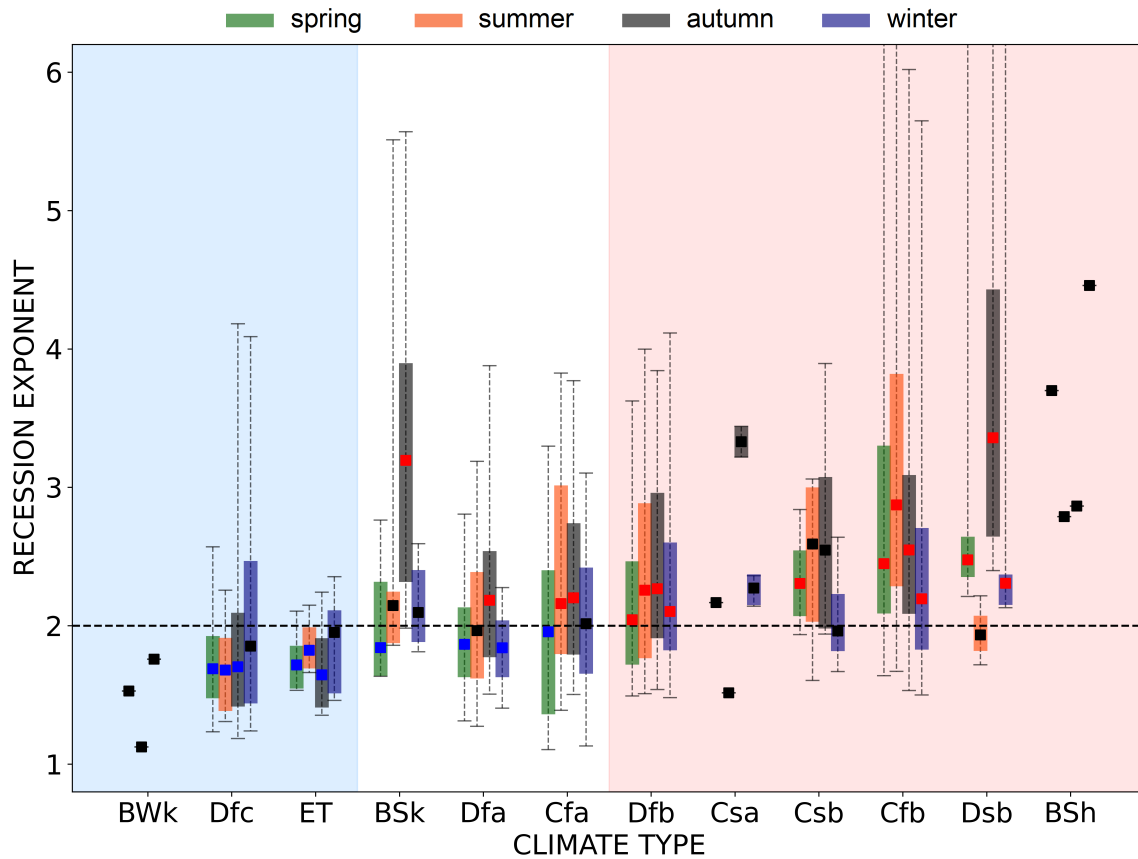


Figure 4.7: Seasonal variations in recession exponents (inferring flood tail behavior) across diverse climate regions Case studies grouped by climate regions based on seasons. Medians of recession exponents in each group are compared with a value of two using Wilcoxon signed-rank test (significance level: 0.05). Red squares indicate significantly heavy-tailed (recession exponents > 2) groups, blue squares indicate significantly nonheavy-tailed (recession exponents < 2) groups, and black squares denote insignificance. Climate regions are categorized as follows: the red area denotes regions with prominent heavy tails across seasons, the blue area denotes regions with prominent nonheavy tails across seasons, and the white area denotes regions with significant seasonality in flood tail behavior.

across seasons, particularly with hot summers, display such dynamics and highlight the role of evapotranspiration in catchments in driving this seasonality.

4.5.4 Factors associated with catchment scales and their role in flood tail behavior

It remains unclear how flood tail behavior varies across catchment scales and what the underlying drivers and mechanisms are (Merz et al., 2022). I employ catchment nonlinearity, represented by recession exponents, to explore the influence of catchment scales on flood tail behavior, as depicted in Figure 4.8. I utilize the categorization of regions characterized by distinct controls on flood tail behavior, primarily influenced by characteristic runoff generation processes (as three groups identified in Figure 4.7), to elucidate the underlying mechanisms. Case studies are categorized into bins based on catchment areas, with the median values represented by squares, interquartile intervals depicted by vertical bars, and catchment area ranges indicated by horizontal dashed bars. Panels a, b, c, and d present results for all regions, regions exhibiting significant heavy tails across seasons, regions with a neutral propensity

and seasonal variations, and regions displaying pronounced nonheavy tails across seasons, respectively. Each panel comprises a total of 30 bins, with approximately 67, 33, 24, and 10 case studies in panels a, b, c, and d, respectively (with minor variations due to rounding).

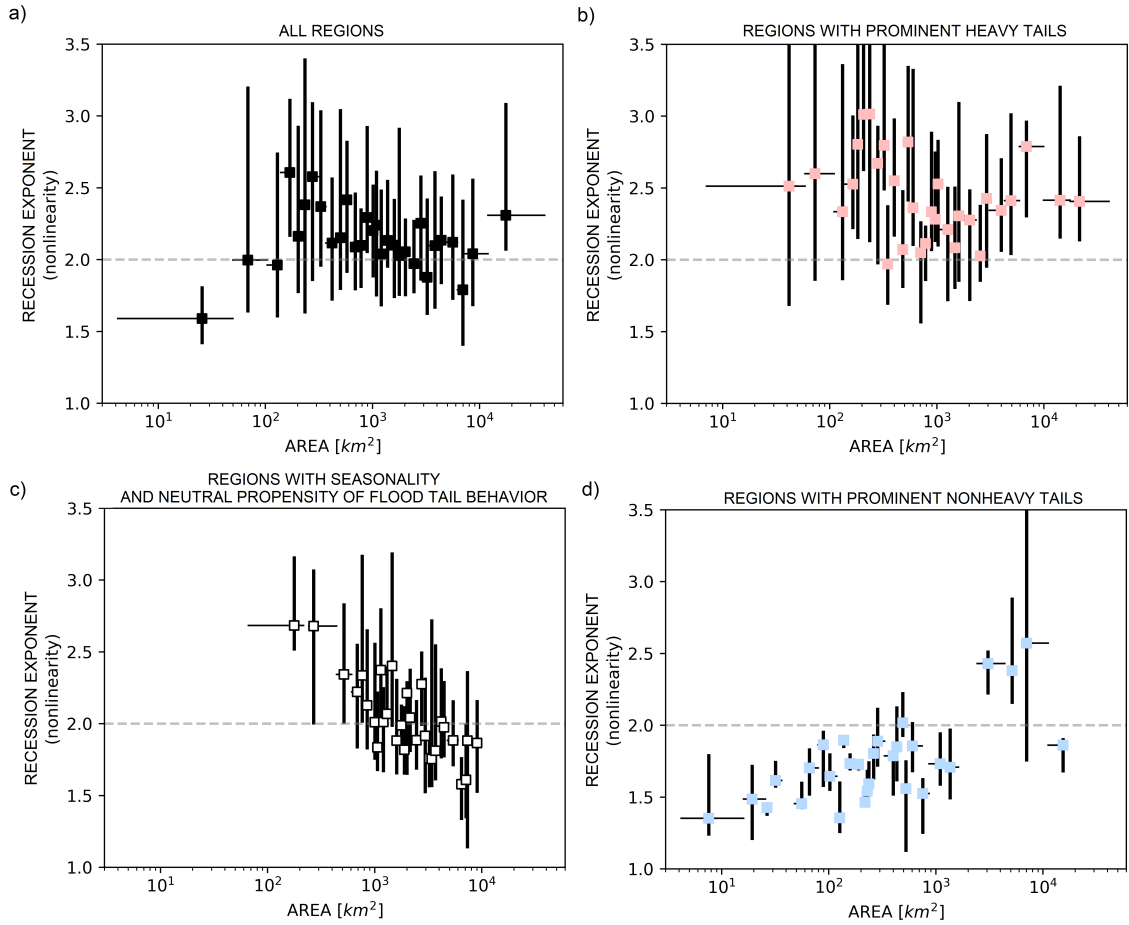


Figure 4.8: Catchment nonlinearity as a function of catchment area The recession exponents, representing catchment nonlinearity, have been evenly grouped into bins based on catchment areas. The squares denote the median values, vertical bars represent the interquartile intervals of the recession exponents, and horizontal dashed bars indicate the catchment area ranges for each bin. (a) All regions (encompassing case studies, $n = 1997$). (b)-(d) show case studies separately according to categorization recognized in Figure 4.7. (b) Regions with prominent heavy tails ($n = 978$). (c) Regions with seasonality and neutral propensity of flood tail behavior ($n = 733$). (d) Regions with prominent nonheavy tails ($n = 286$). In each panel, there are a total of 30 bins, each containing approximately 67, 33, 24, and 10 case studies in panels a, b, c, and d, respectively (with slight variations due to rounding).

From the perspective of all case studies (Figure 4.8a), the pattern appears somewhat unclear. Apart from the case studies involving extremely small and large catchment areas, there seems to be a decrease in nonlinearity as catchment areas increase. Nevertheless, the relationship is rather weak and lacks clarity. These findings align with previous discussions on this matter (e.g., Merz and Blöschl, 2009; Villarini and Smith, 2010; Smith et al., 2018), which have suggested a relatively weak inverse correlation between catchment area and the occurrence of heavy-tailed flood behavior.

However, this study can evidently clarify this relationship by considering the distinct runoff generation processes recognized in different regions. Panel b illustrates that catchment area plays no significant role in catchment nonlinearity in regions characterized by prominent heavy

tails. Whereas a clear inverse relationship between nonlinearity and catchment area is shown in panel c, representing regions characterized by a neutral propensity for heavy and nonheavy tails. In contrast, a proportional relationship between nonlinearity and catchment area is identified in panel d, representing regions characterized by prominent nonheavy tails.

As shown by the previous sections, nonlinearity in neutral regions is primarily driven by high evapotranspiration facilitated by high temperatures. When the catchment area increases, it has a higher chance of encompassing diverse terrain types, including areas with higher altitudes, such as mountainous regions. Increased altitude tends to result in lower temperatures and evapotranspiration rates, negating the evapotranspiration variation and its impact on catchment nonlinearity, which is the main driver of nonlinearity in this region and thus substantiates an inverse relationship (Figure 4.8c). In regions with prominent heavy tails (Figure 4.8b), nonlinearity is generated from the interplay of uneven rainfall and evapotranspiration dynamics, and the enlargement of catchments does not substantively change this relationship. For regions with prominent nonheavy tails (Figure 4.8d), the underlying mechanisms are similar to the neutral regions but work in the opposite direction due to the differently dominant mechanism. Recall that the runoff process in this region is generally dominated by snow dynamics. The region is mainly located in high mountain or high latitude areas. As catchments expand, more diverse terrain is encompassed, potentially introducing a mixture of flood generation processes due to the incorporation of lowland or coastal areas. Particularly, more rain-on-snow events or rainfall-driven events may be encompassed in a same catchment together with snowmelt-driven events (Vormoor et al., 2016). Therefore, an increase in nonlinearity is facilitated due to the mixture of flood types, favoring the emergence of heavier tails in flood distributions (Tarasova et al., 2020a). It should be noted that the tail patterns, based on Figure 4.8d, are still more likely to be nonheavy tails (i.e., recession exponents below two), even though nonlinearity indeed appears to show an increasing tendency along with catchment area.

These findings disentangle the relationship between flood tail behavior (inferred from catchment nonlinearity) and catchment scale, and provide a mechanistic understanding that underscores the role of variability in runoff generation processes introduced by the expansion of catchment area.

4.5.5 Discussion

I have confirmed the effectiveness of the recession exponent in identifying heavy-tailed flood behavior in case studies across countries with varying degrees of the propensity of such behavior: heavy-tailed richness (Germany and the UK), neutrality (the US), and nonheavy-tailed richness (Norway). This validation is substantiated by confirmed power law tailed cases, widely acknowledged as representatives of heavy-tailed distributions (El Adlouni et al., 2008; Clauset et al., 2009), and supported by the significance of catchment nonlinearity as a robust driver of heavy-tailed flood behavior (Fiorentino et al., 2007; Struthers and Sivapalan, 2007; Gioia et al., 2008; Rogger, Pirkl et al., 2012; Basso et al., 2015; Merz et al., 2022; Basso et al., 2023; Wang et al., 2023b).

The findings first indicate that regions with relatively uniform hydroclimatic conditions (Germany, the UK, and Norway) tend to exhibit a single/dominant propensity of flood tail

behavior. Conversely, in regions characterized by diverse conditions (the US), inferred flood tail behavior presents a balance between heavy- and nonheavy-tailed cases in terms of frequency and distribution. Climate conditions have been found shaping the catchment geomorphology (Wu et al., 2023) and river network dynamics (Ward et al., 2020) which contribute to the degree of catchment response nonlinearity (Biswal and Marani, 2010). Meanwhile, the changes in flood generation processes can significantly affect the frequency of large floods (Tarasova et al., 2023), potentially altering flood tail behavior. Our findings in Figure 3e exemplify how different flood generation processes, influenced by the interplay of varied hydrometeorological and terrain conditions, result in opposite flood tail propensities.

I further identify key drivers of heavy-tailed flood behavior by conducting large scale physioclimatic analyses. Specifically, our findings reveal that regions with a pronounced propensity for heavy tails exhibit distinct characteristics: the presence of a dry period and higher temperatures (as shown in Figure 4.4 and Table 4.1). This aligns with previous studies based on the mathematical analysis which associates heavier-tailed flood behavior with a lower frequency of streamflow-triggering rainfall events. Such lower frequency often results from erratic rainfall patterns and higher rates of evapotranspiration, leading to drier catchment conditions (Botter, 2010; Basso et al., 2016). In line with this theory, our large scale analysis provides evidence by showing a prevalent propensity for heavy tails in regions characterized by uneven rainfall patterns throughout the year (i.e., more erratic rainfall), contributing to the presence of dry periods, along with higher potential evapotranspiration rates, as indicated by higher temperatures.

