River Floods in a Changing World

Matthias Kemter

Cumulative dissertation
for the degree

Doctor rerum naturalium
(Dr. rer. nat.)

in Natural Hazards Research

submitted to the

Faculty of Science
at the University of Potsdam

Date of submission: 20th of January 2022

Date of defense: 10th of June 2022

Reviewers

Prof. Dr. Bruno Merz
Helmholtz-Zentrum Potsdam Deutsches Geoforschungszentrum,
Universität Potsdam

Dr. habil. Norbert Marwan
Potsdam Intitut für Klimafolgenforschung
Prof. Dr. Jan Seibert
Universität Zürich

Supervisors

Prof. Dr. Bruno Merz
Prof. Dr. Jürgen Kurths
Dr. habil. Norbert Marwan

Examination Board Members

Prof. Dr. Bruno Merz
Dr. habil. Norbert Marwan
Prof. Dr. Jan Seibert
Prof. Dr. Oliver Korup
Prof. Dr. Annegret Thieken
Prof. Dr. ThorstenWagener

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-55856
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-558564

Declaration of Originality

I hereby declare that this thesis has been produced by myself without any inadmissible or undeclared help from others. I did not use any sources or means other than the ones specified in the text and reference list. All sources have been referred to and the contributions of myself and all co-authors to the included publications are thoroughly declared. Furthermore, this thesis has not been previously submitted to any university.

Location and Date

Matthias Kemter

Acknowledgments

This thesis would not exist without the constant support of various kind people. While many of them would deserve to have their name printed on the front page, this section will have to suffice to express my gratitude to them.

First of all, I want to thank my supervisors for their guidance, help, and especially their trust, which they gave in me despite my lack of prior knowledge in the field of hydrology. Thanks to Prof. Bruno Merz and Dr. habil. Norbert Marwan for the many hours they spent discussing my research, providing ideas and polishing our manuscripts. They were available whenever I needed them and provided supervision far beyond what I had expected.

Thanks to Prof. Jan Seibert for agreeing to act as an external reviewer for this thesis. Thanks to the German Research Foundation DFG for funding this research as a part of the graduate school "Natural Hazards and Risks in a Changing World". Special thanks to Prof. Annegret Thieken and Prof. Axel Bronstert as speakers of the graduate school. This program not only helped me to expand my knowledge and finish my thesis relatively care-free during a devastating pandemic, it also allowed me to find some of my dearest friends among my colleagues. I am beyond grateful for all the more or less productive discussions and activities we had.

Thanks to all of my co-authors, editors, and reviewers for helpful discussions and feedback to my work. Special thanks to all members of the Australia wildfire task-force for contributing to an inspiring and unique project that found its way into this thesis. Thanks to Friederike Körting and Seth Bryant for proofreading this thesis and providing valuable remarks.

The warmest thanks to my parents. Without them I would not be where and who I am now. They supported all of my decisions even when the final goals were not always clear to them or even me. Last, but most importantly I want to thank Elli for walking this path alongside me and providing an open ear across the hallway at all times. I could not have done it without your support.

Summary

River floods are among the most devastating natural hazards worldwide. As their generation is highly dependent on climatic conditions, their magnitude and frequency are projected to be affected by future climate change. Therefore, it is crucial to study the ways in which a changing climate will, and already has, influenced flood generation, and thereby flood hazard. Additionally, it is important to understand how other human influences – specifically altered land cover – affect flood hazard at the catchment scale.

The ways in which flood generation is influenced by climatic and land cover conditions differ substantially in different regions. The spatial variability of these effects needs to be taken into account by using consistent datasets across large scales as well as applying methods that can reflect this heterogeneity. Therefore, in the first study of this cumulative thesis a complex network approach is used to find 10 clusters of similar flood behavior among 4390 catchments in the conterminous United States. By using a consistent set of 31 hydro-climatological and land cover variables, and training a separate Random Forest model for each of the clusters, the regional controls on flood magnitude trends between 1960-2010 are detected. It is shown that changes in rainfall are the most important drivers of these trends, while they are regionally controlled by land cover conditions.

While climate change is most commonly associated with flood magnitude trends, it has been shown to also influence flood timing. This can lead to trends in the size of the area across which floods occur simultaneously, the flood synchrony scale. The second study is an analysis of data from 3872 European streamflow gauges and shows that flood synchrony scales have increased in Western Europe and decreased in Eastern Europe. These changes are attributed to changes in flood generation, especially a decreasing relevance of snowmelt. Additionally, the analysis shows that both the absolute values and the trends of flood magnitudes and flood synchrony scales are positively correlated. If these trends persist in the

future and are not accounted for, the combined increases of flood magnitudes and flood synchrony scales can exceed the capacities of disaster relief organizations and insurers.

Hazard cascades are an additional way through which climate change can influence different aspects of flood hazard. The 2019/2020 wildfires in Australia, which were preceded by an unprecedented drought and extinguished by extreme rainfall that led to local flooding, present an opportunity to study the effects of multiple preceding hazards on flood hazard. All these hazards are individually affected by climate change, additionally complicating the interactions within the cascade. By estimating and analyzing the burn severity, rainfall magnitude, soil erosion and stream turbidity in differently affected tributaries of the Manning River catchment, the third study shows that even low magnitude floods can pose a substantial hazard within a cascade.

This thesis shows that humanity is affecting flood hazard in multiple ways with spatially and temporarily varying consequences, many of which were previously neglected (e.g. flood synchrony scale, hazard cascades). To allow for informed decision making in risk management and climate change adaptation, it will be crucial to study these aspects across the globe and to project their trajectories into the future. The presented methods can depict the complex interactions of different flood drivers and their spatial variability, providing a basis for the assessment of future flood hazard changes. The role of land cover should be considered more in future flood risk modelling and management studies, while holistic, transferable frameworks for hazard cascade assessment will need to be designed.

Zusammenfassung

Flusshochwasser gehören zu den verheerendsten Naturkatastrophen weltweit. Ihre Entstehung hängt von klimatischen Bedingungen ab, weshalb vorhergesagt wird, dass sich ihre Magnituden und Häufigkeit durch den Klimawandel ändern werden. Daher ist es notwendig zu untersuchen, auf welche Art sich ein verändertes Klima – auch im Vergleich mit Effekten durch Landbedeckungsänderungen – auf Hochwasserentstehung und -gefahr auswirken könnte und das bereits getan hat.

Diese kumulative Arbeit beleuchtet drei Teilaspekte dieses Themas. In der ersten Studie werden mittels maschinellen Lernens die wichtigsten Variablen entdeckt und untersucht, die die Änderungen von Hochwassermagnituden in 4390 Einzugsgebieten in den USA von 1960-2010 kontrolliert haben. Es wird gezeigt, dass Änderungen der Regenmengen der entscheidende Faktor waren, während Landnutzung regional von großer Bedeutung war. Die zweite Studie untersucht von 1960-2010 Änderungen in der Distanz innerhalb welcher Hochwasser in verschiedenen Flüssen gleichzeitig auftreten. Daten von 3872 europäischen Flusspegeln zeigen, dass sich die Fläche der gleichzeitigen Überflutung in Westeuropa vergrößert und in Osteuropa verkleinert hat, was auf abnehmende Relevanz der Schneeschmelze bei der Hochwasserentstehung zurückzuführen ist. Die dritte Studie behandelt die Auswirkungen kaskadierender Naturkatastrophen auf Hochwasser am Beispiel der australischen Waldbrände 2019/2020. Die Untersuchung der verschieden stark betroffenen Nebenflüsse des Manning River zeigt, dass in einer Naturgefahrenkaskade selbst gewöhnliche Hochwasser substantielle Auswirkungen haben können.

Diese Arbeit zeigt, dass die Menschheit Hochwassergefahren auf verschiedene Arten und mit räumlich sowie zeitlich variablen Resultaten beeinflusst. Diese Aspekte müssen zukünftig global näher untersucht und ihre Entwicklung für die Zukunft modelliert werden, um fundierte Entscheidungen in Hochwasserschutz treffen zu können. Für Hochwassermagnituden und die Fläche gleichzeitiger Überflutung können hierfür die präsentierten Methoden adaptiert werden.

Contents

List of Figures

List of Tables

Abbreviations

ALE	Accumulated Local Effects
AMS	Annual Maximum Streamflow
AR	Atmospheric River
CAPE	Convective Available Potential Energy
CIN	Convective Inhibition
CONUS	Conterminous United States
dNBR	Differential Normalized Burned Ratio
FSS	Flood Synchrony Scale
GEV	Generalized Extreme Value
MSE	Mean Squared Error
NBR	Normalized Burned Ratio
NIR	Near-Infrared
NSW	New South Wales
NTU	Nephelometric Turbidity Unit
RFE	Recursive Feature Elimination
RTE	Relative Total Effect
RUSLE	Revised Universal Soil Loss Equation
SWIR	Short-Wave Infrared
USGS	United States Geological Survey

1 Introduction

1.1 Flood Hazard and Risk in the 21st Century

River floods are the most common and deadly natural hazard worldwide (Doocy et al. 2013) and global flood damages are estimated at about US$ 104 billion per year (Desai et al. 2015). However, average global flood fatalities and flood-affected people per year have been decreasing since the 1990s (Merz et al. 2021). This is attributed to improved flood protection and building standards, especially in developing countries (Jongman et al. 2015). Past trends in flood damages have high uncertainties and mostly vanish if corrected for inflation and economic growth (Kron et al. 2012). By the end of the 21st century given a 2°C global warming scenario, models project substantial increases in the number of flood-affected people (+76%), flood fatalities (+103%), and economic damages (+520%), but with a high dependence on the scenario of future socioeconomic change (Doocy et al. 2013). Both the past and future trends in flood hazard and risk vary immensely for different parts of the world (Merz et al. 2021).

Depending on how flood generation changes regionally, the effects on flood hazard can differ. The most obvious impacts of flood generation changes on flood hazard, are changes in flood magnitude, i.e. the peak discharge in the river. In different hydro-climatological parts of the world, a changing climate is projected to cause either upward or downward trends in flood magnitudes (Arnell and Gosling 2016; Hirabayashi et al. 2013). Most commonly, researchers focus on the positive trends in flood magnitudes, as higher flood peaks can lead to larger inundation areas and higher damages. Furthermore, flood protection measures need to be improved to cope with the increasing flood hazard, ideally before these increased magnitudes occur. However, negative flood magnitude trends can also influence policy; for example in reducing design criteria and over-spending on flood protection infrastructure. Otherwise, unnecessary measures are taken, draining resources

from public and private budgets, which could be used for protection from other hazards or climate change adaptation.

Generally, flood protection is designed using flood return levels. These are defined as the river height or discharge that will on average occur once in a certain time period. In most countries of Western Europe, flood protection measures are designed to withstand a 100-year return level, while globally the design return period generally varies between 10-500 years (Merz et al. 2021). Return levels are usually estimated by fitting an extreme value distribution to recorded data. These distributions assume the underlying process is stationary in time. If flood generating processes were to change, return levels would be misestimated and flood protection measures would become over- or under-designed (Milly et al. 2008). To withstand future floods and to remain cost-efficient, the design of these measures will need to take future flood hazard changes into account.

A second aspect of flood hazard, which could be affected by changes in flood generation, is the flood synchrony scale (FSS) – also called flood spatial extent – a measure for the area that is simultaneously affected by flooding from different rivers. Changes of the FSS have only been studied recently (Berghuijs et al. 2019; Brunner and Gilleland 2021). If the spatial extent of flood events increases, additional catchments experience floods simultaneously. Most disaster response systems are only designed to handle a certain number of events of a given magnitude at the same time. Therefore, an increase in synchronous flooding could overwhelm the disaster response capacities of regions or countries and deplete the reserves of insurance companies.

While the impact of climate change on floods is a common topic both in the scientific and public discourse, there are less explored ways in which human activity alters flood generation. Some of these are direct, targeted modifications of streamflow in rivers like reservoirs, levees, or channels (Best 2019; Ho et al. 2017; Yang et al. 2016), while others are not directly aimed at the rivers themselves, but secondary effects of alterations of the catchment. One example for these unintended effects, is the increase of surface flow due to manmade impervious surfaces, leading to faster runoff generation in the downstream river (Blum et al. 2020; Oudin et al. 2018). To accurately project changes in flood hazard, it will be necessary to account

for potential river management or land cover changes together with the changing climate.

The final aspect of potential flood hazard dynamics studied in this thesis is the impact of hazard cascades. This term describes a sequence of different hazards, in which early hazards alter the probability, magnitude, or damage potential of subsequent events (Zscheischler, Martius, et al. 2020). If the frequency and/or magnitude of antecedent hazards, for example droughts, changes due to climate change, it can have unexpected consequences for flood hazard and risk (Zscheischler et al. 2018). Flood risk and its future trajectory could be misestimated in regions prone to hazard cascades where hydrologic models neglect these interactions (UNDRR 2019).

To accurately predict how river floods will change in the future – depending on the climate-change and land-use pathways human societies take – we must understand how and why flood generation has changed in the past. This thesis provides insights on how humanity affects river flood magnitude, synchrony, and generation. These insights are meant to aid future studies on projected flood hazard, by highlighting the variety of ways in which changes in flood generation can affect flood risk at various scales, and by providing a set of relevant controls on these effects.

1.2 State of the Art

A well known mechanism by which globally increasing temperatures affect floods is through the increased water storage capacity of a warmer atmosphere called the Clausius–Clapeyron relationship, leading to an intensification of the water cycle (Trenberth 2011). Despite this correlation of atmospheric temperature and precipitation intensity, both positive and negative trends in flood magnitudes have been found across the world (Sharma et al. 2018). Depending on the climate region, this discrepancy is often explained by 1) drier catchment states prior to the floods due to increased evapotranspiration (Ivancic and Shaw 2015) and 2) decreasing availability of snow during the flood season leading to reduced magnitudes of snowmelt-associated floods (Yan et al. 2019). Both model-based projections and empirical studies show a tendency of dry regions to become even drier due to

climate change, while wet regions receive more rain (Allan and Soden 2007; John et al. 2009).

While these atmospheric processes are well studied, previous work aiming to attribute flood magnitude changes has often focused on a small set of potential controls on flood magnitude trends (e.g. Blöschl et al. 2019; Slater and Villarini 2016) or narrowly focused on a hydro-climatologically homogeneous region (e.g. Armstrong et al. 2014; Bertola et al. 2019; Mallakpour and Villarini 2015). However, attributing flood magnitude trends at the continental and global scale requires a dataset and methodology that captures and adapts to the regionally varying factors of flood generation. Furthermore, previous studies have often focused on catchments with minimal human alteration (e.g. Hodgkins et al. 2017; Ye et al. 2017). While this is a useful prerequisite to isolate the effects of climate change, the applicability of the results from such studies for flood hazard assessment is limited, as inhabited catchments of interest are, by definition no longer natural.

To understand changes in flood generation, it is necessary to first analyze the generating processes of historical floods. The most common method to do this is flood type classification, which classifies historical floods by generating process, which usually are rainfall, snowmelt, and soil moisture excess. Most studies from a hydrological perspective use one of two general approaches to classify floods (see Tarasova et al. 2019 for a review). The first is based on a comparison of the flood date and the average date of occurrence of maxima of certain flood generating processes in the catchment (e.g. Berghuijs et al. 2016; Kochanek et al. 2012). While the comparison of dates is computationally simple and largely unbiased due to a lack of fixed thresholds, it is only based on mean dates of the flood generating processes, which are often associated with high variance and do not reflect the actual catchment state prior to the flood (Tarasova et al. 2019). This can lead to a poor representation of unseasonal floods and shifts in timing of both the floods and their generating processes throughout the time series. Additionally, this approach generally only allows for a singular type per event and does not account for mixed types (e.g. rain on snow floods). The second approach is to calculate the catchment state in a certain flood concentration time prior to each flood based on climate data (e.g. Merz and Blöschl 2003; Stein et al. 2020; Tarasova et al. 2020). Commonly, a decision tree is used in this approach to determine the

most likely flood type. In contrast to the first method, the catchment state approach allows for any number of flood types or mixtures, is based on the immediate catchment conditions, and handles all floods equally, independent of the seasonality of flood processes. However, this approach is more dependent on the temporal accuracy of the underlying climate data and the classification trees commonly require thresholds which are based on expert opinion, potentially introducing biases. Given the advantages of the catchment state approach to account for unseasonal floods and temporal shifts in flood generation – which are a focus area of this thesis – it is used in Chapters 2 and 3 to classify flood generating processes.

Most studies have focused on how changes in flood generation influence flood magnitudes but have given little attention to shifts in flood timing. Yet, such temporal shifts could lead to 1) less predictable flood seasonality and 2) changes in the size of the area across which floods occur simultaneously in different catchments. While the Clausius–Clapeyron relationship is often discussed with respect to the increased intensity of rainfall, it also leads to an increased size of storm systems (Lochbihler et al. 2017; Molnar et al. 2015). This can cause growing footprints of rainstorms, leading to larger spatial scales, across which river floods are generated simultaneously. Berghuijs et al. (2019) have shown a general increase in spatial flood extents across Europe, using the FSS. However, they did not address the drivers of this trend or its spatial distribution. Furthermore, a clear relationship between the intensity and spatial extent of rainstorms exists (Skøien et al. 2003). If there is a similar relationship for river floods, are the trends in flood magnitudes and flood extents similar? Are they both caused by the same changes in flood generation? These questions have not been studied so far.

The previously discussed effects are direct consequences of climate change or human alterations. Additionally, climatic changes can have a substantial indirect impact on flood hazard through hazard cascades, which alter the state of the catchment prior to a flood. The scarcity of data and complexity of cascading hazards leads to their general underrepresentation in risk assessment (Zscheischler, Martius, et al. 2020). So far, systemic research about hazard cascades containing floods have been limited to the interaction of different flood types (e.g. coastal and river floods) occurring simultaneously (UNDRR 2019). However, many other hazards can influence flood generation. Wildfires can cause alterations of the soil

surface, making it hydrophobic and thereby increasing surface runoff (Debano and Krammes 1966). During subsequent rainfall, this can lead to faster than usual run-off generation in the downhill river with a characteristic sharp spike in the discharge time series a few minutes to hours prior to the main flood peak (Shakesby and Doerr 2006). This surface runoff also leads to increased soil erosion and carries charred soil and vegetation into the river system, negatively affecting water quality (Smith et al. 2011).

Similar to wildfires, droughts can make soils water-repellent (Gilmour 1968). However, at the catchment scale, this surface runoff enhancing behavior is superimposed by a reduction of the runoff coefficient (Petheram et al. 2011), meaning that the same amount of rainfall leads to less streamflow during drought conditions than during an average catchment state. It was previously shown that droughts can decrease the probability of high streamflow (99^{th} percentile) occurrence even when the extreme rainfall occurrence rate remains unchanged (Petheram et al. 2011).

1.3 Research Questions and Structure

This thesis aims to answer certain aspects of the overarching research question:

- **How and with which consequences does humanity affect flood generation and flood hazard?**

The three research articles that comprise chapters 2-4 provide insights into three different aspects of this question, which are presented as the sub-questions Q1-Q3 below. Each of them focuses on a different part of the world: the Conterminous United States (CONUS), Europe, and Australia, respectively. This is partly due to data availability restrictions, but also allows for a representation of the globally diverse hydro-climatological conditions, which influence flood generation.

Chapter 2 assesses the controls on flood magnitude trends in CONUS from 1960-2010. This study aims to find these controls by applying Random Forests and Accumulated Local Effect plots to a large set of hydro-meteorological and land-use variables. These methods allow for the representation of non-linear relationships and the detection of the most important controls despite the presence of collinearity between the variables. Regional variability in flood generation and its changes are

reflected by clustering catchments based on similar flood behavior using Complex Networks. Based on these methods, Chapter 2 aims to answer the question:

Q1 **Which variables controlled flood magnitude trends in recent decades?**

Chapter 3 deals with a different aspect of flood hazard: spatial flood extents. These are represented through the FSS i.e., the maximum distance, across which at least 50% of catchments record an annual maximum streamflow near-simultaneously (±7 days). Following the calculation of trends of flood magnitudes and FSS for 3872 catchments in Europe, the correlation of both their absolute values and their trends is studied. Using reanalysis climate data, daily catchment average time series of hydro-climatological variables relevant for flood generation (precipitation, snowmelt, soil moisture) are calculated. These are used to classify all flood events by their flood generating process according to the antecedent catchment state. By comparing trends in the regional relevance of these processes and the average flood extents for each process, it is possible to find the drivers of flood extent trends, answering the questions:

Q2 **How is the spatial extent of synchronous river flooding related to flood magnitudes? How have changes in flood generation influenced flood extents?**

In Chapter 4, the cascading hazards surrounding the 2019/2020 "Black Summer" wildfires in Australia are investigated. This series of events ended in widespread flooding, the generation of which is analyzed along the example of the severely affected Manning River catchment. Different tributaries were affected to varying degrees by drought, fire, and extreme rainfall, allowing for the comparison of the local effects. All of these hazards are influenced by climate change. By investigating this hazard cascade, we can learn how alterations from prior hazards can modulate the risk of subsequent flooding and how the effects of a changing climate propagate in a cascade:

Q3 **How can hazard cascades affect floods in the context of climate change?**

With respect to the main research question, these studies can only investigate a certain subset of the multitude of human influences on flood hazard. Nonetheless,

they aim to represent the variety of these interactions without sacrificing the depth required to reflect each of them properly. Chapter 5 provides a comprehensive discussion of the implications that the results from Chapters 2-4 have on the main research question. It contextualizes these results and explores their meaning for flood hazard research and management. Furthermore, it suggests further research that should be conducted to validate and apply these results in other regions of the world.