The underlying mechanism of the emergence of heavy-tailed flood behavior is attributed to variations in catchment water storage. In wetter catchments, relatively stable runoff coefficients are observed due to consistent high levels of soil moisture across events. In contrast, drier catchments exhibit larger variations in runoff coefficients between small and large events (Merz and Blöschl, 2009; Viglione et al., 2009). This increased variability in runoff coefficients results in high nonlinearity of catchment responses, favoring heavy-tailed flood behavior. Previous studies have suggested the prevalence of heavy tails in drier catchments (Molnar et al., 2006; Merz and Blöschl, 2009; Guo et al., 2014). Our findings show that this mechanism is primarily driven by concurrent higher evapotranspiration and lower rainfall in summer, as well as lower evapotranspiration and higher rainfall in winter. These conditions lead to variations in storage, enabling the occurrence of both very small and very large flood events, thereby resulting in heavy-tailed flood behavior. In line with this, Tarasova et al. (2018) observed clear seasonal dynamics of catchment average runoff coefficients in Germany, with higher values in wet winters and lower values in dry summers.

The seasonality of flood tail behavior has been suggested in previous studies but remains less understood (Basso et al., 2015; Smith et al., 2018; Macdonald et al., 2022). It's noteworthy that more than one-third of catchments appear to exhibit inconsistent flood tail behavior across seasons (Figure 4.5). In these catchments, some seasons show a tendency toward nonheavy tails, while others tend to display heavy tails. Identifying these catchments and understanding the factors driving them to exhibit heavy tails is vital for hazard assessment. This understanding allows us to pinpoint catchments where extreme floods could potentially occur, even if methods solely based on annual maximum floods might estimate the flood tail

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

as nonheavy based on annual maxima, when heavy tails can still occur within a single season. We have identified that regions characterized by stronger evapotranspiration dynamics across seasons favor this seasonality of flood tail behavior, as it leads to larger variations in water storage during higher evapotranspiration seasons, such as the growing seasons (highlighted in white in Figure 4.7). This finding aligns with previous studies that have observed similar seasonal dynamics in the nonlinearity of hydrological responses (Tashie et al., 2019; Tarasova et al., 2018).

In this study, I also found that the relationship between flood tail behavior and the expansion of catchment scales can be explained by changes in catchment nonlinearity, which are influenced by distinct flood generation processes. Previous studies have presented diverse perspectives on the relationship between flood tail behavior and catchment scales. While some studies have suggested that smaller catchments tend to exhibit heavier tails (e.g., Meigh et al., 1997; Pallard et al., 2009), others have noted a similar trend but with only a weak correlation (Merz and Blöschl, 2009; Villarini and Smith, 2010). Meanwhile, some studies have found no significant relationship between these two variables (Morrison and Smith, 2002; Smith et al., 2018). These studies have explored this topic without reaching a consensus, and many conclusions lack sufficient evidence and a clear understanding. In contrast, our findings (Figure 4.8) distinctly differentiate between various patterns by considering region classifications based on distinct dominant flood generation processes, thereby providing a mechanistic understanding. As a catchment expands, it encompasses more diverse terrain, which in turn facilitates a wider range of altitudes and flood types. In regions where tail behavior is primarily influenced by evapotranspiration dynamics (Figure 4.8c), the presence of diverse altitudes tends to moderate the effect of higher temperatures, reducing the influence of high evapotranspiration on the emergence of heavy tails. In regions where tail behavior is primarily controlled by snowmelt (Figure 4.8d) (mainly composed of catchments in Norway in this study), it has been shown that larger catchments are more likely to encompass a mix of flood types, including snowmelt-driven and rainfall-driven floods (Vormoor et al., 2016). Merz et al. (2022) suggested that heavier-tailed behavior in rainfall-driven floods tends to dominate in such mixed conditions. My findings support this hypothesis by demonstrating an increase in tail heaviness as catchment area enlarges. In regions where heavy tails are pronounced due to the strong nonlinearity resulting from the interplay of uneven rainfall and high evapotranspiration, there is no significant relationship between catchment nonlinearity and catchment area (Figure 4.8b). This lack of relationship may be because the expansion of the catchment area does not appear to significantly enhance or reduce this interplay.

To summarize the findings and underscore the contributions of this study, we benchmark them against the existing hypotheses proposed in the state-of-the-art review of heavy-tailed flood distributions (Merz et al., 2022). These hypotheses (highlighted in italics) provide a framework for understanding the factors influencing flood tail behavior, and our study sheds light on which of these hypotheses receive stronger support or require further refinement. We acknowledge that this summary does not cover all the hypotheses proposed in the review due to the scope of this study. Instead, it primarily focuses on the compartments of the atmosphere and catchment:

“Hypothesis 2 (of the review paper): The Characteristic Flood Generation Process Shapes the Upper Flood Tail Catchments.” While previous studies have hinted at the possibility that regions where flood generation is dominated by rainfall-driven floods tend to exhibit heavier-tailed flood behavior compared to regions dominated by snowmelt (Bernardara et al., 2008; Thorarinsdottir et al., 2018), more explicit process explanations are desired. In line with this hypothesis, I present further evidence showing that the specific nonlinearity inherent in each flood generation process is the primary driver of flood tail behavior. Specifically, I show that in snowmelt-dominated regions, such as the case studies in Norway, hydrological responses closely resemble linear behavior and thus floods tend to exhibit pronounced nonheavy-tailed behavior. Conversely, heavy-tailed floods are more prominent in regions like the UK, where hydrological responses display nonlinearity (as indicated by recession exponents above two). In these areas, flood generation processes are primarily driven by rainfall events. Furthermore, our study reveals that flood generation processes are significantly influenced by the interplay between regional terrain and meteorological features. These factors, in turn, impact the nonlinearity of hydrological responses and can determine the occurrence of heavy or nonheavy tails in flood distributions (Figure 4.3e).

“Hypothesis 3: Mixture of Flood Event Types Generates Heavy Tails.” One argument presented in this hypothesis is that heavy tails may arise from the presence of a flood type displaying heavy-tailed behavior within a mixture of processes (Morrison and Smith, 2002; Villarini and Smith, 2010). However, studies exploring the relationship between the mixture of flood types and flood tails have been lacking. My research addresses this gap by demonstrating that in regions primarily characterized by nonheavy-tailed floods, driven mainly by snowmelt, the tail heaviness increases as catchment areas expand. This increase is likely attributed to the incorporation of additional flood types, especially those associated with rainfall processes occurring in lowland and coastal areas, as catchment areas expand. Thus, our findings provide evidence that supports this hypothesis.

“Hypothesis 4: Non-Linear Response to Precipitation Causes Heavy Flood Tails.” Studies have consistently highlighted the significance of nonlinearity in hydrological processes within catchments as a key determinant in the emergence of heavy-tailed flood behavior (e.g., Struthers and Sivapalan, 2007; Rogger, Pirkl et al., 2012; Basso et al., 2015). In this research, I contribute by introducing a quantitative approach that employs hydrograph recession exponents as a measure of nonlinearity in flood tail analyses and validate its effectiveness in identifying heavy-tailed flood behavior in a large scale analysis. While nonlinearity has long been acknowledged as a contributing factor, my works uniquely utilizes this driver as a reliable index by establishing a specific recession exponent threshold that robustly discriminates heavy-tailed distributions, characterized by power-law tails, from nonheavy ones, offering a valuable tool to the field. Furthermore, our large scale analysis identifies rainfall unevenness and high temperatures as crucial drivers behind the observed nonlinearity in flood responses. Specifically, they intensify catchment soil dryness and amplify water balance storage variations, thereby facilitating both very small and very large runoff events, translating into heavy-tailed flood behavior.

“Hypothesis 5: Drier Catchments Have Heavier Flood Tails Due To Interaction of Water Balance Processes.” In alignment with previous studies that suggest the water balance processes in drier catchments contribute to the emergence of heavy-tailed flood behavior (e.g., Molnar et

al., 2006; Merz and Blöschl, 2009; Guo et al., 2014), we emphasize the critical interplay between uneven rainfall and evapotranspiration dynamics in facilitating these processes and shaping such the behavior. Specifically, our findings show that heavy-tailed flood behavior is more likely to occur in catchments characterized by lower rainfall and higher evapotranspiration in one season (e.g., summer), contrasted with more rainfall and lower evapotranspiration in another season (e.g., winter). When one of these conditions is lacking, heavy-tailed behavior may be less pronounced. For example, regions classified as BSh and BSk, both of which exhibit semi-arid characteristics based on their rainfall patterns, exhibit differences in the prevalence of heavy-tailed cases. BSk regions, despite their semi-arid status, exhibit fewer pronounced heavy-tailed cases due to colder temperatures (Table 4.1) and only show a higher rate of heavy-tailed cases during the summer (Figure 4.7). This interplay highlights the importance of considering the seasonality of flood tail behavior, particularly in regions that do not experience significant dry periods based on their rainfall patterns. In such regions, heavy tails are still likely to occur in seasons with higher evapotranspiration rates (indicated by the white area in Figure 4.7).

“Hypothesis 6: Smaller Catchments Have Heavier Flood Tails Due To Less Pronounced Spatial Aggregation Effects.” A commonly debated question among hydrologists is whether the roles identified in large catchments are applicable to smaller ones, and vice versa. This issue has also arisen in discussions regarding flood tail heaviness, but evidence on the matter has been scattered. While smaller catchments have been suggested to exhibit heavier tails (Meigh et al., 1997; Pallard et al., 2009), previous research has revealed weak (Merz and Blöschl, 2009; Villarini and Smith, 2010) to no (Morrison and Smith, 2002; Smith et al., 2018) correlations between catchment size and tail heaviness. My findings (Figure 4.8) help clarify the relationship between catchment nonlinearity (used as an indicator of tail heaviness) and catchment sizes. We observe distinct patterns among regions characterized by strong, neutral, and weak conditions of heavy tail behavior. These findings underscore the importance of considering the dominant flood generation processes in each region and elucidate how catchment size interacts with flood tail behavior by influencing these dominant processes—either amplifying, reducing, or having no significant effect.

4.6 Chapter conclusions

I analyze common streamflow dynamics to infer heavy-tailed flood behavior by employing a recently developed index of tail heaviness, namely the hydrograph recession exponent. The wide-ranging dataset allows for unveiling spatial and seasonal patterns of flood tail behavior, and to construct a geography of heavy-tailed flood distributions. We analyze and discuss the underlying influences of hydroclimatic settings on this geographical patterns, as represented by Köppen climate characteristics. The main findings of this study can be summarized as follows:

1. **Capability of Recession Exponents for Detecting Heavy-Tailed Flood Behavior:**

The capability of this index to discern between case studies which display heavy-tailed flood distributions and those exhibit nonheavy-tailed behavior is validated by using empirical data from catchments across Germany, Norway, the UK, and the US. This extensive analysis provides a well-rounded evaluation due to the inclusion of regions with

divergent conditions, such as rainfall-driven floods (Germany, the UK, and the US) versus snowmelt-driven floods (Norway), as well as regions characterized by single/dominant hydroclimates (Germany, the UK, and Norway) versus those with mixed hydroclimates (the US).

2. **Regional Propensity for Heavy-Tailed Flood Behavior:** Germany and the UK are characterized by a propensity for heavy-tailed flood behavior, which is prevalent in these regions. Conversely, a tendency for nonheavy-tailed flood behavior is predominant in Norway under current hydroclimatic conditions, as indicated by the degree of catchment nonlinearity in each region. The US exhibits a mixture of heavy- and nonheavy-tailed behavior. This is likely the results of overarching climatic characteristics, which also shape river network morphology, interacting with diverse regional physioclimatic settings. I emphasize that the relatively more uniform climates in Germany, the UK, and Norway contribute to a dominant presence of heavy or nonheavy-tailed behaviors in these countries, while the US experiences more complex regional patterns due to more diverse hydroclimatic conditions.
3. **Factors Influencing Heavy-Tailed Flood Behavior:** The presence of simultaneous dry periods (defined by uneven rainfall throughout the year) and higher temperatures emerge as the pivotal conditions favoring heavy-tailed flood behavior. Drier catchments alter the runoff generation process, resulting in higher nonlinearity of catchment responses, while higher temperatures elevate evapotranspiration rates, enhancing nonlinearity but also maintaining atmospheric moisture preventing precipitation limitations. The absence of either condition diminishes the prevalence of heavy-tailed flood behavior. More generalized climate categorizations like Arid, Temperate, and Continental exhibit minimal influence on our results.
4. **Seasonality of Flood Tail Behavior:** We contribute to a better understanding of the seasonality of flood tail behavior. Around two-thirds of catchments exhibit consistent behavior across seasons, with the remaining one-third demonstrating seasonality. Heavy-tailed flood behavior is more likely during the growing season (spring to autumn) and diminishes during the dormant season (autumn to winter). These findings hint at the role of temperature-driven evapotranspiration dynamics for the emergence of heavy-tailed flood behavior, which are particularly important in regions which do not experience simultaneous dry conditions and high temperatures.
5. **Influences of Catchment Area on Flood Tail Behavior:** This study elucidates that the impacts of catchment size on flood tail behavior are primarily contingent on the dominant flood generation processes within each region. Specifically, the expansion of catchment area tends to have three distinct effects: (1) It diminishes tail heaviness in regions with moderate nonlinearity, characterized by strong evapotranspiration dynamics and relatively even rainfall throughout the year. This reduction is attributed to the smoothing of evapotranspiration variations. (2) Conversely, in regions with low nonlinearity, characterized by snowfall dynamics, increasing catchment area intensifies tail heaviness. This effect results from the inclusion of diverse flood types, particularly rainfall-driven floods. (3) In regions with high nonlinearity, characterized by simultaneous

4. Geographic Patterns of Heavy-Tailed Floods and Hydroclimatic Interpretation

strong evapotranspiration dynamics and uneven rainfall throughout the year, catchment size expansion appears to have no significant impact on tail heaviness. This lack of effect is likely due to the absence of significant influence on rainfall patterns, which are critical in determining the presence of drier soil conditions.

I propose that a key mechanism driving the emergence of heavy-tailed flood behavior is the temporal variability in catchment storage, primarily induced by simultaneous high evapotranspiration rates and drier soil conditions. This variation in storage can lead to the occurrence of both very small and very large flood events, ultimately resulting in heavy-tailed flood behavior. In contrast, when the catchment remains consistently wet or dry, the magnitude of generated floods tends to fall within a similar range, leading to nonheavy tails in the distribution. It's important to emphasize that this mechanism is influenced by seasonality and catchment sizes, both of which play a role in shaping the variability in catchment storage.

4.7 Supporting information

Table 4.2: Daily Hydrological Data Information \tilde{x} denotes the median value.