1.4 Author Contributions

The three manuscripts that comprise the following three chapters are published or under review for publication in peer-reviewed journals. While the author of this thesis is the main author of all of them, several co-authors contributed to these studies in various ways. Their contributions are as follows:

Chapter 2: Conceptualization: **MK**, BM, NM; Formal Analysis: **MK**; Investigation: **MK**, BM, NM, GV; Methodology: **MK**, BM, NM; Visualization: **MK**; Software: **MK**; Supervision: BM, NM; Writing - Original Draft: **MK**; Writing - review & editing: **MK**, BM, NM, GV

Chapter 3: Conceptualization: **MK**, BM, NM; Formal Analysis: **MK**; Investigation: **MK**, BM, NM, GB, SV; Methodology: **MK**, BM, NM, GB; Visualization: **MK**; Software: **MK**; Supervision: BM, NM; Writing - Original Draft: **MK**; Writing - review & editing: **MK**, BM, NM, GB, SV

Chapter 4: Conceptualization: **MK**, MF, LL, ES, JV, AB, OK; Formal Analysis: **MK**, MF, LL, ES, JV, AB; Investigation: **MK**, MF, LL, ES, JV; Methodology: **MK**, MF, LL, ES, JV; Visualization: **MK**, ES; Software: **MK**, MF, JV; Supervision: OK, KT; Writing - Original Draft: **MK**, MF, LL, ES, JV, AB; Writing - review & editing: **MK**, MF, LL, ES, JV, AB, OK, KT

Additionally, the author contributed to the following publications, which are not included in this thesis:

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., Bertola, M., **Kemter, M.**, Kreibich, H., Lall, U., Macdonald, E. (2021). Causes, impacts and patterns of disastrous river floods. Nature Reviews Earth & Environment, 1-18. https://doi.org/10.1038/s43017-021-00195-3

Thieken, A., **Kemter, M.**, Vorogushyn, S., Berghäuser, L., Sieg, T., Natho, S., Mohor, G., Petrow, T., Merz, B., Bronstert, A. (2021). Extreme Hochwasser bleiben trotz integriertem Risikomanagement eine Herausforderung Einleitung: Hochwasserrisikomanagement in Deutschland. Technical Report. https://www.uni-potsdam.de/fileadmin/projects/natriskchange/Taskforces/Flut2021_StatementThiekenEtAl.pdf

Macdonald, E., Merz, B., Guse, B., Wietzke, L., Ullrich, S., **Kemter, M.**, Ahrens, B., Vorogushyn, S. (2022) Event and Catchment Controls of Heavy Tail Behavior of Floods. Water Resources Research (under review)

Banerjee A., **Kemter, M.**, Goswami, B., Marwan, N., Merz, B., Kurths, J. (2022) Spatial coherence patterns of extreme winter precipitation in the United States. Theoretical and Applied Climatology (submitted)

2 Controls on Flood Trends Across the United States

Submitted to Water Ressources Research: Kemter, M., Marwan, N., Villarini, G. & Merz, B. (2022, under review). Controls on flood trends across the United States.

Abstract

Trends in flood magnitudes have been shown to vary across the conterminous USA (CONUS). There have been attempts to identify what controls these regionally varying trends, but these attempts were limited to certain, e.g. climatic, variables or smaller regions, using different methods and datasets each time. Here we attribute the trends in annual maximum streamflow for 4390 gauging stations across the CONUS in the period 1960-2010, while using a consistent methodology and exploring a large variety of potential controlling variables. Using a novel combination of process-based flood classification and complex networks, we find 10 distinct clusters of catchments with similar flood behavior. We compile a set of 31 hydro-climatological and land use variables as predictors for 10 separate Random Forest models, allowing us to find the main controls on the flood magnitude trends for each cluster. By using Accumulated Local Effect plots, we can understand how the predictors influence the trends in the flood magnitude. We show that hydro-climatologic changes and land use are of similar importance for flood magnitude trends across the CONUS. Static land use variables are more important than their trends, suggesting that land use is able to attenuate (in forested areas) or amplify (in urbanized areas) the effects of climatic changes on flood magnitudes. For some variables, we find opposing effects for different regions, showing that a combination of a comprehensive dataset and a cluster-based methodology is necessary to attribute flood magnitude trends reliably at the continental scale while maintaining the sensitivity to regional controls.

2.1 Introduction

Floods are among the most harmful natural hazards in the conterminous United States (CONUS) and worldwide (Munich Re 2021). The damages they cause could grow, as increasing global temperatures have led to higher water content in the atmosphere and an intensification of precipitation extremes throughout the CONUS (Yin et al. 2018). While we would expect this intensification to lead to higher flood peaks (Sharma et al. 2018), in recent decades annual maximum streamflow (AMS) – which in this study will represent flood magnitudes – has both increased and decreased in different parts of the CONUS (Archfield et al. 2016; Hirsch and Ryberg 2012; Sauer et al. 2021; Slater and Villarini 2016; Villarini and Slater 2017). This is partly caused by the superposition of different climate change effects. It has been suggested that increasing extreme rainfall can be offset by decreasing soil moisture due to higher evapotranspiration (Ivancic and Shaw 2015), reduced meltwater availability from decreasing snowpacks (Yan et al. 2019) or reduced catchment wetness due to smaller rainstorm extents (Sharma et al. 2018). Additionally, the effects of climate change on flood generation can be modulated by direct human influences in the catchment through land use and land use changes (Blum et al. 2020; Hodgkins et al. 2019). Understanding this variety of controls on changes in flood magnitudes and their interactions represents critical information for the predictions of future changes to prevent growing damage and fatalities in regions of increasing AMS magnitudes and avoid unwarranted costs for flood protection in areas with a decreasing flood hazard.

Recent studies have attributed AMS magnitude trends in the CONUS to changes in selected weather variables (Do et al. 2020; Hirsch and Ryberg 2012; Hodgkins et al. 2017; Slater and Villarini 2016; Villarini and Slater 2018), mainly precipitation, soil moisture, snowmelt, and temperature. Others have focused on certain hydrometeorologically homogeneous parts of the CONUS considering only the variables that seem most relevant in that region (Armstrong et al. 2014; Mallakpour and Villarini 2015). Additionally, there has been work on the influence of river regulation and urbanization on flood magnitude trends (Blum et al. 2020; Hodgkins et al. 2019; Vogel et al. 2011). This concentration on either certain controls on flood trends or a smaller subregion was often seen as necessary, given the diversity

in climatology, flood generation processes, and land cover across the CONUS. However, the different catchment selection criteria and methodological approaches limit the comparability among regions and studies.

Here we aim to identify which factors influenced AMS magnitude trends in 4390 catchments for the period 1960-2010 using the same consistent dataset and methods for the entire CONUS (Figure 2.1). We include anthropogenically influenced catchments, which allow us to compare the impacts of direct human influences and a changing climate. Using complex networks, we find clusters of catchments with similar flood behavior. We train separate Random Forest models (Breiman 2001) for each cluster to select the most important variables to explain variability in flood magnitude trends, drawing from the same set of 31 hydro-meteorological and land cover variables in all clusters. Accumulated Local Effect plots are a recent advance in interpretable machine learning (Apley and Zhu 2020), which we use to understand how each selected predictor contributes to AMS magnitude trends within the different regions.

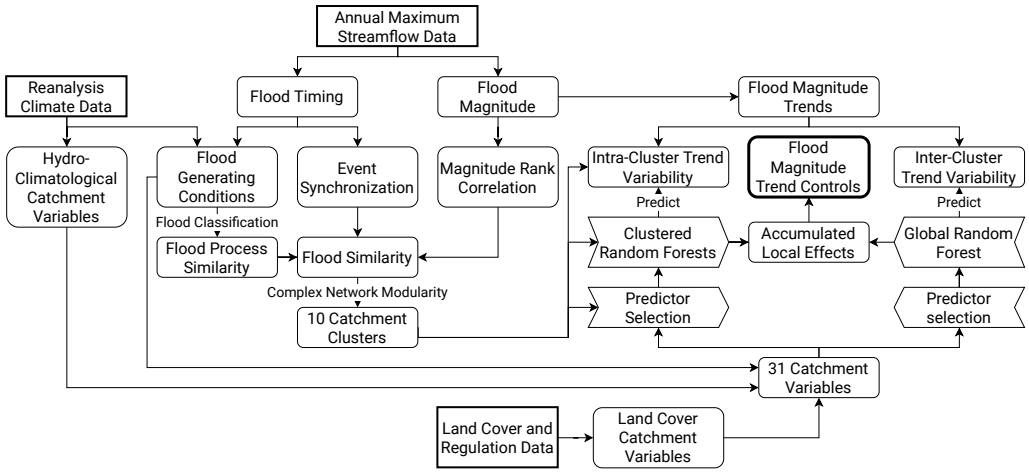

Figure 2.1: Workflow Overview. Input data are shown in square boxes, processed data in rounded squares, processes and models in arrow-shaped boxes, and the final result in the thick-bordered rounded square. We use flood and climate data to cluster catchments by flood similarity. For each cluster (intra-cluster variability) we train a Random Forest to find the most important controls on flood magnitude trends among 31 catchment variables. Additionally, we train one global Random Forest to find CONUS-scale controls on flood magnitude trends (inter-cluster variability). We investigate the effects of the controls on flood magnitude trends, using Accumulated Local Effects. See Section 2.2 and Table 2.A1 for more details on data. See Section 2.3 for more details on methods.

2.2 Data

2.2.1 Streamflow Data

This work is based on annual maximum instantaneous streamflow (AMS) data provided by the United States Geological Survey (USGS). We start with data from 5002 catchments that are located completely within the CONUS and have at least 30 years of record within the period 1960-2010. Data consist of magnitude and date of the AMS, and catchment characteristics (e.g., catchment area) for each gauging station. We selected the period 1960-2010 because: (1) the number of stations with sufficient data declines drastically for earlier and later years, and (2) the reanalysis dataset used to characterize potential controls on flood trends is only available until 2011. For each gauging station, the corresponding catchment boundaries are available from the USGS Streamgage NHDPlus Version 1 dataset.

2.2.2 Climate and Land Use Data

We use the Livneh daily reanalysis dataset with 0.0625° spatial resolution (Livneh et al. 2013), and calculate catchment-averaged daily time series for precipitation (mm), snowmelt (mm of water equivalent), total evapotranspiration (mm), and soil moisture (mm). Snowmelt is calculated from the daily difference in snow water equivalent. Soil moisture refers to the vertical soil level 2 of the underlying hydrological model, which represents the soil layer that reflects daily to weekly changes in weather conditions. The upper boundary of this layer lies at 10 cm depth, while its thickness varies between 50-110 cm for each pixel as it was iteratively estimated during model calibration (Tang et al. 2012). To calculate rainfall from precipitation, we subtract any daily net gain of snow water equivalent (reflecting snowfall) in a cell from that day's precipitation value. From these hydro-climatological variables, we derive two types of data for different purposes: (1) daily time series for the classification of all floods in terms of flood generating processes; and (2) long-term averages (e.g., mean annual rainfall, mean annual snowfall) and trends of the annual values, serving as potential controls on flood trends.

To study the effects of human activity on flood magnitude trends, we also consider averages and trends of land cover data from the GAGES-II dataset (Falcone

2017; Falcone and LaMotte 2016; Homer et al. 2015), which provides catchment averages for multiple land cover variables at different time steps (Table 2.A1) between 1960 and 2010 and is available for 4663 stations of the AMS dataset. The remaining 339 stations are not considered for further analysis.

2.3 Methods

2.3.1 Predictor Variables

To find the controls on the flood magnitude trends we calculate 31 predictor variables. Details on their calculation can be found in Table 2.A1, while we here explain their potential relevance for flood magnitude trends. We divide the predictors into five categories: flood characteristics, climate, flood type, land use, and topography. The flood characteristics category contains variables describing direct flood water contributions from the three main sources (rain, snowmelt, and soil moisture). To estimate their catchment averages prior to the floods, we calculate the concentration time t_c for every catchment:

$$t_c = 0.1 * A^{0.3}, \tag{2.1}$$

where A is the catchment area in km2 (Corradini et al. 1995; Robinson and Sivapalan 1997). For rain and snowmelt, we calculate the mean values during t_c prior to the flood day, which is the time period during which water from these sources can affect the flood generation directly. For soil moisture, we use the value of the day prior to t_c to capture the antecedent catchment wetness state. Using the Theil-Sen estimator (Sen 1968), we calculate trends in these three values, as they can be expected to have a direct impact on flood magnitude trends. Additionally, we consider AMS date trend as a variable in this category, as a shift in AMS timing can represent a change in flood generation that could explain flood magnitude trends. We calculate this based on the Theil-Sen estimator in a form that is adapted to circular values (Blöschl et al. 2017). Finally, the flood characteristics category contains the mean flood magnitude, which is an obvious choice, as we are aiming to understand its changes.

The climate category contains variables describing the long-term availability

of the three flood generating water sources. We calculate the annual mean and the trends of these annual means for rain, snowmelt, and soil moisture. It was previously shown that flood magnitudes change differently depending on catchment wetness or snow availability (Blöschl et al. 2019). Trends in rain, snowmelt, and soil moisture can directly affect flood magnitudes due to increased/decreased water availability. Furthermore, we consider precipitation seasonality (Walsh and Lawler 1981) and its trend, as flood magnitudes in catchments that are mainly influenced by seasonal precipitation events (e.g., hurricanes, atmospheric rivers) could change differently from those in catchments with less seasonal precipitation. The final variable in the climate category is aridity, as the ratio of total annual evapotranspiration and precipitation, which factors in differences in runoff generation in arid/humid catchments (Farquharson et al. 1992; Metzger et al. 2020). Aridity has been shown to have a strong effect on the flood frequency distribution especially in terms of flood magnitude variance (Guo et al. 2014).

The flood type category contains the relative importance of five flood generating processes, which are based on the flood process classification described in Section 2.3.2. Information about the processes controlling flood generation is crucial to understand flood magnitude trends, as catchments will react differently to changes in a flood cause that is often causing AMS, than to changes in a cause that is less relevant for flood generation in those catchments (Tarasova et al. 2019). Similarly, how the relevance of these processes has changed over time could be a crucial factor in flood magnitude trends (Blöschl et al. 2019; Kemter et al. 2020), which is why the trends in their occurrence comprise the remaining five variables in this category.

The land use category includes six variables representing natural and anthropogenic land cover, its changes, as well as direct human influences on streamflow. The relative area of the catchment covered by canopy can affect flood magnitudes, for example due to higher water storage capacity or delayed snowmelt in forested areas (Blöschl et al. 2007; Marks et al. 1998). As no data on changes in canopy cover are available for most of our study period, we only consider a single value in 2011. Impervious surfaces greatly affect flood generation, by decreasing flood concentration times and increasing surface flow (Blum et al. 2020). Therefore, we use both mean catchment imperviousness and its trend as variables for our

study. Comparably, cropland cover alters soil surface properties, affecting runoff generation, evapotranspiration, and thereby flood magnitudes (Schilling et al. 2014), which is why we consider cropland cover and its trend as variables. Finally, water regulation and diversion are targeted human alterations of streamflow, often directly aimed at reducing flood magnitudes (Magilligan and Nislow 2005).

We do not include purely spatial variables (i.e. latitude, longitude, and elevation) as they do not have an effect on floods themselves, but would only be a proxy for other variables such as snowfall or aridity. Their inclusion could prevent the selection of the actual flood magnitude trend controls in the Random Forests. Therefore, the topography category only contains the catchment area, which has been shown to have an impact on flood magnitude trends (Bertola et al. 2020). Analogously to the purely spatial variables, we do not consider the mean flood date as a variable, which is often used to link floods to their generating processes (e.g. Berghuijs et al. 2016), as these processes are already reflected in the variables of the flood type and flood characteristics category. Nevertheless, we calculate mean AMS dates for all catchments using the circular mean, to contextualize some of our results in the discussion.

2.3.2 Flood Similarity and Clustering

As climate, topography, flood generation, land use, and the temporal variations of potential controls vary regionally throughout the CONUS, we assume that controls on flood magnitude trends vary regionally as well. Therefore, we divide the catchments into clusters with similar behavior in terms of flood generating processes, flood magnitude fluctuations and flood timing. This approach does not aim to find clusters of similar flood magnitude trends. First, we use catchment-average reanalysis data and a simple classification tree (Figure 2.A1) to estimate the importance of the five different flood generating processes (i.e., long rain, short rain, snowmelt, rain-on-snow and soil moisture excess) in each catchment. This classification is adapted from a previous study in Europe (Kemter et al. 2020). It is based on the catchment and weather conditions during the concentration time t_c prior to the day of the flood peak. We classify each AMS and estimate the flood generating process similarity $PS_{x,y}$ for each pair of catchments x and y as:

$$PS_{x,y} = \frac{1}{N}\sum_{n=1}^{N} p \; with \; p = \begin{cases} 1 \; if \; P_{x,n} = P_{y,n} \\ 0 \; else, \end{cases} \tag{2.2}$$

where N is the total number of years for which data exist for the two stations, n is one of those years, and P is the classified generating process for a flood at a given station in a certain year. PS is one (zero) for stations that have all (no) processes in common.

Second, we test each pair of stations x and y for correlation in terms of flood magnitudes, which we call $RC_{x,y}$. We use the Spearman rank correlation, which compares the magnitudes solely based on relative rank, as absolute flood magnitudes for catchments with similar flood conditions can vary greatly, mainly because of catchment area. We normalize all $RC_{x,y}$ values to the range 0-1 by adding 1 to each value and dividing it by 2. Therefore, values of zero describe complete anti-correlation between the ranks of the flood magnitudes at two stations, while one denotes perfect correlation.

Third, we calculate the event synchronization (ES) of AMS events between all stations (Malik et al. 2010). We do not limit this analysis to commonly recorded years, as this would lead to high ES values at station pairs with low temporal overlap in the flood record. Let t_l^x and t_m^y be vectors of the flood dates at the stations x and y, where $l = 1, 2, 3, ...L_x$ and $m = 1, 2, 3, ...M_y$, with L_x and M_y as the total number of flood events at the respective stations. We allow for a time lag τ of up to 7 days between floods at two stations for them to be considered synchronous. We use a fixed time lag instead of the more commonly applied variable time lag (Agarwal et al. 2017), because our use of AMS values – as opposed to peak over threshold events – prevents multiple counting of synchronous events. We count the amount of events $ES_{x,y}$ that occur at both stations within τ:

$$ES_{x,y} = \sum_{l=1}^{L_x}\sum_{m=1}^{M_y} c \; with \; c = \begin{cases} 1 \; if \; 0 \leq t_l^x - t_m^y \leq \tau \\ 0 \; else \end{cases} \tag{2.3}$$

We then define the event synchronization $C_{x,y}$ as:

$$C_{x,y} = \frac{ES_{x,y}}{\sqrt{L_x * M_y}} \tag{2.4}$$

A value of zero means that no events are synchronous between two stations, while a value of one indicates that all floods at one station occurred within a week of a flood at the other station and vice versa.

Finally, the three similarity measures are combined into the flood similarity measure $FS_{x,y}$ by multiplication:

$$FS_{x,y} = PS_{x,y} * RC_{x,y} * C_{x,y} \tag{2.5}$$

This measure combines information about flood generating processes, flood magnitude distribution and flood timing. We use *FS* to find clusters of similar flood behavior using a complex network approach. A complex network consists of nodes, in this case the gauging stations, and links which connect certain nodes based on real-world connection or similarity (Boccaletti et al. 2006; Newman 2003). Complex networks are therefore ideal for this clustering, as they represent the pairwise similarity of catchments in an intuitive way. We link only those nodes that are exceptionally similar to each other. We define exceptional similarity as the 99th percentile of all *FS* values. The links, for which *FS* exceeds this similarity threshold, are denoted as ones in a binary adjacency matrix in which each row and column represents one catchment. To find clusters within this network, we compute modularity, a measure for how well the nodes within a cluster are connected to each other in comparison to their connections to nodes outside of the cluster (Clauset et al. 2004). Modularity has been successfully used to find clusters of similar flood and extreme rainfall behavior in previous studies (Conticello et al. 2020; Conticello et al. 2018; Yang et al. 2019). We use the Louvian modularity algorithm implemented in the *Brain Connectivity* Toolbox for *MATLAB* (Rubinov and Sporns 2010) to maximize the modularity of the clusters. To have a decent sample size to train the Random Forests, we merge small clusters (details on the threshold in Section 2.4.1) with the respective large cluster to which they have the most links. Very small clusters

(<10 stations) and stations that are not connected to any other station within the network are excluded from further analysis.

2.3.3 Trend Detection and Attribution

We use the Theil-Sen slope estimator to estimate the values and quantify the significance ($p<0.05$) of the trends in flood magnitudes for each station. To understand what controls the variability in flood magnitude trends (regardless of their significance) across the CONUS we use Random Forest models (Breiman 2001). This method can be used for the regression of a response variable based on several predictor variables. The prediction combines the results of a large number – in this case 500 – of decision trees, which can chose from a random subset of all predictors at each split of the tree. This kind of ensemble regression with a random element reduces overfitting, i.e. improves the model performance on new data (Breiman 2001). We use Random Forests, as we cannot a priori reject the possibility that some controls affect magnitude trends in a non-linear way, which would not be captured in a linear regression model. Furthermore, we employ forests based on conditional inference trees, which have the advantage of resulting in less biased predictor selection for correlated predictors compared to commonly used partition trees (Hothorn et al. 2006). This allows us to avoid biases in initial variable selection. We use the *cforest_unbiased* function of the *party*-package in *R*, which sets all user parameters of the Random Forests to further minimize bias (Strobl et al. 2007).

Here we use our set of 31 variables as predictors for Random Forests to detect flood magnitude trend controls, i.e. those variables that explain most of the variability in flood magnitude trends. It is not necessary to focus only on significant trends here, as non-significant magnitude trends can be caused by significant opposing effects from different controls, which are averaged at the catchment scale (Kormann et al. 2015). These mixing effects at the seemingly non-significant center of the trend distribution within each cluster are important for fitting the Random Forest models to all flood magnitude trends and their inclusion is crucial for successful trend attribution (Kormann et al. 2015). We then train one Random Forest for each cluster and interpret their results separately. To evaluate the performance of our clustered Random Forest approach we measure the predictive performance of the

model using mean squared errors (MSE). MSE is computed based on out-of-bag errors, meaning that the calculated error for the prediction of a magnitude trend value is only based on the trees that did not include that value in their training data. This is similar to a training- and test-data split and reduces overfitting of the model. We compare the distribution of squared errors of the clustered Random Forests with those of a global Random Forest trained on all data without cluster information and use a Mann-Whitney test to determine significance of the difference in error distributions (Mann and Whitney 1947).

2.3.4 Predictor Selection

To avoid overfitting by the Random Forests (i.e., good predictive capability for training data, but poor results for new data), it is advisable to reduce the amount of model predictors to as few as possible while retaining predictive accuracy. This process is called feature selection, where feature is synonymous with predictor. For this purpose, we use the Recursive Feature Elimination (RFE) algorithm of the *featureselection* function in the *R*-package *moreparty*. RFE is a method that is well suited to select the best set of predictors, even if some of them are correlated (Gregorutti et al. 2017). We use an iterative approach to reduce the number of predictors both at the intra-cluster level and across all clusters. This is accomplished by first applying RFE to each cluster separately and removing predictors using the one standard deviation rule of featureselection. All predictors that have not been selected for any of the clusters are then removed from the entire set of predictors. Afterwards, we again apply RFE to all clusters using this reduced set of predictors and repeat this process until only those predictors are left that were selected in at least one cluster. For each cluster, this final set of selected predictors is used to train a Random Forest. The iterative use of RTE reduces the number of considered predictors in each iteration, decreasing the uncertainty of the predictor selection and reducing the final number of selected predictors per cluster to an interpretable amount. For the global Random Forest, we only select the predictors once, skipping the iterative process.