Region	Germany	UK	Norway	US
Gauge Number	98	82	82	313
Catchment Size [km^2]	110 – 23,843 (\tilde{x} : 1,195)	15 – 9,948 (\tilde{x} : 283)	4 – 40,504 (\tilde{x} : 234)	66 – 9,935 (\tilde{x} : 1,769)
Streamflow Record Length [year]	35 – 63 (\tilde{x} : 58)	50 – 138 (\tilde{x} : 59)	50 – 148 (\tilde{x} : 96)	24 – 55 (\tilde{x} : 55)
Streamflow Record Duration	1951 – 2013	1883 – 2021	1871 – 2019	1948 – 2002
Number of Case Study	386	325	306	980
(spring/summer autumn/winter)	(97/96 98/95)	(82/81 81/81)	(76/76 76/78)	(285/267 288/140)

4.8 Data availability

I express my gratitude to the following organizations for providing the discharge data: the Bavarian State Office of Environment (LfU) in Germany (<https://www.gkd.bayern.de/de/fluesse/abfluss>), the Global Runoff Data Centre (GRDC) prepared by the Federal Institute for Hydrology (BfG) in the UK and Norway (<http://www.bafg.de/GRDC>), and the National Oceanic and Atmospheric Administration (NOAA) Office of Global Programs (MOPEX) in the US (<http://hydrology.nws.noaa.gov/pub/gcip>). We obtained the digital elevation model from the Shuttle Radar Topography Mission (SRTM) (<http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1>). Köppen climate classification were sourced from the high-resolution present-day Köppen climate map presented by Beck et al. (2018) (<https://doi.org/10.1038/sdata.2018.214>). The dataset of dams used in this study is available from the GeoDAR v.1.0 (<https://doi.org/10.5281/zenodo.6163413>). For characteristics of separated rainfall-runoff events for each streamflow gauge used in the analysis, please refer to Data Set S1 of Tarasova et al., 2018 (<https://doi.org/10.1029/2018WR022588>).

5

Conclusions and Future Avenues

In this section, the overview of the understanding of the emergence of heavy-tailed flood behavior based on the key findings of previous chapters is summarized. Additionally, potential future outlooks, practical applications, and directions for further research stemming from this work are outlined.

5.1 Conclusions

In this dissertation, the fundamental factor contributing to the tail heaviness of flood distributions, which leads to the occurrence of extreme floods, is demonstrated to be the nonlinearity of catchment hydrological responses (chapters 3 and 4). This factor can be influenced by several environmental mechanisms, which can be summarized into three main categories: hydrological processes, climate conditions, and catchment geographical features. The first two categories, hydrological processes and climate conditions, often directly affect catchment nonlinearity, while catchment geographical features are more likely to either enhance or diminish the effects of the first two. Figure 5.1 provides a graphical summary of these factors, with factors related to hydrological processes in green, climate conditions in blue, and catchment geographical features in yellow. These factors (identified in this dissertation) are also placed into primary, secondary, and tertiary groups based on their effects on catchment nonlinearity in Figure 5.1. The interactions between each pair of factors are illustrated by arrows, with red arrows representing a positive effect (enhancement) and blue arrows representing a negative effect (reduction). The light-green-shaded arrow signifies the transition from snow-dominant floods to mixed floods. The main interactions between these factors and the underlying mechanisms are outlined below:

1. Spatially heterogeneous rainfall tends to generate heavy-tailed behavior in flood distributions but only beyond a certain threshold (section 2.4.1). Catchments can exhibit resilience, possibly due to the asymmetry in the distributions of runoff-routing pathways, influenced by the river network structure. More contributions from longer and more distant pathways within river networks (due to the asymmetry distributions)

can mitigate the impacts of rainfall heterogeneity, raising the threshold at which spatial rainfall heterogeneity significantly influences flood tail heaviness.

2. Catchment size and shape are related to catchment resilience via the characteristic distributions of runoff-routing pathways. Larger catchment areas or more circular shapes tend to provide more resilience and better withstand the effects of heterogeneous rainfall on promoting heavier flood tails (sections 2.4.2 and 2.5).
3. On the contrary, we observed that rainfall with strong spatial autocorrelation tends to amplify the impact of rainfall heterogeneity on generating heavier flood tails. This may be because spatially autocorrelated rainfall reduces the likelihood of simultaneous contributions from upstream and downstream river networks, preventing their offsetting effects (which can buffer the spatial heterogeneity of rainfall) of the runoff routing through river networks (sections 2.4.4 and 2.5).
4. Catchments characterized by significant aridity exhibit pronounced nonlinearity in hydrological responses, favoring the emergence of heavy-tailed flood behavior. Such arid conditions are often associated with uneven rainfall patterns throughout the year and elevated evapotranspiration due to high temperatures (sections 4.4.2 and 4.5).
5. Furthermore, regions with higher evapotranspiration during growing seasons tend to display more pronounced catchment nonlinearity, favoring heavy-tailed flood behavior during these periods of the year. This could be attributed to greater wet-dry variations in water storage within catchments, allowing for the occurrence of both very small and very large events, thus promoting catchment nonlinearity. This highlights the significance of considering flood tail heaviness seasonality (sections 4.4.3 and 4.5).
6. Regions primarily characterized by snowmelt-triggered flood generation processes, as opposed to rainfall-triggered events, typically exhibit lower nonlinearity in catchment responses, corresponding to nonheavy-tailed flood behavior (sections 4.4.1 and 4.5).
7. While expanding catchment areas may not directly determine flood tail heaviness, they can alter the characteristic flood generation process, indirectly impacting flood tail heaviness. In this study, regions where flood tail heaviness is mainly influenced by evapotranspiration dynamics tend to show less nonlinearity when catchments expand. This occurs because such expansion often encompasses more diverse terrain with various altitudes, reducing the variation in average evapotranspiration. Conversely, in regions characterized by prevalent snowmelt-triggered flood generation processes, expanding catchments often results in a mix of flood types, particularly rainfall-driven floods in lowland and coastal regions. This amalgamation of flood types tends to increase catchment nonlinearity (sections 4.4.4 and 4.5).

The nonlinearity of catchment responses can be effectively measured using the hydrograph recession exponent, making it a feasible indicator of heavy-tailed flood behavior. This newly proposed index is grounded in classical hydrological processes, and its efficacy is not only theoretically supported (section 3.2) but also empirically validated on an extensive dataset (sections 3.4 and 4.4.1). The index leverages common streamflow dynamics to infer the

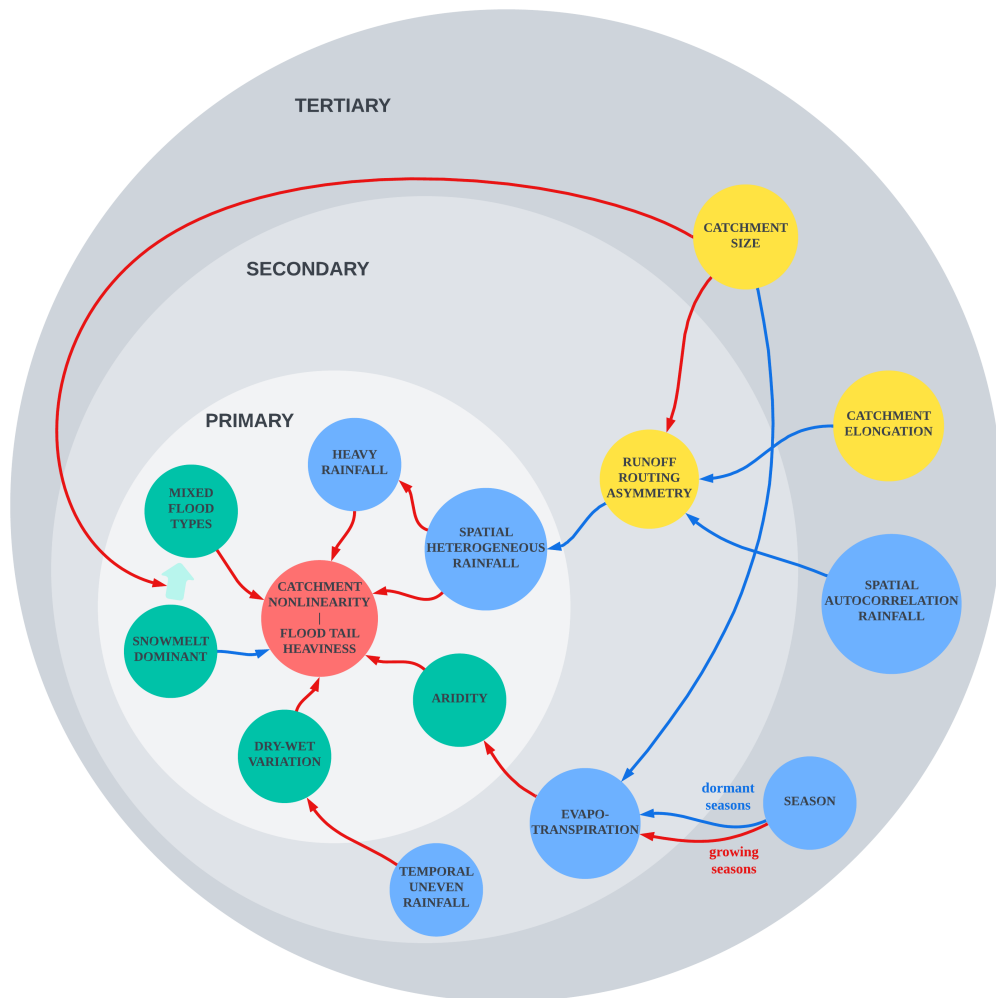


Figure 5.1: Summary of the emergence of flood tail heaviness Catchment nonlinearity is found to be the fundamental factor in emerging flood tail heaviness. Other contributing factors identified in this dissertation are categorized into primary, secondary, and tertiary groups based on their effects on catchment nonlinearity. Green bubbles indicate factors related to hydrological processes, blue bubbles denote those related to climate conditions, and yellow bubbles signify those associated with geographical characteristics. Red arrows represent a positive effect (enhancement), while blue arrows represent a negative effect (reduction). The light-green-shaded arrow signifies the transition from snow-dominant floods to mixed floods.

emergence of heavy-tailed flood behavior, providing robust estimations across datasets with varying record lengths. Notably, it excels in datasets with shorter record lengths, highlighting its utility in flood hazard research across a wide range of applications (sections 3.4 and 3.5; figure 3.3). The findings of this study contribute to comprehending the mechanisms behind heavy-tailed flood behavior. Moreover, a physically-based robust approach is introduced for assessing flood tail heaviness, thereby mitigating the sensitivity to limited data, which is a common challenge in such analyses, and broadening the applicability across regions. These outcomes pave the way for numerous prospective applications in the realm of hazard research and engineering, which are exemplified in the subsequent subsection.

5.2 Future avenues

Application in flood frequency guideline

One direct application of this work is the selection of appropriate candidate probability distributions for flood frequency analysis. This selection is critical for accurately determining the return periods required for effective flood risk management (Nerantzaki and Papalexiou, 2022). Misinterpreting the tail behavior of the underlying distribution in flood analysis can result in significant underestimation of extreme events and lead to severe hazards (Merz et al., 2021; Merz et al., 2022). It is advisable to employ methods that pre-estimate the tail behavior and carefully select suitable candidate probability distributions for flood frequency analysis (DWA, 2012; Mushtaq et al., 2022). The outcomes of this research, such as accurately identifying and potentially predicting heavy-tailed flood behavior, allow this application.

Assessing the climate-change related impact

The intensification of extreme flood events resulting from climate change (Tabari, 2020) underscores the need for an accurate assessment of evolving flood likelihood. These assessments play a crucial role in informing flood risk management and decision-making processes (Arnell and Gosling, 2016). However, traditional methods for estimating the tail heaviness of flood distributions rely on long-term data records (Cai and Hames, 2010; Németh et al., 2019; Wietzke et al., 2020), which may not be readily available or sensitive enough for climate change studies (Lins, 2008). This limitation arises due to the significant global warming resulting from human alteration of the planet's energy budget since the 1970s (Sarkar and Maity, 2021). As a result, the sensitivity of traditional methods in assessing changing flood behavior is restricted by this relatively short historical timeframe.

This work presents an innovative approach that leverages readily available short-term data recordings, such as common streamflow dynamics, for the assessment of flood tail heaviness (Wang et al., 2023a). This method enables the sensing of the impacts of changing conditions and allows us to effectively tackle this issue. Figure 5.2 illustrates an increasing tendency of hydrograph recession exponents based on 75 years of recent data from case studies in Norway. Notably, the visual representation shows a more pronounced increase since the late 1970s to 1990s, particularly, for the cases in winter. These results provide a consistent indication that as global warming gradually alters the dominant runoff generation from snow dynamics to a mixture type (e.g., Tarasova et al., 2023) in Norway, the corresponding nonlinearity of catchment response changes, potentially exacerbating the risk of extreme floods. Given that the changes are recognized via such an approach, it could offer valuable insights into estimating flood tail behavior in the context of climate change.

Flood hazard classification and quantitative flood hazard assessment

Quantitative flood hazard assessment involves analyzing spatial flood indicators, classifying hazard levels based on these analyses, and creating flood risk maps (Maranzoni et al., 2023). However, in the current state of research, a standardized flood hazard classification is highly desirable due to the variability in flood scales and data availability worldwide (de Moel et al.,

2009; van Alphen et al., 2009; Maranzoni et al., 2023). The tail heaviness of flood distributions can serve as an objective benchmark of the flood extremeness using for hazard level classification. Given that the tail heaviness can be reliably assessed by hydrograph recession analysis across regions with diverse physiographical conditions and data record lengths, it serves as an useful tool and allows for further applications in flood hazard assessment systems.

Understanding the impact of human activities

Assessing catchment nonlinearity helps uncover the relationship between extreme floods and various factors, including those influenced by human activities. Scenarios like reservoir or dam construction (e.g., Volpi et al., 2018), driven by water resource shortages or urban development, demand an understanding of their impact on flood behavior. However, the operation regulations of such structures can vary widely based on regional management and their intended purposes, making it challenging to provide a general assessment of their effects on hydrological responses. Analyzing common streamflow dynamics and catchment nonlinearity allows for a comprehensive evaluation of these activities' overall impact and their influence on the likelihood of extreme flood occurrences.