2.3.5 Predictor Effects

To understand how the predictors in the Random Forest models affect the flood magnitude trends, we use Accumulated Local Effects (ALE) plots, which describe how a model predictor affects the model prediction throughout the predictor's value range (Apley and Zhu 2020). In hydrology, this method has been recently used to study the effects of catchment variables on flood generating processes (Stein et al. 2021). The uncentered ALE $\hat{g}_{j,ALE}(x)$ is calculated based on differences in the prediction within quantiles of the predictor:

$$\hat{g}_{j,ALE}(x) = \sum_{k=1}^{K_j} \frac{1}{n_j(k)} \sum_{i:x_{i,j} \in N_j(k)} \left[g(z_{k,j}, x_{i,\setminus j}) - g(z_{k-1,j}, x_{i,\setminus j}) \right] \tag{2.6}$$

where x is one of the values of the predictor j for which the ALE plot is calculated, k is one of K_j quantiles into which the range of x is divided, $n_j(k)$ is the number of values of x that fall into this quantile $N_j(k)$ ranging from $i = 1, 2, ...n_j(k)$, and $z_{k,j}$ denotes the values of x at the boundary of that quantile. Furthermore, g is the output of the prediction model and $x_{i,\setminus j}$ are the values of instance i for all other predictors except for j. This means that for each quantile we calculate the mean difference in the model response between the upper and the lower margin of that quantile. We use a variable number of quantiles depending on the number of stations within a cluster. We chose the number of quantiles so that each of them contains at least 15 stations, but limit the maximum number of quantiles at 15, to prevent the appearance of unrealistically high resolution and artifacts in the ALE plots. The uncenterd ALE values are centered to retrieve $\hat{f}_{j,ALE}(x)$ by subtracting its mean across all quantiles:

$$\hat{f}_{j,ALE}(x) = \hat{g}_{j,ALE}(x) - \frac{1}{K_j} \sum_{k=1}^{K_j} \hat{g}_{j,ALE}(x_k) \tag{2.7}$$

Furthermore, we classify the predictor effect as "positive," "negative" or "non-monotonic." For this, we first test for the existence of a significant ($p<0.05$) monotonic pattern in the ALE data using the Mann-Kendall test (Mann 1945). If a significant trend exists, we use the sign of its slope to assign "positive" or "negative" accordingly. If the Mann-Kendall test finds no significant trend, we assign "non-

monotonic". Note that this does not mean non-monotonic in the strict mathematic sense, but rather a distinction between a linear relationship and all other types of relationship. While the trend of the ALE plot gives us information about the directional effect of a predictor on the magnitude trends, we use the mean absolute ALE values to derive the absolute predictor importance, similar to Stein et al. (2021). For each cluster, we divide these values by the largest value within this cluster and call the resulting measure of predictor importance the relative total effect (RTE) with a range between 0 and 1.

2.3.6 Inter-Cluster Variability

With our setup, the Random Forests explain the variability of flood magnitude trends within each cluster, which we call intra-cluster variability. This approach however has a potential shortcoming for clusters in which the intra-cluster variability is not evenly spread around zero, but highly skewed toward the negative or positive. For example, if all flood magnitude trends in a cluster were positive (negative), the Random Forests could only explain why some catchments show more positive trends than others, but not why the overall trend of the cluster is positive (negative). We call this overall trend the inter-cluster variability, as it reflects the cluster's difference from other clusters with an overall trend closer to zero. Scenarios can be imagined in which the inter-cluster variability is controlled by different processes than the intra-cluster variability, which makes it important to study both types of variability.

For clusters with a clear majority (at least two in three stations) of trends with a certain sign, we therefore additionally analyze the inter-cluster variability. For this, we use boxplots of the 31 catchment variables, grouped by clusters, to find the most prominent differences to neighboring clusters. Additionally, we study the selected features and ALE plots of the global Random Forest, which contains more information about inter-cluster variability, as it includes all catchments from every cluster.

2.4 Results

2.4.1 Clustering Results

The clustering algorithm found 13 clusters. There were 273 catchments that were not assigned to any of the clusters and were excluded from further analysis, leaving a total of 4390 catchments. Three of the 13 clusters consisted of less than 90 catchments each: 58 in a cluster on the Florida peninsula, 64 along the eastern flank of the Rocky Mountains, and 88 on the Colorado Plateau, compared to 218 for the next smallest cluster. We considered these sample sizes as too small to guarantee reliable results, and the clusters were therefore merged with the cluster to which they had the most connections within the complex network, the Southeast, Rocky Mountains, and California clusters, respectively. This produced a final set of 10 clusters consisting of 218–654 catchments (Figure 2.2a). Flood magnitude trends (Figure 2.2b) range from −5.5%/yr to 4.6%/yr with a median of −0.0009%/yr. The 5th and 95th percentile of trends are −1.4%/yr and 1.1%/yr, respectively. There are 743 stations that show significant trends ($p<0.05$, Theil Sen), but even for those without significant trends there are distinct regions of predominantly positive and negative trends. The clusters Midwest, Southeast, and New England consist mainly (i.e. at least 66.6%) of catchments with one direction of trends (Figure 2.3a). In the seven other clusters, both positive and negative flood magnitude trends are almost equally common.

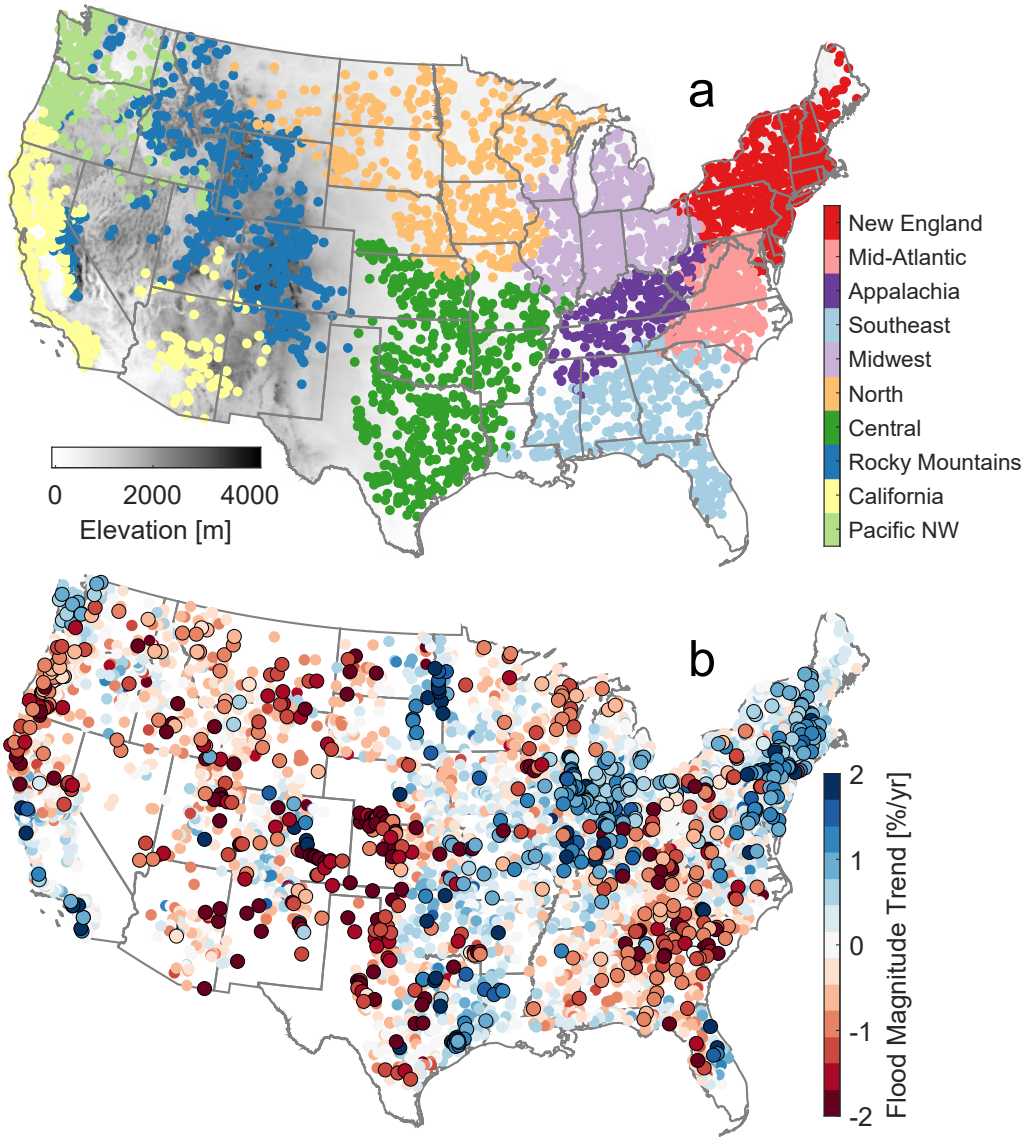

Figure 2.2: Flood Behavior Clustering and Flood Magnitude Trends for the Conterminous United States. a) Based on flood generating processes, flood magnitude distribution, and flood timing, 10 distinct clusters of catchments have been found. Clusters are based on complex network modularity, meaning that floods in catchments within a cluster are similar to each other, but distinct from those in all other clusters. Note that no spatial information is part of the clustering process and that the clustering does not aim to group catchments of similar flood magnitude trends. b) Trends in the magnitudes of annual maximum floods for 4390 stations between 1960-2010 based on Sen's slope. Large dots mark significant trends ($p<0.05$). Values are relative to the mean of the station time series. Color bar range is truncated to improve readability.

The distribution of catchment variables shows that the clusters differ substantially from each other (Figure 2.3). All clusters are unique in at least one aspect and are clearly distinct from their neighboring clusters in several metrics. As the clustering is solely based on flood behavior, this highlights how a river is aggregating the different hydro-climatological and land cover processes of a catchment.

2.4.2 Variable Selection and Effects

During the iterative predictor selection process for the Random Forests, 16 predictors were chosen in at least one cluster, following the fourth iteration, which found no unselected predictors. We only show the three most important predictors here (Figure 2.4), while the detailed ALE plots and exact importance values can be found in Section 2.5.2 (Figure 2.5) and in Section 2.A (Figures 2.A3 to 2.A11). The 15 predictors that have not been selected in any cluster are mean flood magnitude, flood date trend, relevance of long rain, rain on snow and soil moisture floods, all trends in process relevance, annual snowfall trend, annual soil moisture trend, flood generating snowmelt trend, cropland trend, and imperviousness trend. If any of these variables is important for flood magnitude trends, it is not at the intra-cluster level or only for very few catchments within a cluster.