Applications in research of environmental sustainability

Streamflow dynamics influence various environmental variables such as suspended sediment, nutrients, driftwood, and macro plastic through river transport and export. These variables can lead to sustainability issues. For instance, high suspended sediment concentrations can adversely affect both human livelihoods, straining drinking water purification systems (Braga and Filion, 2022), and freshwater aquatic ecosystems (Kjelland et al., 2015). Changing the magnitude and frequency of floods can also lead to a change in soil erosion rate, which is substantially influenced by flood dynamics (Le Gall et al., 2017). The dynamics of nutrient storage and transport in catchments can be affected by varying frequency of flood and drought (Johnston and Maher, 2022). Examining the relationship between hydrological response nonlinearity and the dynamics of these variables reveals the potential impacts of heavy-tailed flood behavior on environmental sustainability and allows for related research applications.

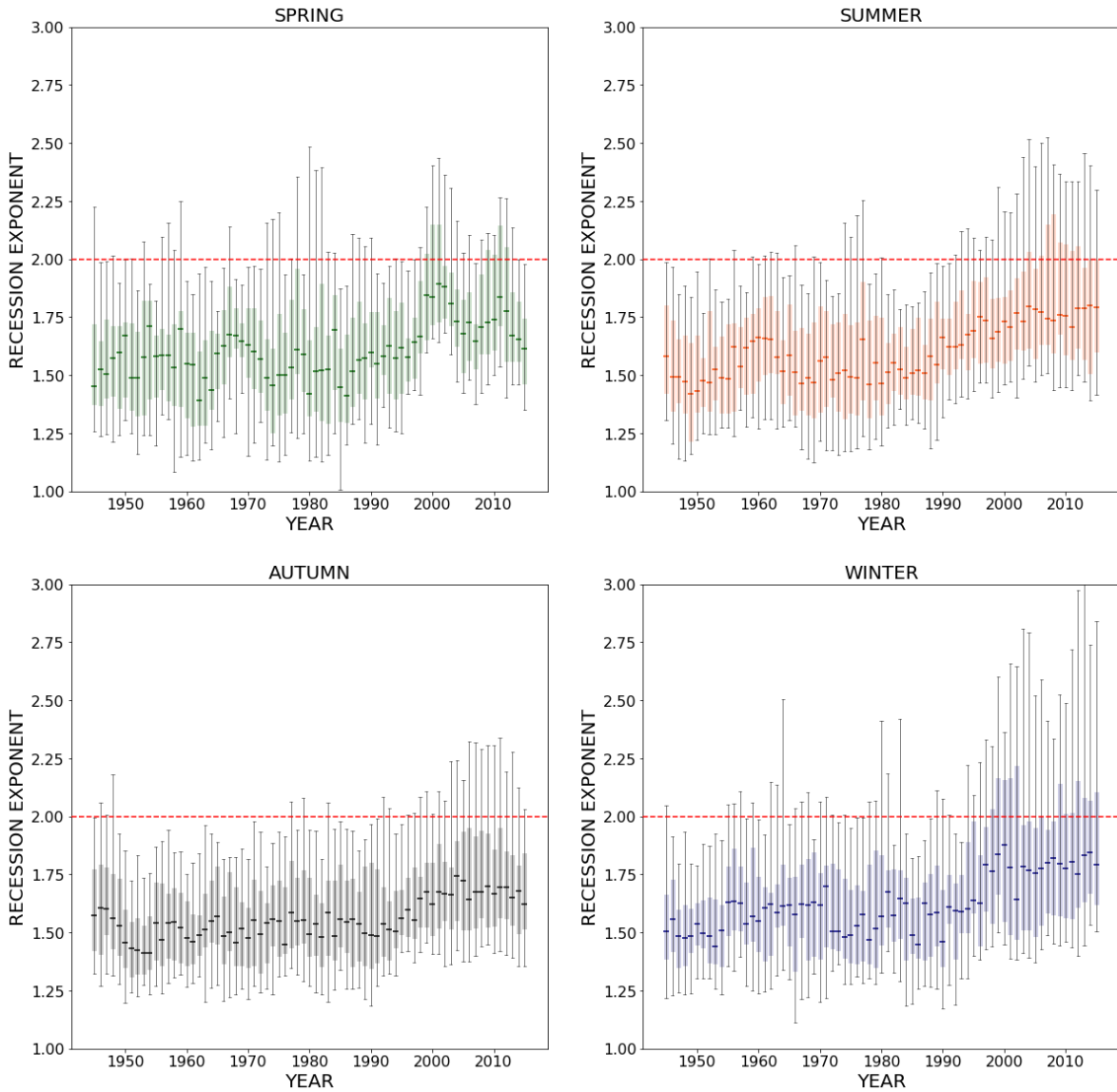


Figure 5.2: Streamflow recession patterns in recent decades in Norway A sliding window approach is utilized to estimate the hydrograph recession exponent for each specific year, considering streamflow data from the preceding five years. The dot, box, and whisker represent the 50th, 35th65th, and 20th80th percentiles, respectively, of the studied catchments. The red dashed line indicate the critical value of the assessment of the heavy-tailed behavior. The analysis is based on 49 catchments across Norway using daily streamflow data from 1940 to 2015.

References

- Ács, F., Horváth, Á., Breuer, H., & Rubel, F. (2010). Effect of soil hydraulic parameters on the local convective precipitation. *Meteorologische Zeitschrift*, *19*(2), 143–153.
<https://doi.org/10.1127/0941-2948/2010/0435>
- Adams, R., Western, A. W., & Seed, A. W. (2012). An analysis of the impact of spatial variability in rainfall on runoff and sediment predictions from a distributed model. *Hydrological Processes*, *26*(21), 3263–3280. <https://doi.org/10.1002/hyp.8435>
- Anderson, P. L., & Meerschaert, M. (1998). Modeling river flows with heavy tails. *Water Resources Research*, *34*(9), 2271–2280.
<https://doi.org/http://doi.org/10.1029/98WR01449>
- Arai, R., Toyoda, Y., & Kazama, S. (2020). Runoff recession features in an analytical probabilistic streamflow model. *Journal of Hydrology*, *597*, 125745.
<https://doi.org/10.1016/j.jhydrol.2020.125745>
- Archfield, S. A., Hirsch, R. M., Viglione, A., & Blöschl, G. (2016). Fragmented patterns of flood change across the United States. *Geophysical Research Letters*, *43*(10), 10232–10239. <https://doi.org/10.1002/2016GL070590>
- Arnell, N. W., & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387–401. <https://doi.org/10.1007/s10584-014-1084-5>
- Assani, A. A., Stichelbout, É., Roy, A. G., & Petit, F. (2006). Comparison of impacts of dams on the annual maximum flow characteristics in three regulated hydrologic regimes in Québec (Canada). *Hydrological Processes*, *20*(16), 3485–3501.
<https://doi.org/10.1002/hyp.6150>
- Ayalew, T. B., Krajewski, W. F., & Mantilla, R. (2014). Connecting the power-law scaling structure of peak-discharges to spatially variable rainfall and catchment physical properties. *Advances in Water Resources*, *71*, 32–43.
<https://doi.org/10.1016/j.advwatres.2014.05.009>

REFERENCES

- Bacchi, B., & Kottegoda, N. T. (1995). Identification and calibration of spatial correlation patterns of rainfall. *Journal of Hydrology*, *165*(1–4), 311–348.
[https://doi.org/10.1016/0022-1694\(94\)02590-8](https://doi.org/10.1016/0022-1694(94)02590-8)
- Basso, S., Botter, G., Merz, R., & Miniussi, A. (2021). PHEV! The PHysically-based Extreme Value distribution of river flows. *Environmental Research Letters*, *16*(12).
<https://doi.org/10.1088/1748-9326/ac3d59>
- Basso, S., Schirmer, M., & Botter, G. (2015). On the emergence of heavy-tailed streamflow distributions. *Advances in Water Resources*, *82*, 98–105.
<https://doi.org/10.1016/j.advwatres.2015.04.013>
- Basso, S., Schirmer, M., & Botter, G. (2016). A physically based analytical model of flood frequency curves. *Geophysical Research Letters*, *43*(17), 9070–9076.
<https://doi.org/10.1002/2016GL069915>
- Basso, S., Merz, R., Tarasova, L., & Miniussi, A. (2023). Extreme flooding controlled by stream network organization and flow regime. *Nature Geoscience*, *16*(April), 339–343.
<https://doi.org/10.1038/s41561-023-01155-w>
- Bates, B. C., & Aryal, S. K. (2014). A similarity index for storm runoff due to saturation excess overland flow. *Journal of Hydrology*, *513*, 241–255.
<https://doi.org/10.1016/j.jhydrol.2014.03.021>
- Baudin, M., Dutfoy, A., Iooss, B., & Popelin, A.-L. (2017). OpenTURNS: An Industrial Software for Uncertainty Quantification in Simulation BT - Handbook of Uncertainty Quantification. In R. Ghanem, D. Higdon, & H. Owhadi (Eds.) (pp. 2001–2038). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-12385-1_64
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 180214. <https://doi.org/10.1038/sdata.2018.214>
- Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., De Waal, D., & Ferro, C. (2004). *Statistics of extremes: Theory and applications*. Wiley.
<https://doi.org/https://doi.org/10.1002/0470012382>
- Berghuijs, W. R., Sivapalan, M., Woods, R. A., & Savenije, H. H. G. (2014). Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales. *Water Resources Research*, *50*(7), 5638–5661.
<https://doi.org/https://doi.org/10.1002/2014WR015692>

- Bernardara, P., Schertzer, D., Sauquet, E., Tchiguirinskaia, I., & Lang, M. (2008). The flood probability distribution tail: How heavy is it? *Stochastic Environmental Research and Risk Assessment*, 22(1), 107–122. <https://doi.org/10.1007/s00477-006-0101-2>
- Beurton, S., & Thielen, A. H. (2009). Seasonality of floods in Germany. *Hydrological Sciences Journal*, 54(1), 62–76. <https://doi.org/10.1623/hysj.54.1.62>
- Bevere, L., & Remondi, F. (2022). *Natural catastrophes in 2021: the floodgates are open. Swiss Re Institute sigma research.*
- Bezak, N., Brilly, M., & Šraj, M. (2014). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, 59(5), 959–977. <https://doi.org/10.1080/02626667.2013.831174>
- Biswal, B. (2021). Decorrelation is not dissociation: There is no means to entirely decouple the Brutsaert-Nieber parameters in streamflow recession analysis. *Advances in Water Resources*, 147, 103822. <https://doi.org/https://doi.org/10.1016/j.advwatres.2020.103822>
- Biswal, B., & Kumar, D. N. (2014). Study of dynamic behaviour of recession curves. *Hydrological Processes*, 792(November 2012), 784–792. <https://doi.org/10.1002/hyp.9604>
- Biswal, B., & Marani, M. (2010). Geomorphological origin of recession curves. *Geophysical Research Letters*, 37(24), 1–5. <https://doi.org/10.1029/2010GL045415>
- Borga, M., Boscolo, P., Zanon, F., & Sangati, M. (2007). Hydrometeorological analysis of the 29 August 2003 flash flood in the eastern Italian Alps. *Journal of Hydrometeorology*, 8(5), 1049–1067. <https://doi.org/10.1175/JHM593.1>
- Botter, G. (2010). Stochastic recession rates and the probabilistic structure of stream flows. *Water Resources Research*, 46(12). <https://doi.org/10.1029/2010WR009217>
- Botter, G., Basso, S., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). Natural streamflow regime alterations: Damming of the Piave river basin (Italy). *Water Resources Research*, 46(6), 1–14. <https://doi.org/10.1029/2009WR008523>
- Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007a). Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. climate regimes. *Water Resources Research*, 43(11), 1–10. <https://doi.org/10.1029/2007WR006162>

REFERENCES

- Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007b). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. *Water Resources Research*, *43*(2), 1–14. <https://doi.org/10.1029/2006WR005043>
- Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2009). Nonlinear storage-discharge relations and catchment streamflow regimes. *Water Resources Research*, *45*(10), 1–16. <https://doi.org/10.1029/2008WR007658>
- Botter, G., & Rinaldo, A. (2003). Scale effect on geomorphologic and kinematic dispersion. *Water Resources Research*, *39*(10), 1–10. <https://doi.org/10.1029/2003WR002154>
- Bowers, M. C., Tung, W. W., & Gao, J. B. (2012). On the distributions of seasonal river flows: Lognormal or power law? *Water Resources Research*, *48*(5), 1–12. <https://doi.org/10.1029/2011WR011308>
- Boyd, M. J. (1978). A storage-routing model relating drainage basin hydrology and geomorphology. *Water Resources Research*, *14*(5), 921–928.
- Bracken, L., Cox, N., & Shannon, J. (2008). The relationship between rainfall inputs and flood generation in south-east Spain. *Hydrological Processes*, *22*(5). <https://doi.org/10.1002/hyp.6641>
- Braga, A. S., & Filion, Y. (2022). The interplay of suspended sediment concentration, particle size and fluid velocity on the rapid deposition of suspended iron oxide particles in PVC drinking water pipes. *Water Research X*, *15*(January), 100143. <https://doi.org/10.1016/j.wroa.2022.100143>
- Brath, A., Montanari, A., & Toth, E. (2004). Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. *Journal of Hydrology*, *291*(3–4), 232–253. <https://doi.org/10.1016/j.jhydrol.2003.12.044>
- Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research*, *13*(3), 637–643. <https://doi.org/10.1029/WR013i003p00637>
- Cai, Y., & Hames, D. (2010). Minimum sample size determination for generalized extreme value distribution. *Communications in Statistics: Simulation and Computation*, *40*(1), 87–98. <https://doi.org/10.1080/03610918.2010.530368>