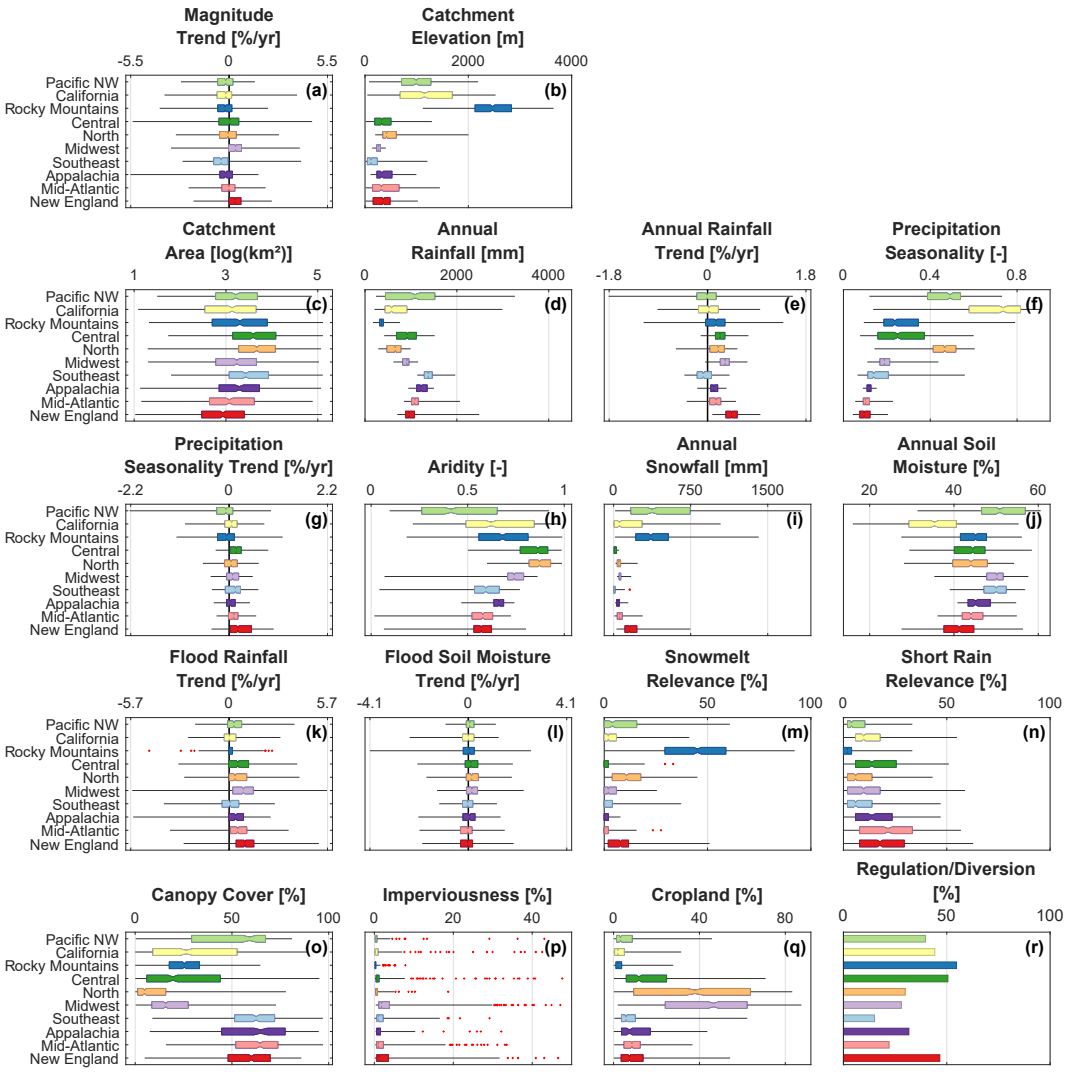

Figure 2.3: Cluster Property Comparison. Distribution of flood magnitude trends (a), catchment elevation (b), and the final set of 16 selected variables grouped by clusters (d-r). Panel titles indicate the values depicted on the x-axes. Boxes range from the 25^{th} to the 75^{th} percentile with a notch indicating the median. Whiskers reach to ten times the interquartile range or to the furthest data point, if the latter is closer. Outliers marked in red. The binary variable regulation/diversion is shown as a bar plot (r). All unselected variables are shown in Figure 2.A2.

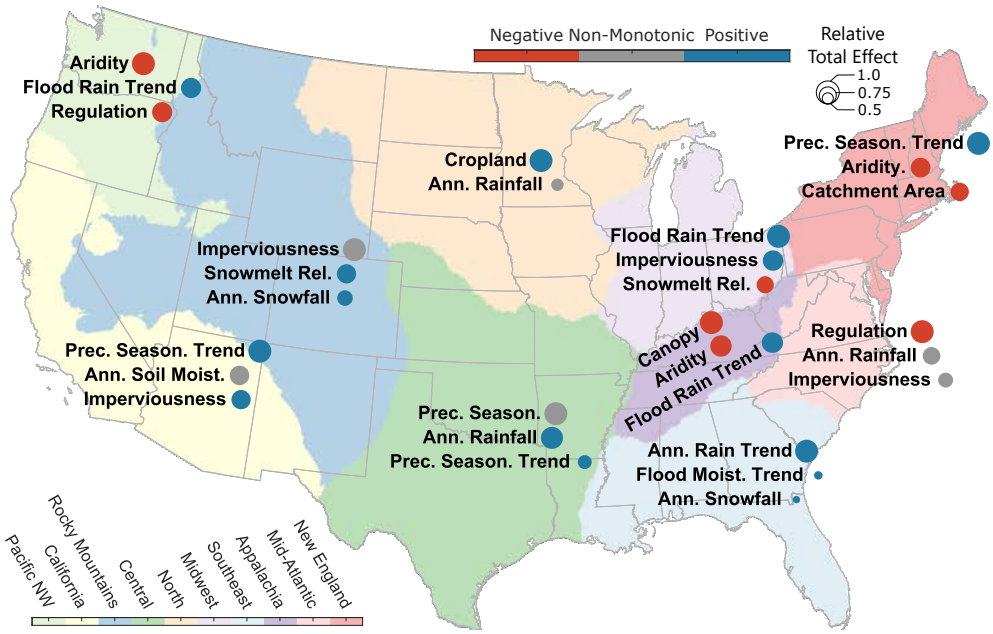

Figure 2.4: Catchment Variable Effects on Flood Magnitude Trends. Overall predictor importance (dot size) of the three most important variables in each cluster (see Figure 2.5 and Figures 2.A3 to 2.A11 for exact values and variables of lower importance). Predictor importance is calculated as the relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. The directional effect (dot color) is estimated from the ALE trend. A positive (negative) effect means that the predicted flood magnitude trend increases (decreases) with increasing variable values. Interpolated cluster map in the background depicts the most common cluster among the 20 nearest stations at each 0.05°×0.05° pixel.

During the iterative predictor selection process, between two and five predictors were selected per cluster. Catchment area, annual rainfall trend, and flood soil moisture trend were only selected in one cluster each. The variables imperviousness (4 times), mean annual rainfall, aridity, precipitation seasonality trend, flood rainfall trend, and canopy cover (3 times each) were selected most often. The categories climate and land use are almost equally important, with predictors from the respective categories being selected 15 and 11 times. At least one climate predictor was selected in all but one cluster (Midwest).

The combined prediction performance of the clustered Random Forest models (MSE=0.409) is significantly better ($p<0.05$, Mann-Whitney test) than that of the global Random Forest for the entire CONUS (MSE=0.445). In the global Random

Forest, only five predictors were selected: mean annual rainfall, annual rainfall trend, aridity, flood rainfall trend, and imperviousness. All of these are important in at least one of the clustered Random Forests as well, and four are among the most frequently selected predictors in the clustered models (all but annual rainfall trend).

2.4.3 Controls on Flood Magnitude Trends

Based on Figure 2.4 and the underlying ALE plots (Figure 2.5, Figures 2.A3 to 2.A11) we can draw conclusions on the processes that control intra-cluster variability in flood magnitude trends for the different regions. In this section, we present the controls of intra-cluster variability for the clusters in the context of previous work. For the clusters with at least 66.6% of catchments showing a trend of the same sign, we additionally discuss inter-cluster variability based on Figure 2.3.

In the Pacific NW cluster, the most important predictor is aridity, which affects flood magnitude trends negatively (Figure 2.5). Simultaneously, flood generating rainfall trends have a positive effect. This means that flood magnitude trends are the highest in very wet catchments, in which flood generating rainfall has increased, and magnitude trends are lower in other catchments. The very dry catchments in which magnitude trends are high are situated in the rain shadow of the Cascade Range in the eastern part of the cluster, where precipitation patterns and atmospheric river (AR) influence differ greatly from the coastal part of the cluster (Hu et al. 2017). This could also be linked to an increasing amount of precipitation falling as rain instead of snow (Hatchett 2018), mainly due to increasing temperatures (Yan et al. 2019) and partly due to warming ARs – which are essential for flood generation (Barth et al. 2017; Neiman et al. 2011) in the region. The positive flood magnitude trends are mostly (82.0%) limited to unregulated catchments, which shows in the negative effect of regulation/diversion.

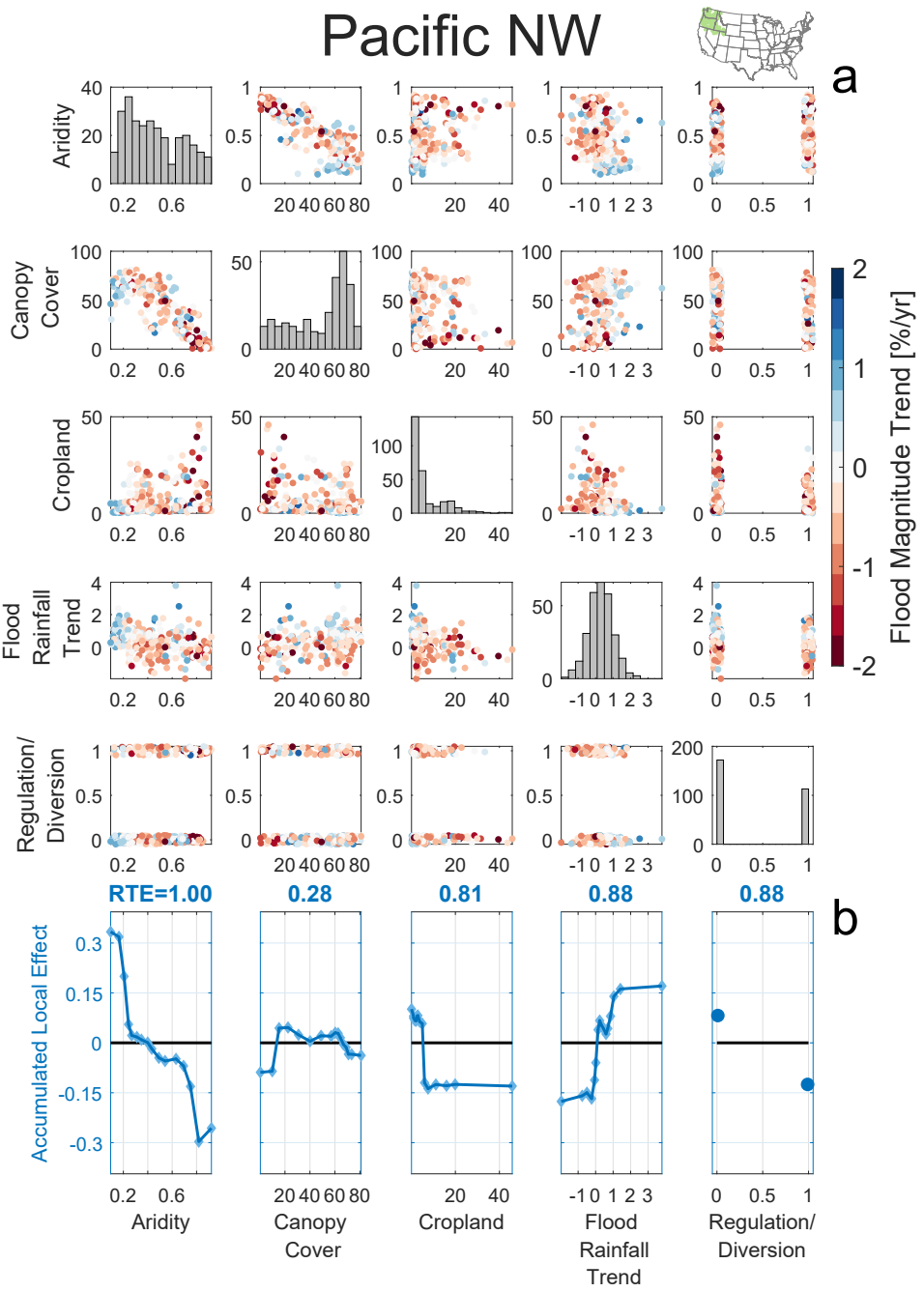

Figure 2.5: Predictor Effects for the Pacific NW Cluster. a) Bivariate scatter plots for the five predictors that were selected for this cluster. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. Random noise added to regulation/diversion values to improve visibility. Histograms plotted on the diagonal. Y-axes labels on the left apply to the entire row, except for the histograms, which show counts on the y-axes. X-axis labels in the bottom of panel b apply to the entire column. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, highlighting the variable quantile ranges.

In the California cluster (Figure 2.A3), the most important predictor is the trend in precipitation seasonality, meaning that relatively high (low) magnitude trends occur in catchments in which precipitation has become more (less) seasonal. The second highest importance is found for mean annual soil moisture, which shows a distinct, non-linear change point in the ALE plots, with a positive effect for catchments with values below 33%. This means that AMS trends are high for the dryer catchments in southern California and the Colorado Plateau, where short rain events (positive ALE) are more important for flood generation. This indicates that the different types of flood generation in this cluster (i.e., short-lived spring rain in the south and east; long rain, atmospheric rivers and a snowmelt component in the north) have changed differently, which can lead to drastic changes in flood magnitudes (Davenport et al. 2020). Additionally, there is a strong positive effect of imperviousness mainly for the highly urbanized areas of Los Angeles and the San Francisco Bay.

In the Rocky Mountains cluster (Figure 2.A4), we find that imperviousness has the highest importance, but that its highly positive effect is limited to the few (2.6%) catchments with values above 2%. The mean annual snowfall and the relevance of snowmelt have a positive effect on flood magnitude trends. These two effects fit previous findings, suggesting that in low- to medium-elevation catchments in this region, increasing temperature has led to an earlier onset of snowmelt and reduced snowpacks, causing lower streamflow in spring (Rood et al. 2016), which is the region's main AMS season. At the higher elevations with more annual snowfall and a higher relevance of snowmelt for flood generation, snowpacks and thereby floods are not yet affected by increasing temperatures (Rood et al. 2016).

The Central and North cluster show similarities in the spatial distribution of flood magnitude trends (Figure 2.2b). Both have an average magnitude trend of about zero, but with a high variance in both directions and some of the highest and lowest trends in the entire CONUS (Figure 2.3a). There is a division roughly along the 100th western meridian separating negative AMS magnitude trends to the west and positive trends to the east, except for a patch of negative trends in the northeast of the North cluster. This east-west separation is common among climatological variables in this region, as it separates the arid highland prairie of the "Great Plains" in the west from the more humid, low-lying regions in the east.

An influence on flood magnitudes that is not directly represented in our model is the lowering of the groundwater table due to human activity, which has been shown to have decreased peak streamflow in Kansas and Nebraska (Rasmussen and Perry 2001), which are part of these two clusters. Slater and Villarini (2016) have shown that this water withdrawal has decreased basin wetness in this area.

Our model finds similar importance for the non-linear negative effect of precipitation seasonality and the positive effect of mean annual rainfall on flood magnitude trends in the Central cluster (Figure 2.A5b). Both variables are highly anti-correlated here (Figure 2.A5a). Precipitation seasonality trend, with its almost linear, positive effect, is the third most important predictor. This suggests that, where precipitation is converging towards the flooding season, flood magnitudes increase, while catchments in which seasonality remains constant or diverges show no or negative flood magnitude trends. In the eastern half of the cluster, mean flood dates lie in late winter and early spring, when ARs cause most AMS (Lavers and Villarini 2013; Nayak and Villarini 2017), while supercell thunderstorm and hurricanes cause many floods in summer and fall in the west and south (Aryal et al. 2018; Smith et al. 2001). It is noteworthy that none of the land use variables we considered was selected in this cluster, suggesting that human activity is less important than hydro-climatological changes for flood magnitude trends, despite 48.2% of rivers being regulated/diverted. This high percentage does however not say anything about the degree of regulation at each of these rivers or whether water management has changed in the study period – aspects that we cannot study with this dataset. Furthermore, this seeming irrelevance of human activity does not include potential effects from water withdrawal of the Ogallala Aquifer, which is an important water source for rivers in the cluster and has been depleting between 1950 and 2015 (McGuire 2017).

In the North cluster, cropland cover is the most important variable (Figure 2.A6). It has a positive effect with a step change at a value of 30%, suggesting a direct human influence on floods through surface alteration and tile drainage, which only manifests itself in the flood magnitude trends if a large enough part of the catchment is affected. In contrast to the Central cluster, annual rainfall has a more complex effect on AMS magnitude trends here. The effect is negative for very dry catchments, positive for medium dry ones (400-600mm/yr), neutral for

medium wet ones (700-900mm/yr), and positive again for wetter catchments. It has been previously shown that precipitation is the most reliable predictor for flood frequency trends in this region (Neri et al. 2019), but for magnitude trends the situation seems to be more complicated.

In the Midwest cluster flood magnitudes have increased in 74.0% of catchments and its mean flood magnitude trend of 0.34%/yr is the highest among all clusters (Figure 2.3a). Annual rainfall has increased in 98.2% of catchments (Figure 2.3e), which is most likely a key contributor to these AMS magnitude trends. This cluster shows a narrow window of mean AMS dates with 89% of them in March and April, suggesting that flood generation is very homogeneous. It is the cluster with the highest mean imperviousness and cropland cover (Figure 2.3p and q), as well as river regulation or diversion in 31.8% of catchments, indicating a strong anthropogenic influence on flood magnitudes. The most important variable in this cluster is the trend of flood generating rainfall, which has a positive effect (Figure 2.A7). Flood generating rainfall has increased in 82.2% of the catchments, likely driven by the generally increasing rainfall amounts mentioned above. Previous work suggests that about 50% of AMS in this cluster are related to ARs (Nayak and Villarini 2017). Even though no trends in the frequency or duration of the ARs in this region could be found (Nayak and Villarini 2017), an intensification of the associated rainfall could have driven the rainfall and flood magnitude trends. The second most important variable is imperviousness, which has a strong positive effect for the 94 catchments with imperviousness values above 6%, 44 of which are situated in the Chicago metropolitan area. Increasing imperviousness due to urbanization has been linked to increasing flood frequencies in the Chicago area (Villarini et al. 2013). Canopy cover mainly has a negative effect for catchments with more than 40% canopy cover. This means that flood magnitude trends are lower in the less urbanized or cultivated catchments of this cluster, hinting at a buffering effect of forested areas for increasing rainfall sums. Snowmelt relevance has a negative effect, with catchments in which snowmelt generates more than 12% of floods showing mainly negative flood magnitude trends.

Flood magnitudes have decreased in 78.3% of the catchments in the Southeast cluster (Figure 2.3a). The trend in annual rainfall has by far the highest importance and shows a positive effect (Figure 2.A8). As 76.6% of stations in this cluster have

their mean AMS date in February or March, we calculated rainfall trends for these two months. With a cluster-wide mean of −0.43%/yr these are more than seven times lower than the annual trends (−0.059%/yr). This suggests that especially the decreasing rainfall amounts in February and March are responsible for the generally negative flood magnitude trends in the cluster. This is likely related to the decreasing number and intensity of extra-tropical cyclones (Wang et al. 2013), which are important for spring precipitation in this region (Hawcroft et al. 2012). With a much lower importance, the trend in flood generating soil moisture (positive effect) was selected. As high antecedent soil moisture is responsible for 35.7% of AMS in this cluster – which is the highest value among all clusters (Figure 2.A2e) – a change in soil moisture can be expected to affect flood magnitudes substantially here. Annual rainfall sums have decreased the most in the catchments along the Atlantic Coast, east of the 86th Meridian. We find especially negative flood magnitude trends in the catchments in the north of the cluster, which are fed by precipitation in the Appalachians, explaining the negative effect of mean annual snowfall.

In the Appalachia cluster, canopy cover is the most important predictor, and has a negative effect in the Random Forest model (Figure 2.A9). This effect stems from the highly forested regions along the western flank of the Appalachians. Aridity has the second highest effect, with relatively dry catchments showing the lowest flood magnitude trends. The trend in flood generating rainfall is of similar importance but shows a positive effect. It has been previously shown that ARs are responsible for 50-80% of AMS in this region (Nayak and Villarini 2017). Similar to the Midwest cluster, a change in the intensity of ARs could have had a significant impact on flood generating rainfall and thereby flood magnitudes in the cluster.

In the Mid-Atlantic cluster, regulation/diversion was found to be the most important predictor (Figure 2.A10). The predictor effect is negative and only 12.3% of the stations with a positive flood magnitude trend are regulated. Annual rainfall, the second most important predictor, has a non-linear effect. Annual rainfall trends are gradually increasing from the southwest to the northeast of the cluster. The effect of imperviousness is overall minor, but highly positive for the 13.0% of catchments with an imperviousness above 4%, which are mainly located in the Washington D.C. and Richmond metropolitan areas.

Most catchments in the New England cluster show positive flood magnitude trends (73.5%) and the mean flood magnitude trend of 0.31%/yr is the second highest among all clusters (Figure 2.3a). These generally positive trends are most likely due to increasing annual rainfall in all the catchments (Figure 2.3e). There is also a high mean flood generating rainfall trend of 0.92%/yr (Figure 2.3k), which, together with the decreasing relevance of rain on snow or snowmelt for flood generation in 66.2% of catchments (Figure 2.A2h and Figure 2.A2j), suggests a general shift from snow- to rain-associated floods. The ALE plots found the precipitation seasonality trend (positive) as the most important predictor in the cluster (Figure 2.A11). To reiterate, a low precipitation seasonality means that precipitation is spread out evenly throughout the year, meaning that flood magnitude trends are the highest in catchments in which precipitation has become more seasonal. Similar to the nearby Appalachia cluster, high aridity has a negative effect on flood magnitude trends here. The model found a negative, linear effect for catchment area, meaning that large catchments are experiencing smaller flood magnitude trends. Furthermore, short rain relevance has a minor positive effect. In this region, short rain floods are often caused by late spring or summer storms, which are highly sensitive to the local topography and land cover heterogeneity (Yeung et al. 2011).

2.5 Discussion

2.5.1 Clustering

The clustering process produced ten spatially distinct (Figure 2.2a) and hydro-meteorologically meaningful clusters (Figure 2.3). The number of clusters is below those in previous studies (Brunner et al. 2020; Yang et al. 2019), who found 15 clusters; their analyses were based on different datasets, clustering approaches, and similarity measures. The difference can be explained by the higher number of stations in our dataset and our relatively high minimum cluster size, which is necessary to guarantee a sufficient sample size for Random Forest training. Compared to the results of Brunner et al. (2020), our clusters are similar in extent and some of them are combinations of two or three clusters in their paper. A comparison to Yang et al. (2019) is less straightforward, as their limited dataset of

242 catchments leads to some very small clusters. Their six largest clusters, which contain 226 catchments, are similar in extent to our Pacific NW, Rocky Mountains, North, Midwest, and New England clusters as well as a combination of the Southeast, Mid-Atlantic and Appalachian clusters. In comparison to both results, our method generated less spatial overlap between clusters. These comparisons show that our clustering approach produced physically reasonable, distinct clusters, in spite of the larger size of the dataset and the ignorance of spatial information during the clustering process.