- Ceola, S., Botter, G., Bertuzzo, E., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). Comparative study of ecohydrological streamflow probability distributions. *Water Resources Research*, 46(9), 1–12. <https://doi.org/10.1029/2010WR009102>
- Cerdan, O., Le Bissonnais, Y., Govers, G., Lecomte, V., Van Oost, K., Couturier, A., et al. (2004). Scale effect on runoff from experimental plots to catchments in agricultural areas in Normandy. *Journal of Hydrology*, 299(1–2), 4–14. <https://doi.org/10.1016/j.jhydrol.2004.02.017>
- Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, 14(4), 601–604.
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703. <https://doi.org/10.1137/070710111>
- Collischonn, W., Fleischmann, A., Paiva, R. C. D., & Mejia, A. (2017). Hydraulic Causes for Basin Hydrograph Skewness Walter. *Water Resources Research*, 53(10), 603–618. <https://doi.org/10.1002/2017WR021543>
- Cooke, R. M., Nieboer, D., & Misiewicz, J. (2014). *Fat-Tailed Distributions: Data, Diagnostics and Dependence* (volume 1). John Wiley & Sons.
- Cooke, R. M., & Nieboer, D. (2011). Heavy-Tailed Distributions: Data, Diagnostics, and New Developments. *Resources for the Future Discussion Paper, No. 11-19*. <https://doi.org/dx.doi.org/10.2139/ssrn.1811043>
- Cunderlik, J. M., & Burn, D. H. (2002). The use of flood regime information in regional flood frequency analysis. *Hydrological Sciences Journal*, 47(1), 77–92. <https://doi.org/10.1080/02626660209492909>
- Darras, T., Borrell Estupina, V., Kong-A-Siou, L., Vayssade, B., Johannet, A., & Pistre, S. (2015). Identification of spatial and temporal contributions of rainfalls to flash floods using neural network modelling: Case study on the Lez basin (southern France). *Hydrology and Earth System Sciences*, 19(10), 4397–4410. <https://doi.org/10.5194/hess-19-4397-2015>
- De Moel, H., Van Alphen, J., & Aerts, J. C. J. H. (2009). Flood maps in Europe - Methods, availability and use. *Natural Hazards and Earth System Science*, 9(2), 289–301. <https://doi.org/10.5194/nhess-9-289-2009>
- Derrick, B., Toher, D., & White, P. (2016). Why Welchs test is Type I error robust. *The Quantitative Methods for Psychology*, 12(1), 30–38. <https://doi.org/doi:10.20982/tqmp.12.1.p030>

REFERENCES

- Dingman, S. L. (2002). Physical hydrology (p. 646). Upper Saddle River, New Jersey: Prentice Hall.
- Diop, A., & Lô, G. S. (2009). Ratio of Generalized Hill's estimator and its asymptotic normality theory. *Mathematical Methods of Statistics*, 18(2), 117–133.
<https://doi.org/10.3103/S1066530709020021>
- D'Odorico, P., & Rigon, R. (2003). Hillslope and channel contributions to the hydrologic response. *Water Resources Research*, 39(5), 1–9.
<https://doi.org/10.1029/2002WR001708>
- Donat, M. G., Angélic, O., & Ukkola, A. M. (2019). Intensification of precipitation extremes in the world's humid and water-limited regions. *Environmental Research Letters*, 14(6).
<https://doi.org/10.1088/1748-9326/ab1c8e>
- Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. *Nature Climate Change*, 6(5), 508–513. <https://doi.org/10.1038/nclimate2941>
- Doulatyari, B., Betterle, A., Basso, S., Biswal, B., Schirmer, M., & Botter, G. (2015). Predicting streamflow distributions and flow duration curves from landscape and climate. *Advances in Water Resources*, 83, 285–298.
<https://doi.org/10.1016/j.advwatres.2015.06.013>
- Dunne, T., & Black, R. D. (1970). Partial Area to Storm Runoff in a Small New England Watershed Dense Forest. *Water Resources Research*, 6(5), 1296–1311. Retrieved from <http://soilandwater.bee.cornell.edu/Research/VSA/papers/DunneWRR70.pdf>
- Durrans, S. R., Eiffe, M. A., Thomas, W. O., & Goranflo, H. M. (2003). Joint Seasonal /Annual Flood Frequency Analysis. *Journal of Hydrologic Engineering*, 8(4), 181–189.
[https://doi.org/10.1061/\(asce\)1084-0699\(2003\)8:4\(181\)](https://doi.org/10.1061/(asce)1084-0699(2003)8:4(181))
- DWA. (2012). Merkblatt DWA-M 552 - Ermittlung von Hochwasserwahrscheinlichkeiten. Hennef: Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V. (DWA).
- El Adlouni, S., Bobée, B., & Ouarda, T. B. M. J. (2008). On the tails of extreme event distributions in hydrology. *Journal of Hydrology*, 355(1–4), 16–33.
<https://doi.org/10.1016/j.jhydrol.2008.02.011>
- Eliazar, I., & Sokolov, I. (2010). Gini characterization of extreme-value statistics. *Physica A - Statistical Mechanics and Its Applications - PHYSICA A*, 389, 4462–4472.
<https://doi.org/10.1016/j.physa.2010.07.005>

- Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extreme events for insurance and finance*. Springer Berlin Heidelberg.
- Engeland, K., Hisdal, H., & Frigessi, A. (2004). Practical extreme value modelling of hydrological floods and droughts: A case study. *Extremes*, 7(1), 5–30. <https://doi.org/10.1007/s10687-004-4727-5>
- ESRI. (2009). Arc Hydro Tools - Tutorial. *New York*, (January), 136.
- Fatichi, S., Ivanov, V. Y., Paschalis, A., Peleg, N., Molnar, P., Rimkus, S., et al. (2016). Uncertainty partition challenges the predictability of vital details of climate change. *Earth's Future*, 4(5), 240–251. <https://doi.org/10.1002/2015EF000336>
- Felgueiras, M. M. (2012). Explaining the seismic moment of large earthquakes by heavy and extremely heavy tailed models. *GEM - International Journal on Geomathematics*, 3(2), 209–222. <https://doi.org/10.1007/s13137-012-0042-5>
- Fiorentino, M., Manfreda, S., & Iacobellis, V. (2007). Peak runoff contributing area as hydrological signature of the probability distribution of floods. *Advances in Water Resources*, 30(10), 2123–2134. <https://doi.org/10.1016/j.advwatres.2006.11.017>
- Fischer, S., & Schumann, A. (2016). Robust flood statistics: comparison of peak over threshold approaches based on monthly maxima and TL-moments. *Hydrological Sciences Journal*, 61(3), 457–470. <https://doi.org/10.1080/02626667.2015.1054391>
- Gaume, E. (2006). On the asymptotic behavior of flood peak distributions. *Hydrology and Earth System Sciences*, 10(2), 233–243. <https://doi.org/10.5194/hess-10-233-2006>
- Gioia, A., Iacobellis, V., Manfreda, S., & Fiorentino, M. (2012). Influence of infiltration and soil storage capacity on the skewness of the annual maximum flood peaks in a theoretically derived distribution. *Hydrology and Earth System Sciences*, (16), 937–951. <https://doi.org/10.5194/hess-16-937-2012>
- Gioia, A., Iacobellis, V., Manfreda, S., & Fiorentino, M. (2008). Runoff thresholds in derived flood frequency distributions. *Hydrology and Earth System Sciences*, 12(6), 1295–1307. <https://doi.org/10.5194/hess-12-1295-2008>
- Godrèche, C., Majumdar, S. N., & Schehr, G. (2015). Statistics of the longest interval in renewal processes. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(3). <https://doi.org/10.1088/1742-5468/2015/03/P03014>

REFERENCES

- Grillenzoni, C. (2022). Spatial Autocorrelation Function of a Process in the Plane. Retrieved from <https://www.mathworks.com/matlabcentral/fileexchange/113695-spatial-autocorrelation-function-of-a-process-in-the-plane>
- Gumbel, E. J. (1958). *Statistics of Extremes*. Columbia University Press.
<https://doi.org/doi:10.7312/gumb92958>
- Guo, J., Li, H.-Y., Leung, L. R., Guo, S., Liu, P., & Sivapalan, M. (2014). Links between flood frequency and annual water balance behaviors: A basis for similarity and regionalization. *Water Resources Research*, *50*, 937–953.
<https://doi.org/http://dx.doi.org/10.1002/2013WR014374>
- Gupta, N., & Chavan, S. R. (2023). Investigating the tail behaviour and associated risk with daily discharges in South Indian Rivers. *Stochastic Environmental Research and Risk Assessment*, *0123456789*. <https://doi.org/10.1007/s00477-023-02453-w>
- Gupta, V. K., & Waymire, E. C. (1998). Spatial Variability and Scale Invariance in Hydrologic Regionalization. In G. Sposito (Ed.), *Scale Dependence and Scale Invariance in Hydrology* (pp. 88–135). Cambridge University Press.
<https://doi.org/https://doi.org/10.1017/CBO9780511551864.005>
- Haddad, K., & Rahman, A. (2011). Selection of the best fit flood frequency distribution and parameter estimation procedure: A case study for Tasmania in Australia. *Stochastic Environmental Research and Risk Assessment*, *25*(3), 415–428.
<https://doi.org/10.1007/s00477-010-0412-1>
- Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., et al. (2014). Understanding flood regime changes in Europe: A state-of-the-art assessment. *Hydrology and Earth System Sciences*, *18*(7), 2735–2772. <https://doi.org/10.5194/hess-18-2735-2014>
- Harman, C. J., Sivapalan, M., & Kumar, P. (2009). Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research*, *45*(9), 1–13. <https://doi.org/10.1029/2008WR007392>
- Hodgkins, G. A., Whitfield, P. H., Burn, D. H., Hannaford, J., Renard, B., Stahl, K., et al. (2017). Climate-driven variability in the occurrence of major floods across North America and Europe. *Journal of Hydrology*, *552*, 704–717.
<https://doi.org/10.1016/j.jhydrol.2017.07.027>

- Horton, R. E. (1932). Drainage-basin characteristics. *Transactions of the American Geophysical Union*, 13(1), 350–361. <https://doi.org/10.1029/TR013i001p00350>
- Houska, T., Kraft, P., Chamorro-Chavez, A., & Breuer, L. (2015). SPOTting model parameters using a ready-made python package. *PLoS ONE*, 10(12), 1–22. <https://doi.org/10.1371/journal.pone.0145180>
- Hu, L., Nikolopoulos, E. I., Marra, F., & N., A. E. (2023). Toward an improved estimation of flood frequency statistics from simulated flows. *Journal of Flood Risk Management*, 1–13. <https://doi.org/10.1111/jfr3.12891>
- Huntingford, C., Marsh, T., Scaife, A. A., Kendon, E. J., Hannaford, J., Kay, A. L., et al. (2014). Potential influences on the United Kingdom's floods of winter 2013/14. *Nature Climate Change*, 4(9), 769–777. <https://doi.org/10.1038/nclimate2314>
- Ibragimov, M., Ibragimov, R., & Walden, J. (2015). *Heavy-Tailed Distributions and Robustness in Economics and Finance*. Springer Cham. <https://doi.org/10.1007/978-3-319-16877-7>
- Jachens, E. R., Rupp, D. E., Roques, C., & Selker, J. S. (2020). Recession analysis revisited: Impacts of climate on parameter estimation. *Hydrology and Earth System Sciences*, 24(3), 1159–1170. <https://doi.org/10.5194/hess-24-1159-2020>
- Joel, A., Messing, I., Seguel, O., & Casanova, M. (2002). Measurement of surface water runoff from plots of two different sizes. *Hydrological Processes*, 16(7), 1467–1478. <https://doi.org/10.1002/hyp.356>
- Johnston, S. G., & Maher, D. T. (2022). Drought, megafires and flood - climate extreme impacts on catchment-scale river water quality on Australia's east coast. *Water Research*, 218(March), 118510. <https://doi.org/10.1016/j.watres.2022.118510>
- Karlsen, R. H., Bishop, K., Grabs, T., Ottosson-Löfvenius, M., Laudon, H., & Seibert, J. (2019). The role of landscape properties, storage and evapotranspiration on variability in streamflow recessions in a boreal catchment. *Journal of Hydrology*, 570(2019), 315–328. <https://doi.org/10.1016/j.jhydrol.2018.12.065>
- Katz, R. W. (2001). Do Weather or Climate Variables and Their Impacts Have Heavy-Tailed Distributions? *16th Conf. on Probability and Statistics in the Atmospheric Sciences*, Orlando, FL, Amer. Meteor. Soc., J3.5. Available online at https://ams.confex.com/ams/annual2002/techprogram/paper_26949.htm.

REFERENCES

- Katz, R. W., & Brown, B. G. (1992). Extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, *21*(3), 289–302. <https://doi.org/10.1007/BF00139728>
- Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8–12), 1287–1304. [https://doi.org/10.1016/S0309-1708\(02\)00056-8](https://doi.org/10.1016/S0309-1708(02)00056-8)
- Kay, A. L., Reynard, N. S., & Jones, R. G. (2006). RCM rainfall for UK flood frequency estimation. I. Method and validation. *Journal of Hydrology*, *318*(1–4), 151–162. <https://doi.org/10.1016/j.jhydrol.2005.06.012>
- Kim, B. S., Kim, B. K., & Kim, Soo, H. (2008). Flood simulation using the gauge-adjusted radar rainfall and physics-based distributed hydrologic model. *Hydrological Processes*, *22*, 4400–4414. <https://doi.org/10.1002/hyp.7043>
- Kirkby, M. (1975). Hydrograph modelling strategies, in *Processes in Human and Physical Geography*. In R. Peel, M. Chisholm, & P. Haggett (Eds.) (pp. 69–90). London: Heinemann.
- Kjeldsen, T. R., Macdonald, N., Lang, M., Mediero, L., Albuquerque, T., Bogdanowicz, E., et al. (2014). Documentary evidence of past floods in Europe and their utility in flood frequency estimation. *Journal of Hydrology*, *517*, 963–973. <https://doi.org/10.1016/j.jhydrol.2014.06.038>
- Kjelland, M. E., Woodley, C. M., Swannack, T. M., & Smith, D. L. (2015). A review of the potential effects of suspended sediment on fishes: potential dredging-related physiological, behavioral, and transgenerational implications. *Environment Systems and Decisions*, *35*(3), 334–350. <https://doi.org/10.1007/s10669-015-9557-2>
- Kling, H., & Gupta, H. (2009). On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability. *Journal of Hydrology*, *373*(3–4), 337–351. <https://doi.org/10.1016/j.jhydrol.2009.04.031>
- Kondor, D., Pósfai, M., Csabai, I., & Vattay, G. (2014). Do the rich get richer? An empirical analysis of the Bitcoin transaction network. *PLoS ONE*, *9*(2). <https://doi.org/10.1371/journal.pone.0086197>

- Koutsoyiannis, D. (2022). *Stochastics of Hydroclimatic Extremes – A Cool Look at Risk. Kallipos* (2nd ed.). Athens: Open Academic Editions. <https://doi.org/10.57713/kallipos-1>
- Koutsoyiannis, D. (2004a). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrological Sciences Journal*, 49(4), 591–610. <https://doi.org/10.1623/hysj.49.4.591.54424>
- Koutsoyiannis, D. (2004b). Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrological Sciences Journal*, 49(4), 575–590. <https://doi.org/10.1623/hysj.49.4.575.54430>
- Kumar, M., Sharif, M., & Ahmed, S. (2020). Flood estimation at Hathnikund Barrage, River Yamuna, India using the Peak-Over-Threshold method. *ISH Journal of Hydraulic Engineering*, 26(3), 291–300. <https://doi.org/10.1080/09715010.2018.1485119>
- Laio, F., Porporato, A., Fernandez-Illescas, C. P., & Rodriguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress IV. Discussion of real cases. *Advances in Water Resources*, 24(7), 745–762. [https://doi.org/10.1016/S0309-1708\(01\)00007-0](https://doi.org/10.1016/S0309-1708(01)00007-0)
- Lang, M., Ouarda, T. B. M. J., & Bobée, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3–4), 103–117. [https://doi.org/10.1016/S0022-1694\(99\)00167-5](https://doi.org/10.1016/S0022-1694(99)00167-5)
- Le Gall, M., Evrard, O., Foucher, A., Laceby, J. P., Salvador-Blanes, S., Manière, L., et al. (2017). Investigating the temporal dynamics of suspended sediment during flood events with ^{7}Be and ^{210}Pb xs measurements in a drained lowland catchment. *Scientific Reports*, 7(September 2016), 1–10. <https://doi.org/10.1038/srep42099>
- Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., et al. (2011). High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, 9(9), 494–502. <https://doi.org/10.1890/100125>
- Li, B., Li, G., & Luo, J. (2021). *Latent but not absent: The “long tail” nature of rural special education and its dynamic correction mechanism. PLoS ONE* (Vol. 16). <https://doi.org/10.1371/journal.pone.0242023>

REFERENCES

- Li, C., Zwiers, F., Zhang, X., Chen, G., Lu, J., Li, G., et al. (2019). Larger Increases in More Extreme Local Precipitation Events as Climate Warms. *Geophysical Research Letters*, *46*(12), 6885–6891. <https://doi.org/10.1029/2019GL082908>
- Li, H., & Sivapalan, M. (2011). Effect of spatial heterogeneity of runoff generation mechanisms on the scaling behavior of event runoff responses in a natural river basin. *Water Resources Research*, *47*(5), 1–20. <https://doi.org/10.1029/2010WR009712>
- Li, L., Zou, Y., Li, Y., Lin, H., Liu, D. L., Wang, B., et al. (2020). Trends, change points and spatial variability in extreme precipitation events from 1961 to 2017 in China. *Hydrology Research*, *51*(3), 484–504. <https://doi.org/10.2166/nh.2020.095>
- Lins, H. F. (2008). Challenges to hydrological observations. *WMO Bulletin*, *57*(January), 55–58.
- Lombardo, F., Napolitano, F., Russo, F., & Koutsoyiannis, D. (2019). On the Exact Distribution of Correlated Extremes in Hydrology. *Water Resources Research*, *55*(12), 10405–10423. <https://doi.org/10.1029/2019WR025547>
- Looper, J. P., & Vieux, B. E. (2012). An assessment of distributed flash flood forecasting accuracy using radar and rain gauge input for a physics-based distributed hydrologic model. *Journal of Hydrology*, *412–413*, 114–132. <https://doi.org/10.1016/j.jhydrol.2011.05.046>
- Lu, P., Smith, J. A., & Lin, N. (2017). Spatial characterization of flood magnitudes over the drainage network of the Delaware river basin. *Journal of Hydrometeorology*, *18*(4), 957–976. <https://doi.org/10.1175/JHM-D-16-0071.1>
- Macdonald, E., Merz, B., Guse, B., Wietzke, L., Ullrich, S., Kemter, M., et al. (2022). Event and Catchment Controls of Heavy Tail Behavior of Floods. *Water Resources Research*, *58*(6), 1–25. <https://doi.org/10.1029/2021wr031260>
- Malamud, B. D. (2004). Tails of natural hazards. *Physics World*, *17*(8), 31–35. <https://doi.org/10.1088/2058-7058/17/8/35>
- Malamud, B. D., & Turcotte, D. L. (2006). The applicability of power-law frequency statistics to floods. *Journal of Hydrology*, *322*(1–4), 168–180. <https://doi.org/10.1016/j.jhydrol.2005.02.032>

- Marani, M., & Ignaccolo, M. (2015). A metastatistical approach to rainfall extremes. *Advances in Water Resources*, *79*, 121–126. <https://doi.org/10.1016/j.advwatres.2015.03.001>
- Maranzoni, A., D’Oria, M., & Rizzo, C. (2023). Quantitative flood hazard assessment methods: A review. *Journal of Flood Risk Management*, *16*(1), 1–31. <https://doi.org/10.1111/jfr3.12855>
- Marchi, L., Borga, M., Preciso, E., & Gaume, E. (2010). Characterisation of selected extreme flash floods in Europe and implications for flood risk management. *Journal of Hydrology*, *394*(1–2), 118–133. <https://doi.org/10.1016/j.jhydrol.2010.07.017>
- Marra, F., Nikolopoulos, E. I., Anagnostou, E. N., & Morin, E. (2018). Metastatistical Extreme Value analysis of hourly rainfall from short records: Estimation of high quantiles and impact of measurement errors. *Advances in Water Resources*, *117*, 27–39. <https://doi.org/10.1016/j.advwatres.2018.05.001>
- Martinez-Villalobos, C., & Neelin, J. D. (2021). Climate models capture key features of extreme precipitation probabilities across regions. *Environmental Research Letters*, *16*(2). <https://doi.org/10.1088/1748-9326/abd351>
- McCuen, R. H., & Smith, E. (2008). Origin of Flood Skew. *Journal of Hydrologic Engineering*, *13*(9), 771–775. [https://doi.org/10.1061/\(asce\)1084-0699\(2008\)13:9\(771\)](https://doi.org/10.1061/(asce)1084-0699(2008)13:9(771))
- McDermott, T. K. J. (2022). Global exposure to flood risk and poverty. *Nature Communications*, *13*(1), 6–8. <https://doi.org/10.1038/s41467-022-30725-6>
- Mei, Y., Anagnostou, E. N., Stampoulis, D., Nikolopoulos, E. I., Borga, M., & Vegara, H. J. (2014). Rainfall organization control on the flood response of mild-slope basins. *Journal of Hydrology*, *510*, 565–577. <https://doi.org/10.1016/j.jhydrol.2013.12.013>
- Meigh, J. R., Farquharson, F. A. K., & Sutcliffe, J. V. (1997). A worldwide comparison of regional flood estimation methods and climate. *Hydrological Sciences Journal*, *42*(2), 225–244. <https://doi.org/10.1080/02626669709492022>
- Mejía, A., Daly, E., Rossel, F., Javanovic, T., & Gironás, J. (2014). A stochastic model of streamflow for urbanized basins. *Water Resources Research*, *50*, 1984–2001. <https://doi.org/10.1002/2013WR014834>

REFERENCES

- Menabde, M., & Sivapalan, M. (2001). Linking space-time variability of river runoff and rainfall fields: A dynamic approach. *Advances in Water Resources*, 24(9–10), 1001–1014. [https://doi.org/10.1016/S0309-1708\(01\)00038-0](https://doi.org/10.1016/S0309-1708(01)00038-0)
- Merz, B., Basso, S., Fischer, S., Lun, D., Blöschl, G., Merz, R., et al. (2022). Understanding heavy tails of flood peak distributions. *Water Resources Research*, 1–37. <https://doi.org/10.1029/2021wr030506>
- Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth and Environment*, 2(9), 592–609. <https://doi.org/10.1038/s43017-021-00195-3>
- Merz, R., & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, 39(12), 1–20. <https://doi.org/10.1029/2002WR001952>
- Merz, R., & Blöschl, G. (2009). Process controls on the statistical flood moments - a data based analysis. *Hydrological Processes*, 23(5), 675–696. <https://doi.org/10.1002/hyp>
- Miniussi, A., & Marani, M. (2020). Estimation of Daily Rainfall Extremes Through the Metastatistical Extreme Value Distribution: Uncertainty Minimization and Implications for Trend Detection. *Water Resources Research*, 56(7), 1–18. <https://doi.org/10.1029/2019WR026535>
- Miniussi, A., Marani, M., & Villarini, G. (2020). Metastatistical Extreme Value Distribution applied to floods across the continental United States. *Advances in Water Resources*, 136(December 2019), 103498. <https://doi.org/10.1016/j.advwatres.2019.103498>
- Molnar, P., Anderson, R. S., Kier, G., & Rose, J. (2006). Relationships among probability distributions of stream discharges in floods, climate, bed load transport, and river incision. *Journal of Geophysical Research: Earth Surface*, 111(2), 1–10. <https://doi.org/10.1029/2005JF000310>
- Morin, E., & Yakir, H. (2014). Hydrological impact and potential flooding of convective rain cells in a semi-arid environment. *Hydrological Sciences Journal*, 59(7), 1353–1362. <https://doi.org/10.1080/02626667.2013.841315>
- Morrison, J. E., & Smith, J. A. (2002). Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resources Research*, 38(12), 41-1-41–12. <https://doi.org/10.1029/2001wr000502>

- Moussa, R. (2003). On morphometric properties of basins, scale effects and hydrological response. *Hydrological Processes*, 17(1), 33–58. <https://doi.org/10.1002/hyp.1114>
- Müller, M. F., Dralle, D. N., & Thompson, S. E. (2014). Analytical model for flow duration curves in seasonally dry climates. *Water Resources Research*, 50, 5510–5531. <https://doi.org/10.1002/2014WR015301>
- Müller, M. F., Roche, K. R., & Dralle, D. N. (2021). Catchment processes can amplify the effect of increasing rainfall variability. *Environmental Research Letters*, 16(8). <https://doi.org/10.1088/1748-9326/ac153e>
- Mushtaq, S., Miniussi, A., Merz, R., & Basso, S. (2022). Reliable estimation of high floods: A method to select the most suitable ordinary distribution in the Metastatistical extreme value framework. *Advances in Water Resources*, 161(September 2021), 104127. <https://doi.org/10.1016/j.advwatres.2022.104127>
- Mushtaq, S., Miniussi, A., Merz, R., Tarasova, L., Marra, F., & Basso, S. (2023). Prediction of extraordinarily high floods emerging from heterogeneous flow generation processes. *Geophysical Research Letters* (accepted). <https://doi.org/10.1029/2023GL105429>
- Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, M., Marelle, L., Samset, B. H., et al. (2019). Frequency of extreme precipitation increases extensively with event rareness under global warming. *Scientific Reports*, 9(1), 1–10. <https://doi.org/10.1038/s41598-019-52277-4>
- Németh, L., Hübnerová, Z., & Zempléni, A. (2019). Trend detection in GEV models, 1–13. Retrieved from <http://arxiv.org/abs/1907.09435>
- Nerantzaki, S. D., & Papalexiou, S. M. (2022). Assessing extremes in hydroclimatology: A review on probabilistic methods. *Journal of Hydrology*, 605(December 2021), 127302. <https://doi.org/10.1016/j.jhydrol.2021.127302>
- Nerantzaki, S. D., & Papalexiou, S. M. (2019). Tails of extremes: Advancing a graphical method and harnessing big data to assess precipitation extremes. *Advances in Water Resources*, 134. <https://doi.org/10.1016/j.advwatres.2019.103448>
- Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323–351. <https://doi.org/10.1080/00107510500052444>

REFERENCES

- Nicótina, L., Alessi Celegon, E., Rinaldo, A., & Marani, M. (2008). On the impact of rainfall patterns on the hydrologic response. *Water Resources Research*, *44*(12), 1–14. <https://doi.org/10.1029/2007WR006654>
- Obled, C., Wendling, J., & Beven, K. (1994). The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of Hydrology*, *159*(1–4), 305–333. [https://doi.org/10.1016/0022-1694\(94\)90263-1](https://doi.org/10.1016/0022-1694(94)90263-1)
- Öztekin, T. (2005). Comparison of parameter estimation methods for the three-parameter generalized Pareto distribution. *Turkish Journal of Agriculture and Forestry*, *29*(6), 419–428. <https://doi.org/10.3906/tar-0409-4>
- Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G. J., et al. (2011). Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature*, *470*(7334), 382–385. <https://doi.org/10.1038/nature09762>
- Pallard, B., Castellarin, A., & Montanari, A. (2009). *A look at the links between drainage density and flood statistics*. *Hydrol. Earth Syst. Sci* (Vol. 13). Retrieved from www.hydrol-earth-syst-sci.net/13/1019/2009/
- Pan, X., Rahman, A., Haddad, K., & Ouarda, T. B. M. J. (2022). Peaks-over-threshold model in flood frequency analysis: a scoping review. *Stochastic Environmental Research and Risk Assessment*, *36*(9), 2419–2435. <https://doi.org/10.1007/s00477-022-02174-6>
- Papalexiou, S. M., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, *17*(2), 851–862. <https://doi.org/10.5194/hess-17-851-2013>
- Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions : A global survey on extreme daily rainfall. *Water Resources Research*, *49*(1), 187–201. <https://doi.org/10.1029/2012WR012557>
- Papalexiou, S. M., & Montanari, A. (2019). Global and Regional Increase of Precipitation Extremes Under Global Warming. *Water Resources Research*, *55*, 4901–4914. <https://doi.org/10.1029/2018WR024067>
- Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S., & Burlando, P. (2014). On the effects of small scale space-time variability of rainfall on basin flood response. *Journal of Hydrology*, *514*, 313–327. <https://doi.org/10.1016/j.jhydrol.2014.04.014>

- Peleg, N., Ban, N., Gibson, M. J., Chen, A. S., Paschalis, A., Burlando, P., & Leitão, J. P. (2022). Mapping storm spatial profiles for flood impact assessments. *Advances in Water Resources*, *166*(June), 104258. <https://doi.org/10.1016/j.advwatres.2022.104258>
- Peleg, N., Blumensaat, F., Molnar, P., Fatichi, S., & Burlando, P. (2017). Partitioning the impacts of spatial and climatological rainfall variability in urban drainage modeling. *Hydrology and Earth System Sciences*, *21*(3), 1559–1572. <https://doi.org/10.5194/hess-21-1559-2017>
- Peleg, N., Marra, F., Fatichi, S., Molnar, P., Morin, E., Sharma, A., & Burlando, P. (2018). Intensification of convective rain cells at warmer temperatures observed from high-resolution weather radar data. *Journal of Hydrometeorology*, *19*(4), 715–726. <https://doi.org/10.1175/JHM-D-17-0158.1>
- Peleg, N., & Morin, E. (2014). Stochastic convective rain-field simulation using a high-resolution synoptically conditioned weather generator (HiReS-WG). *Water Resources Research*, *50*, 2124–2139. <https://doi.org/10.1002/2013WR014836>
- Pendergrass, A. G. (2018). What precipitation is extreme? *Science*, *360*(6393), 1072–1073. <https://doi.org/10.1126/science.aat1871>
- Pokhrel, P., & Gupta, H. V. (2011). On the ability to infer spatial catchment variability using streamflow hydrographs. *Water Resources Research*, *47*(8). <https://doi.org/10.1029/2010WR009873>
- Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *American Naturalist*, *164*(5), 625–632. <https://doi.org/10.1086/424970>
- Post, D. A., & Jakeman, A. J. (1996). Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments. *Hydrological Processes*, *10*(6), 877–892. [https://doi.org/10.1002/\(SICI\)1099-1085\(199606\)10:6<877::AID-HYP377>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1099-1085(199606)10:6<877::AID-HYP377>3.0.CO;2-T)
- Pumo, D., Viola, F., La Loggia, G., & Noto, L. V. (2014). Annual flow duration curves assessment in ephemeral small basins. *Journal of Hydrology*, *519*(PA), 258–270. <https://doi.org/10.1016/j.jhydrol.2014.07.024>

REFERENCES

- Qiu, J., Shen, Z., Leng, G., & Wei, G. (2021). Synergistic effect of drought and rainfall events of different patterns on watershed systems. *Scientific Reports*, *11*(1), 1–18. <https://doi.org/10.1038/s41598-021-97574-z>
- Rajah, K., O’Leary, T., Turner, A., Petrakis, G., Leonard, M., & Westra, S. (2014). Changes to the temporal distribution of daily precipitation. *Geophysical Research Letters*, *41*(24), 8887–8894. <https://doi.org/10.1002/2014GL062156>
- Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., & Gratzki, A. (2013). A Central European precipitation climatology - Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, *22*(3), 235–256. <https://doi.org/10.1127/0941-2948/2013/0436>
- Rentschler, J., Salhab, M., & Jafino, B. A. (2022). Flood exposure and poverty in 188 countries. *Nature Communications*, *13*(1), 3527. <https://doi.org/10.1038/s41467-022-30727-4>
- Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer US.
- Rigon, R., Bancheri, M., Formetta, G., & deLavenne, A. (2016). The geomorphological unit hydrograph from a historical-critical perspective. *Earth Surface Processes and Landforms*, *41*(1), 27–37. <https://doi.org/10.1002/esp.3855>
- Rinaldo, A., Botter, G., Bertuzzo, E., Uccelli, A., Settin, T., & Marani, M. (2006). Transport at basin scales: 1. Theoretical framework. *Hydrology and Earth System Sciences*, *10*(1), 19–29. <https://doi.org/10.5194/hess-10-19-2006>
- Rinaldo, A., Marani, A., & Rigon, R. (1991). Geomorphological dispersion. *Water Resources Research*, *27*(4), 513–525. <https://doi.org/10.1029/90WR02501>
- Rinaldo, A., & Rodriguez-Iturbe, I. (1996). Geomorphological theory of the hydrological response. *Hydrological Processes*, *10*(6), 803–829. [https://doi.org/10.1002/\(SICI\)1099-1085\(199606\)10:6<803::AID-HYP373>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1099-1085(199606)10:6<803::AID-HYP373>3.0.CO;2-N)
- Rodríguez-Iturbe, I., & Valdés, J. B. (1979). The geomorphologic structure of hydrologic response. *Water Resources Research*, *15*(6), 1409–1420. <https://doi.org/10.1029/WR015i006p01409>
- Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., et al. (2012). Runoff models and flood frequency statistics for design flood estimation in Austria - Do they

- tell a consistent story? *Journal of Hydrology*, 456–457, 30–43.
<https://doi.org/10.1016/j.jhydrol.2012.05.068>
- Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., & Merz, R. (2012). Step changes in the flood frequency curve : Process controls. *Water Resources Research*, 48, 1–15. <https://doi.org/10.1029/2011WR011187>
- Rossi, M. W., Whipple, K. X., & Vivoni, E. R. (2016). Precipitation and evapotranspiration controls on daily runoff variability in the contiguous United States and Puerto Rico. *Journal of Geophysical Research: Earth Surface*, 128–145.
<https://doi.org/doi:10.1002/2015JF003446>
- Rowinski, P. M., Strupczewski, W. G., & Singh, V. P. (2002). Applicabilité des distributions de probabilité log-Gumbel et log-logistique dans les analyses hydrologiques: I. Pdf connues. *Hydrological Sciences Journal*, 47(1), 107–122.
<https://doi.org/10.1080/02626660209492911>
- Roy, J., Tejedor, A., & Singh, A. (2022). Dynamic Clusters to Infer Topologic Controls on Environmental Transport of River Networks. *Geophysical Research Letters*, 49, 1–11.
<https://doi.org/10.1029/2021GL096957>
- Saharia, M., Kirstetter, P. E., Vergara, H., Gourley, J. J., Emmanuel, I., & Andrieu, H. (2021). On the Impact of Rainfall Spatial Variability, Geomorphology, and Climatology on Flash Floods. *Water Resources Research*, 57(9), 1–18.
<https://doi.org/10.1029/2020WR029124>
- Saharia, M., Kirstetter, P. E., Vergara, H., Gourley, J. J., Hong, Y., & Giroud, M. (2017). Mapping flash flood severity in the united states. *Journal of Hydrometeorology*, 18(2), 397–411. <https://doi.org/10.1175/JHM-D-16-0082.1>
- Santos, A. C., Portela, M. M., Rinaldo, A., & Schaepli, B. (2018). Analytical flow duration curves for summer streamflow in Switzerland. *Hydrology and Earth System Sciences*, 22(4), 2377–2389. <https://doi.org/10.5194/hess-22-2377-2018>
- Sarkar, S., & Maity, R. (2021). Global climate shift in 1970s causes a significant worldwide increase in precipitation extremes. *Scientific Reports*, 11(1), 1–11.
<https://doi.org/10.1038/s41598-021-90854-8>

REFERENCES

- Sartori, M., & Schiavo, S. (2015). Connected we stand: A network perspective on trade and global food security. *Food Policy*, *57*, 114–127.
<https://doi.org/https://doi.org/10.1016/j.foodpol.2015.10.004>
- Sassolas-Serrayet, T., Cattin, R., & Ferry, M. (2018). The shape of watersheds. *Nature Communications*, *9*(1), 1–8. <https://doi.org/10.1038/s41467-018-06210-4>
- Schaefli, B., Rinaldo, A., & Botter, G. (2013). Analytic probability distributions for snow-dominated streamflow. *Water Resources Research*, *49*(5), 2701–2713.
<https://doi.org/10.1002/wrcr.20234>
- Schumm, S. A. (1956). Evolution of drainage systems and slopes in Badlands at Perth Amboy, New Jersey. *Geological Society of America Bulletin*, *67*, 597–646.
- Seckin, N., Haktanir, T., & Yurtal, R. (2011). Flood frequency analysis of Turkey using L-moments method. *Hydrological Processes*, *25*(22), 3499–3505.
<https://doi.org/10.1002/hyp.8077>
- Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If Precipitation Extremes Are Increasing, Why Aren't Floods? *Water Resources Research*, *54*(11), 8545–8551.
<https://doi.org/10.1029/2018WR023749>
- Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, *11*, 1649–1669.
- Sivapalan, M., Blöschl, G., Merz, R., & Gutknecht, D. (2005). Linking flood frequency to long-term water balance: Incorporating effects of seasonality. *Water Resources Research*, *41*(6), 1–17. <https://doi.org/10.1029/2004WR003439>
- Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., et al. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, *48*(6), 857–880. <https://doi.org/10.1623/hysj.48.6.857.51421>
- Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L., & Bates, P. (2018). Strange Floods: The Upper Tail of Flood Peaks in the United States. *Water Resources Research*, *54*(9), 6510–6542. <https://doi.org/10.1029/2018WR022539>
- Smith, M. B., Koren, V. I., Zhang, Z., Reed, S. M., Pan, J. J., & Moreda, F. (2004). Runoff response to spatial variability in precipitation: An analysis of observed data. *Journal of Hydrology*, *298*(1–4), 267–286. <https://doi.org/10.1016/j.jhydrol.2004.03.039>

- Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*(1), 72–101. <https://doi.org/10.2307/1412159>
- St. George, S., & Mudelsee, M. (2019). The weight of the flood-of-record in flood frequency analysis. *Journal of Flood Risk Management*, *12*(S1), 1–8. <https://doi.org/10.1111/jfr3.12512>
- Strahler, A. N. (1964). Quantitative geomorphology of drainage basins and channel networks. In *Chow by VenTe (ed) Handbook of applied hydrology*. New York: McGraw Hill Book Company.
- Struthers, I., & Sivapalan, M. (2007). A conceptual investigation of process controls upon flood frequency: Role of thresholds. *Hydrology and Earth System Sciences*, *11*(4), 1405–1416. <https://doi.org/10.5194/hess-11-1405-2007>
- Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *Annals of Statistics*, *35*(6), 2769–2794. <https://doi.org/10.1214/009053607000000505>
- Tabari, H. (2020). Climate change impact on flood and extreme precipitation increases with water availability. *Scientific Reports*, *10*(1), 1–10. <https://doi.org/10.1038/s41598-020-70816-2>
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. New York: Random House. <https://doi.org/10.4324/9781912281206>
- Tarasova, L., Basso, S., & Merz, R. (2020a). Transformation of Generation Processes From Small Runoff Events to Large Floods. *Geophysical Research Letters*, *47*(22). <https://doi.org/10.1029/2020GL090547>
- Tarasova, L., Basso, S., Wendi, D., Viglione, A., Kumar, R., & Merz, R. (2020b). A Process-Based Framework to Characterize and Classify Runoff Events: The Event Typology of Germany. *Water Resources Research*, *56*(5). <https://doi.org/10.1029/2019WR026951>
- Tarasova, L., Basso, S., Zink, M., & Merz, R. (2018). Exploring Controls on Rainfall-Runoff Events: 1. Time Series-Based Event Separation and Temporal Dynamics of Event Runoff Response in Germany. *Water Resources Research*, *54*(10), 7711–7732. <https://doi.org/10.1029/2018WR022587>
- Tarasova, L., Lun, D., Merz, R., Blöschl, G., Basso, S., Bertola, M., et al. (2023). Shifts in flood generation processes exacerbate regional flood anomalies in Europe.

REFERENCES

Communications Earth and Environment, 4(1), 1–12. <https://doi.org/10.1038/s43247-023-00714-8>

Tashie, A., Pavelsky, T., & Band, L. E. (2020a). An Empirical Reevaluation of Streamflow Recession Analysis at the Continental Scale. *Water Resources Research*, 56(1), 1–18. <https://doi.org/10.1029/2019WR025448>

Tashie, A., Pavelsky, T., & Emanuel, R. E. (2020b). Spatial and Temporal Patterns in Baseflow Recession in the Continental United States. *Water Resources Research*, 56(3), 1–18. <https://doi.org/10.1029/2019WR026425>

Tashie, A., Scaife, C. I., & Band, L. E. (2019). Transpiration and subsurface controls of streamflow recession characteristics. *Hydrological Processes*, 33(19), 2561–2575. <https://doi.org/10.1002/hyp.13530>

Thorarinsdottir, T. L., Hellton, K. H., Steinbakk, G. H., Schlichting, L., & Engeland, K. (2018). Bayesian Regional Flood Frequency Analysis for Large Catchments. *Water Resources Research*, 54(9), 6929–6947. <https://doi.org/10.1029/2017WR022460>

Thornthwaite, C. W. (1948). An Approach toward a Rational Classification of Climate. *Geographical Review*, 38(1), 55. <https://doi.org/10.2307/210739>

Van Alphen, J., Martini, F., Loat, R., Slomp, R., & Passchier, R. (2009). Flood risk mapping in Europe, experiences and best practices. *Journal of Flood Risk Management*, 2(4), 285–292. <https://doi.org/10.1111/j.1753-318X.2009.01045.x>

Vázquez, A., Oliveira, J. G., Dezsö, Z., Goh, K. Il, Kondor, I., & Barabási, A. L. (2006). Modeling bursts and heavy tails in human dynamics. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 73(3), 1–19. <https://doi.org/10.1103/PhysRevE.73.036127>

Viglione, A., Merz, R., & Blöschl, G. (2009). On the role of the runoff coefficient in the mapping of rainfall to flood return periods. *Hydrology and Earth System Sciences*, 13(5), 577–593. <https://doi.org/10.5194/hess-13-577-2009>

Viglione, A., Chirico, G. B., Woods, R., & Blöschl, G. (2010). Generalised synthesis of space-time variability in flood response: An analytical framework. *Journal of Hydrology*, 394(1–2), 198–212. <https://doi.org/10.1016/j.jhydrol.2010.05.047>

Villarini, G., & Smith, J. A. (2010). Flood peak distributions for the eastern United States. *Water Resources Research*, 46(6), 1–17. <https://doi.org/10.1029/2009WR008395>

Villarini, G., Smith, J. A., Baeck, M. L., Marchok, T., & Vecchi, G. A. (2011). Characterization of rainfall distribution and flooding associated with U.S. landfalling

- tropical cyclones: Analyses of Hurricanes Frances, Ivan, and Jeanne (2004). *Journal of Geophysical Research Atmospheres*, 116(23). <https://doi.org/10.1029/2011JD016175>
- Vivekanandan, N. (2015). Flood frequency analysis using method of moments and L-moments of probability distributions. *Cogent Engineering*, 2(1). <https://doi.org/10.1080/23311916.2015.1018704>
- Volpi, E., Di Lazzaro, M., Bertola, M., Viglione, A., & Fiori, A. (2018). Reservoir Effects on Flood Peak Discharge at the Catchment Scale. *Water Resources Research*, 54(11), 9623–9636. <https://doi.org/10.1029/2018WR023866>
- Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., & Koutsoyiannis, D. (2019). Save hydrological observations! Return period estimation without data decimation. *Journal of Hydrology*, 571(February), 782–792. <https://doi.org/10.1016/j.jhydrol.2019.02.017>
- Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D., & Wong, W. K. (2016). Evidence for changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway. *Journal of Hydrology*, 538, 33–48. <https://doi.org/10.1016/j.jhydrol.2016.03.066>
- Vorogushyn, S., & Merz, B. (2013). Flood trends along the Rhine: The role of river training. *Hydrology and Earth System Sciences*, 17(10), 3871–3884. <https://doi.org/10.5194/hess-17-3871-2013>
- Wang, H., Merz, R., & Basso, S. (2023a). Constructing a geography of heavy-tailed flood distributions : insights from common streamflow dynamics. *Earth and Space Science Open Archive (Preprint)*. <https://doi.org/10.22541/essoar.169651331.19199134/v1>
- Wang, H.-J., Merz, R., Yang, S., & Basso, S. (2023b). Inferring heavy tails of flood distributions from common discharge dynamics. *EGUsphere (Preprint)*, 2023(May), 1–24. <https://doi.org/https://doi.org/10.5194/egusphere-2023-660>
- Wang, H., Merz, R., Yang, S., Tarasova, L., & Basso, S. (2022). Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. *Advances in Water Resources Journal*, 171(104359). <https://doi.org/10.1016/j.advwatres.2022.104359>
- Ward, A. S., Wondzell, S. M., Schmadel, N. M., & Herzog, S. P. (2020). Climate Change Causes River Network Contraction and Disconnection in the H.J. Andrews Experimental Forest, Oregon, USA. *Frontiers in Water*, 2(April), 1–10. <https://doi.org/10.3389/frwa.2020.00007>
- Wasko, C., & Sharma, A. (2017). Global assessment of flood and storm extremes with increased temperatures. *Scientific Reports*, 7(1), 1–8. <https://doi.org/10.1038/s41598-017-08481-1>

REFERENCES

- Wasko, C., Sharma, A., & Westra, S. (2016). Reduced spatial extent of extreme storms at higher temperatures. *Geophysical Research Letters*, *43*(8), 4026–4032. <https://doi.org/10.1002/2016GL068509>
- Welch, B. L. (1947). The generalization of “Student’s” problem when several different population variances are involved. *Biometrika*, *34*(1–2), 28–35. <https://doi.org/doi:10.1093/biomet/34.1-2.28>. MR 0019277. PMID 20287819
- Werner, T., & Upper, C. (2002). Time Variation in the Tail Behaviour of Bund Futures.
- Wietzke, L. M., Merz, B., Gerlitz, L., Kreibich, H., Guse, B., Castellarin, A., & Vorogushyn, S. (2020). Comparative analysis of scalar upper tail indicators. *Hydrological Sciences Journal*, *65*(10), 1625–1639. <https://doi.org/10.1080/02626667.2020.1769104>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83. <https://doi.org/10.2307/3001968>
- Wilson, P. S., & Toumi, R. (2005). A fundamental probability distribution for heavy rainfall. *Geophysical Research Letters*, *32*(14), 1–4. <https://doi.org/10.1029/2005GL022465>
- Withanage, N. S., Dayawansa, N. D. K., & De Silva, R. P. (2014). Morphometric analysis of the Gal Oya river basin using spatial data derived from GIS. *Tropical Agricultural Research*, *26*(1), 175–188. <https://doi.org/10.4038/tar.v26i1.8082>
- Wittenberg, H. (1999). Baseflow recession and recharge as nonlinear storage processes, *726*(January 1998), 715–726.
- Woltemade, C. J., & Potter, K. W. (1994). A watershed modeling analysis of fluvial geomorphic influences on flood peak attenuation. *Water Resources Research*, *30*(6), 1933–1942.
- Wu, Q., Ke, L., Wang, J., Pavelsky, T. M., Allen, G. H., Sheng, Y., et al. (2023). Satellites reveal hotspots of global river extent change. *Nature Communications*, *14*(1). <https://doi.org/10.1038/s41467-023-37061-3>
- Yang, L., Yang, Y., & Smith, J. (2021). The Upper Tail of Flood Peaks Over China: Hydrology, Hydrometeorology, and Hydroclimatology. *Water Resources Research*, *57*(11). <https://doi.org/10.1029/2021WR030883>

- Yokoo, Y., & Sivapalan, M. (2011). Towards reconstruction of the flow duration curve: Development of a conceptual framework with a physical basis. *Hydrology and Earth System Sciences*, *15*(9), 2805–2819. <https://doi.org/10.5194/hess-15-2805-2011>
- Zhang, X. S., Amirthanathan, G. E., Bari, M. A., Laugesen, R. M., Shin, D., Kent, D. M., et al. (2016). How streamflow has changed across Australia since the 1950s: Evidence from the network of hydrologic reference stations. *Hydrology and Earth System Sciences*, *20*(9), 3947–3965. <https://doi.org/10.5194/hess-20-3947-2016>
- Zhao, F., Zhang, L., Chiew, F. H. S., Vaze, J., & Cheng, L. (2013). The effect of spatial rainfall variability on water balance modelling for south-eastern Australian catchments. *Journal of Hydrology*, *493*, 16–29. <https://doi.org/10.1016/j.jhydrol.2013.04.028>
- Zhu, Z., Wright, D. B., & Yu, G. (2018). The Impact of Rainfall Space-Time Structure in Flood Frequency Analysis. *Water Resources Research*, *54*(11), 8983–8998. <https://doi.org/10.1029/2018WR023550>
- Zink, M., Kumar, R., Cuntz, M., & Samaniego, L. (2017). A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty. *Hydrology and Earth System Sciences*, *21*(3), 1769–1790. <https://doi.org/10.5194/hess-21-1769-2017>
- Zocatelli, D., Borga, M., Viglione, A., Chirico, G. B., & Blöschl, G. (2011). Spatial moments of catchment rainfall: Rainfall spatial organisation, basin morphology, and flood response. *Hydrology and Earth System Sciences*, *15*(12), 3767–3783. <https://doi.org/10.5194/hess-15-3767-2011>
- Zocatelli, D., Borga, M., Zanon, F., Antonescu, B., & Stancalie, G. (2010). Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania. *Journal of Hydrology*, *394*(1–2), 148–161. <https://doi.org/10.1016/j.jhydrol.2010.07.019>
- Zocatelli, D., Borga, M., Zanon, F., Antonescu, B., Stancalie, G., Viglione, A., et al. (2014). On the tails of extreme event distributions in hydrology. *Water Resources Research*, *13*(1), 1–10. <https://doi.org/10.1016/j.jhydrol.2008.02.011>



List of Publications and Author Contributions

This section outlines the chapters in this dissertation, which are based on published papers and have been modified. The author contributions, aligned with the Contributor Roles Taxonomy suggested by Allen et al. (2019)¹, are detailed below.

Chapter 2

In a modified version of the published paper: Wang, H.-J., Merz, R., Yang, S., Tarasova, L., and Basso, S. (2023). Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. *Advances in Water Resources Journal*, 171(104359). <https://doi.org/10.1016/j.advwatres.2022.104359>, with copyright permission.

Author contributions

- HW: Conceptualization, Methodology, Validation, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing, Visualization, Funding acquisition
- RM: Conceptualization, Writing - review and editing, Supervision, Project administration, Funding acquisition
- SY: Methodology, Writing - review and editing, Supervision
- LT: Data curation, Writing - review and editing
- SB: Conceptualization, Methodology, Writing - review and editing, Supervision, Project administration, Funding acquisition

Chapter 3

In a modified version of the accepted paper: Wang, H.-J., Merz, R., Yang, S., and Basso, S. (2023). Inferring heavy tails of flood distributions through hydrograph recession analysis. Hydrology and Earth System Sciences, with copyright permission.

Author contributions

- HW: Conceptualization, Methodology, Validation, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing, Visualization, Funding acquisition
- RM: Writing - review and editing, Supervision, Project administration, Funding acquisition
- SY: Methodology, Writing - review and editing, Supervision
- SB: Conceptualization, Methodology, Investigation, Writing - review and editing, Supervision, Project administration, Funding acquisition

¹Allen, L., O'Connell, A., and Kiermer, V. (2019). How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing*, 32(1), 71–74. <https://doi.org/10.1002/leap.1210>

B

Declaration under Oath

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Unterschrift des Antragstellers
Signature of the applicant

Datum
Date



Curriculum Vitae

PERSONAL INFORMATION

Name Hsing-Jui Wang
Date of Birth 23 June 1987
Place of Birth Tainan City, Republic of China, ROC (Taiwan)
Sex Male
Nationality Republic of China, ROC (Taiwan)

WORK EXPERIENCE

03.2018 – 06.2019 **Hydrotech Research Institute**
Taipei, Taiwan *Research Assistant*

08.2011 – 07.2017 **The Church in Taipei Assembly Hall**
Taipei, Taiwan *Youth Counselor*

08.2010 – 07.2011 **National Taiwan University**
Taipei, Taiwan *Student Teaching Assistant*
Course: Environmental Chemistry and Lab

EDUCATION

08.2019 – present **Helmholtz Centre for Environmental Research (UFZ)**
Halle (Saale), Germany Department Catchment Hydrology
PhD Student

C. Curriculum Vitae

11.2017 – 12.2017
Hamburg, Germany

DeutschAkademie Hamburg
German studies

08.2017 – 09.2017
Taipei, Taiwan

Chinese Institute of European Languages
German studies

08.2011 – 07.2009
Taipei, Taiwan

National Taiwan University
Department of Bioenvironmental Systems Engineering
Master of Science
Thesis: Analysis of biofilm formation and detachment
on oyster shells in low velocity flow
Grade: 1.27 (converted to German system)

08.2005 – 07.2009
Taipei, Taiwan

National Taiwan University
Department of Bioenvironmental Systems Engineering
Bachelor of Science in Engineering
Grade: 1.53 (converted to German system)

08.2002 – 07.2005
Tainan, Taiwan

National Tainan First Senior High School
Diploma of High School
Gifted classes: Advanced math program

SKILLS

Languages

English (*proficient*), Chinese (*native*)
German (*basic*), Japanese (*basic*)

Programming

Python (*proficient*), Matlab (*proficient*)
R (*basic*), Fortran (*basic*)

Software

ArcGIS, LaTeX, GitHub, Mendeley

Laboratory

Water Quality Analysis, Soil Mechanics, Environmental Microbiology

Instruments

Acoustic Doppler Current Profiler (ADCP)
Atomic Absorption Spectroscopy (AAS)
Scanning Electron Microscope (SEM)

Driving Licences

Taiwan, Germany

SCIENTIFIC PUBLICATIONS

Peer-Reviewed Journal Papers

Hsing-Jui Wang, Ralf Merz, Soohyun Yang, & Stefano Basso (2023). Inferring heavy tails of flood distributions through hydrograph recession analysis. *Hydrology and Earth System Sciences*. (accepted)

Wang, H., Merz, R., Yang, S., Tarasova, L., & Basso, S. (2023). Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. *Advances in Water Resources Journal*, 171(104359). <https://doi.org/10.1016/j.advwatres.2022.104359>

Hsing-Jui Wang, Po-Kang Shih, & Wen-Lian Chang (2011). Dispersion effects analysis and constructed wetlands treatment of oyster shell contacted beds water purification systems by tracer test. *Journal of Taiwan Agricultural Engineering*. 57(3), 17–31. <https://doi.org/10.29974/JTAE.201109.0002>

Conference Papers

Wang, H.-J., Merz, R., Yang, S., & Basso, S. (2023). Exploring spatial and seasonal patterns of heavy-tailed flood behavior with catchment nonlinear responses. AGU Annual Meeting 2023, San Francisco, US, 11–15 Dec, AGU1324519. (accepted as oral presentation)

Wang, H.-J., Merz, R., Yang, S., & Basso, S. (2023). Common streamflow dynamics unraveled the heavy-tailed flood distributions. EGU General Assembly 2023, Vienna, Austria, 24–28 Apr, EGU23-2835. <https://doi.org/10.5194/egusphere-egu23-2835>, 2023. (Oral presentation)

Wang, H.-J., Yang, S., Merz, R., & Basso, S. (2022). The role of spatial rainfall variability for the emergence of heavy tails in streamflow distributions. EGU General Assembly 2022, Vienna, Austria, 23–27 May, EGU22-5194. (Oral presentation)

Wang, H.-J., Yang, S., Merz, R., & Basso, S. (2021). Investigating rainfall and catchment attributes promoting heavy-tailed distributions of river flows. EGU General Assembly 2021, Online, 19–30 Apr, EGU21-2682. (Oral presentation)

Wang, H.-J., Yang, S., Merz, R., & Basso, S. (2021). Identifying drivers of hydrological hazards of river basins in spatial heterogeneities of rainfall and catchment attributes. 2nd International Conference on Natural Hazards and Risks in a Changing World 2021, Potsdam, DE, 5-6 Oct. (Poster)

Lee, F.-Z., Lai, J.-S., Chang, W.-Y., Yang, S.-Y., Liu, C.-C., **Wang, H.-J.**, et al. (2019). Flow field measurement and movement estimation of turbidity current in a reservoir. Abstract Presented in the AOGS 16th Annual Meeting, Singapore, SG, 28 Jul–2 Aug.

Lai, J.-S., Lee, F.-Z., Yang, S.-Y., Chang, W.-Y., Liu, C.-C., **Wang, H.-J.**, et al. (2019). UAV application on the flow field measurement between the upstream and downstream of Bitan Weir. Abstract Presented in the AOGS 16th Annual Meeting, Singapore, SG, 28 Jul–2 Aug.

Lee, F.-Z., Lai, J.-S., **Wang, H.-J.**, & Liu, C.-C. (2018). Efficiency of hydraulic desiltation. Taiwan Agricultural Engineers Society Annual Meeting. Oral presentation, Kaohsiung, ROC.

C. Curriculum Vitae

Thesis

Wang, H.-J. (2011). Analysis of biofilm formation and detachment on oyster shells in low velocity flow. Master thesis. National Taiwan University, ROC. (Master thesis)

INVITED TALKS

- Seminar "Investigating the impacts of spatially variable rainfall on flood hazards applying a river-network-based stochastic-distributed model (English)": January 19, 2022, from 11:00 AM to 12:30 PM at the Hydrotech Research Institute, National Taiwan University.
- Seminar "Introduction to Stochastic Models in Advanced Hydrology": January 11, 2022, from 11:00 AM to 12:00 PM at the Department of Water Resources and Environmental Engineering, Tamkang University.

JOURNAL REVIEWERS

- Water Resources Research (WRR)
- Hydrology and Earth System Sciences (HESS)

AWARDS

- 2022 Water Data Applications Competition – Team Bronze Prize, Water Resources Agency, Taiwan
- 2020 Government Scholarship to Study Abroad (Full PhD scholarship), Ministry of Education, Taiwan
- 2011 Outstanding Paper Award (Master Degree), Chi-Seng Water Management - Research & Development Foundation, Taiwan