2.5.2 General Predictor Effects

We find that in all clusters in which regulation/diversion was selected, it has a negative effect on flood magnitude trends. While it can be expected due to the role of reservoirs in attenuating flood peaks to prevent damage, note that this predictor is a static value with no information about whether regulation has increased during the study period. As the GAGES data only provides a start and an end year for the influence codes, a more detailed analysis of the effects of trends in regulation/diversion is not possible in this study. The negative effects of canopy cover suggests that – similar to the effect of regulation/diversion – the flood attenuating effects of forests could also mitigate the effects of a changing climate on flood magnitudes. The ALE plots reveal that imperviousness generally has almost no effect at low values but a very high positive effect above a certain threshold (between 2-10% depending on the cluster). Both of these findings are in agreement with previous work (Hodgkins et al. 2019; Vogel et al. 2011). Like in the case of regulation/diversion, this predictor is a static value for the entire study period. In this case, we additionally use imperviousness trends as a possible predictor, which was however not selected in any of the clusters nor in the global Random Forest model. This suggests that trends in imperviousness – all of which are positive in our dataset – have a smaller effect on flood magnitudes than the amplification of other trends in flood generating mechanisms through existing impervious surfaces. Note however, that the trends of imperviousness and cropland are only based on 5 and 6 time steps, respectively (Table 2.A1). Therefore, the calculated trends might not be accurate. A dataset at a higher temporal resolution would be needed to confirm

our finding of a smaller effect of imperviousness trends. To our knowledge, no such datset exists for the study period. Cropland cover, which was selected in two clusters, shows strong opposing effects in the Pacific NW and North cluster. While 98% of catchments in the Pacific NW cluster have less than 30% cropland cover, this number is exceeded in 59% of catchments in the North cluster (Figure 2.3q). The 30% threshold is relevant because the strong positive effect of cropland cover in the North cluster mainly occurs from this value upward (Figure 2.A6b). This discrepancy in cropland value ranges explains the opposing effects and hints at a non-linear step change in the influence of cropland cover on flood generation.

The generally positive effects of annual rainfall trends, flood rainfall trends, and flood soil moisture trends are unsurprising, as a higher (lower) amount of available water in a catchment will naturally lead to increasing (decreasing) flood magnitudes. More interestingly, high precipitation seasonality generally has a negative effect, while its trend has a positive effect. This means that in some clusters, catchments with a more seasonal precipitation distribution have lower magnitude trends, while in other clusters increasing seasonality is linked to higher flood magnitude trends. In the Central cluster, both effects exist simultaneously, which seems contradictory and should be studied further.

In three clusters, aridity has a strong negative effect on flood magnitude trends, while mean annual rainfall sums have a positive or non-linear effect in three others, which is consistent with previous findings that climate change is making wet seasons wetter in rainy regions, but not in dry ones (Chou et al. 2013), and that wet regions in the CONUS have become wetter in the past decades (Greve et al. 2014). The positive effect of short rain relevance in two clusters can be explained by the increase in daily extreme rainfall magnitude across the CONUS (Yin et al. 2018), which affects short rain floods the most, which we define here as floods that are triggered mainly by the rainfall from a single day. Mean annual snowfall and snowmelt relevance show opposing effects in the two respective clusters in which they were selected. For both variables, the effects are positive in the Rocky Mountains cluster, which stems – as mentioned in Section 2.4.3 – from the effects of decreasing snowpacks and earlier snowmelt in the relatively low-lying regions (Rood et al. 2016), where less precipitation falls as snow than in the higher regions. In the two clusters with negative effects for these variables (Midwest and Southeast),

snow is in general less relevant for flood generation and in the catchments where it is relevant its decreasing abundance seems to lead to lower flood magnitude trends.

2.5.3 Inter-Cluster Variability

While the predictive performance of the clustered Random Forests is higher than that of the global model (see Section 2.4.2), the characteristics of the latter can still yield additional insights into controls on flood magnitude trends at the CONUS scale. In the global Random Forests four climatic predictors and only one land cover variable were selected (Figure 2.6), in contrast to the clustered Random Forests in which climate and land cover predictors are of almost similar importance. This suggests that climatic changes were of higher importance for past flood magnitude trends at the CONUS scale, while land cover conditions were almost equally important as climatic changes at the regional scale.

In the three clusters with predominantly negative (South Atlantic) or positive (Midwest and New England) flood magnitude trends, we identified a respective negative or positive trend in annual and flood generating rainfall sums as the most likely controls on these predominant trends. This result is reflected in the ALE plots for the global Random Forest, in which these two variables have the highest importance and a positive effect on flood magnitudes.

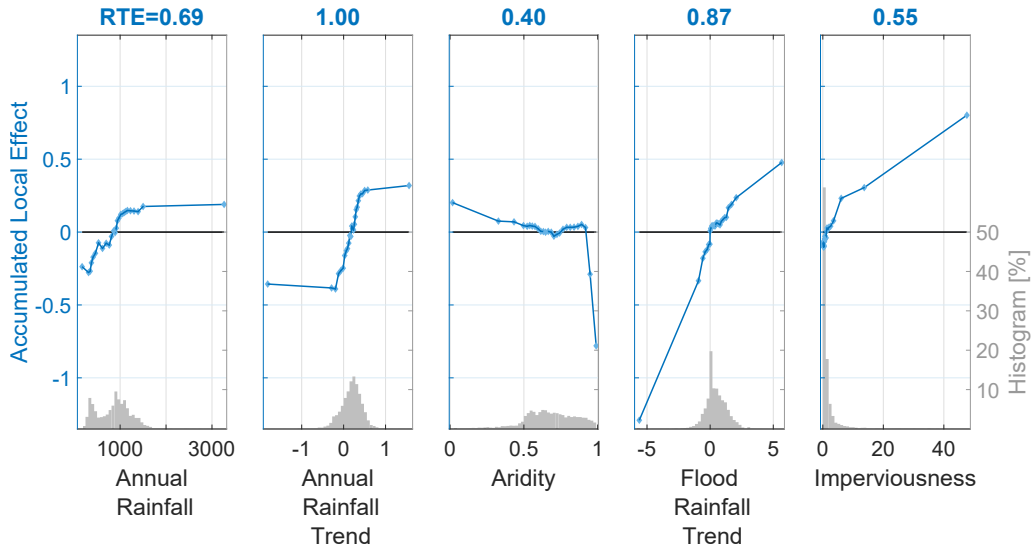

Figure 2.6: Accumulated Local Effects for the Predictors of the Global Random Forest. Blue lines show the ALE values for the five variables that were selected in the CONUS-wide Random Forest. Grey bars show a histogram of each variable. The trends in flood generating and annual rainfall are the most important predictors of flood magnitude trends at the CONUS level. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

2.5.4 Uncertainty Assessment

All climatological variables are based on a single reanalysis dataset (Livneh et al. 2013). We note that this dataset was not initially calibrated to study flood events and therefore might not represent all aspects of flood generation correctly. However, previous studies have successfully used the dataset to model floods and high flows in Texas (Zhao et al. 2016) and across the CONUS (Oudin et al. 2018). The fact that the controls we found in many clusters fit well to results from previous regional studies (see Section 2.4.3) further supports the feasibility of the reanalysis data for this study. Nevertheless, all results have to be seen with this uncertainty in mind, meaning that variable importance should be viewed in relative, rather than absolute terms.

The flood classification approach we are using is a simplification of the real world in several ways. Firstly, we only consider five different types of flood generation and treat them as distinct even though real world floods are always the result

of the interplay of water from multiple sources. This degree of simplification is in line with most flood type classifications (Berghuijs et al. 2016; Tarasova et al. 2019). More nuanced classifications exist (e.g. Tarasova et al. 2020), but the aim of this study – to investigate a large, heterogeneous region with the same set of variables – required a simpler approach. Furthermore, the calculation of trends in flood types requires a certain number of events of a given type at each station within the study period. A larger number of flood types would therefore decrease the chance of gaining a sufficient number of events per type. Secondly, the thresholds in the classification scheme are to some extent arbitrary, even though they are based on previous studies (Kemter et al. 2020; Stein et al. 2020). A sensitivity analysis (not shown here) to small changes in these thresholds showed no substantial impact on our main results. Thirdly, the classification is based on a simple representation of catchment concentration time t_c. While more realistic representations of t_c are available (Tarasova et al. 2019), which take into account the flood generating process, we decided to use the approach we present here. We did so because one aim of t_c in this study is to detect the flood generating process itself, which excludes it from the information that could be used to calculate t_c. Furthermore, we did not consider the variation of t_c between AMS events by analysing event hydrographs. This would complicate the comparison and trend calculation of flood contributions from different water sources (rain, snowmelt, soil moisture), which we also calculate based on t_c. In addition, such an approach would not be possible for the smaller, fast-responding catchments in the dataset, as our analysis is based on daily streamflow data.

Sensitivity analysis (not shown here) of the clustering approach shows that it can be prone to shifting cluster assignment due to changes of connections in the complex network, leading to a shift of a group of catchments from one class to another or the merging or splitting of certain clusters. However, we found that, apart from these occasional shifts of small groups, the boundaries between clusters are stable. We furthermore found, that these changes in cluster assignment do not affect the main results of the Random Forests substantially: the ratio of predictor selections between the categories is stable, the most important predictors are always similar, and in cases when groups are transferred to other clusters the

importance of a predictor with which this group is highly associated migrates with it to the new cluster.

Given the sparse data availability for land cover changes (only 5 to 6 time steps from 1960 to 2010) and no information about canopy cover or regulation/diversion changes, the studied effects of land cover changes carry high uncertainties. The fact that static land cover variables were selected in 7 out of 10 clusters suggests that their trends – which were not selected at all – should also be of some relevance for flood magnitude trends. Trends in land use should be assessed in future studies focusing on more recent decades for which more data are avaialable from remote sensing products.

2.6 Conclusions

Through flood behavior clustering and Random Forests, we found that hydro-climatological and anthropogenic factors have both substantially contributed to past changes in AMS magnitudes across the CONUS. Their similar importance highlights the necessity to represent land cover and water regulation as well as their potential changes in models of future flood risk, which so far focus mainly on climatic changes. Furthermore, it suggests that land use change and land use planning can be crucial aspects of flood risk management, for instance through reforestation. The land use factors that have been found as important controls on flood magnitude trends are always static indicators (for imperviousness and cropland) or the available information does not allow to quantify whether the factor has changed within the study period (for regulation/diversion and canopy cover). The result that static land use factors are more important than their trends carries some uncertainty given the sparse data availability in the study period, but suggests that land use is able to attenuate (in the case of forested areas) or amplify (in the case of urbanized areas) the effects of climatic changes on flood magnitudes. The adequate consideration of both static and dynamic land cover variables in planning will require additional, dedicated studies, as flood magnitude trend controls can have non-linear effects – like in the case of imperviousness, for which in four clusters catchments beyond a certain degree of urbanization have witnessed significantly higher flood magnitude trends than those below the threshold. Therefore, land cover

changes and effects should be studied more thoroughly in hydrologic modelling research.

Through a novel approach to Random Forest interpretation (ALE plots), we were able to show how the different controls influence the flood magnitude trends. We identified variables that generally had a negative or positive effect on flood magnitude trends independent of the cluster, while other variables showed opposing effects for different regions due to the diverse hydro-climatological conditions. Such regionally varying effects of flood trend controls can remain undetected in continental scale studies, potentially leading to incomplete interpretations. The ALE plots furthermore allowed us to uncover non-monotonic relationships between catchment variables and flood magnitude trends.

The clustered Random Forests have a higher predictive ability than a global Random Forest trained on the same data. As the clustered Random Forests focus on attributing flood magnitude trends within the clusters, the global Random Forest model is nonetheless useful to identify the controls on inter-cluster variability, i.e., what controls the sign of the trend at the large scale. Our results show that a combination of a clustering approach and a global model is necessary to fully grasp controls on flood trends at a continental scale without neglecting regional differences.

Our findings highlight the importance of a holistic view on flood magnitude trends, as hydro-climatological and land cover conditions can regionally interact in complex ways, which can go unnoticed when entire continents are studied or only certain aspects, subregions, or catchment types are taken into account. Models of future changes in flood risk will have to consider these interactions to fully capture the interconnected nature of flood generation and more reliably project trends.

2.7 Acknowledgements and Data Availability

This research was funded by the DFG Research Training Group “Natural Hazards and Risks in a Changing World" (NatRiskChange GRK 2043). NM acknowledges funding by the BMBF project climXtreme (01LP1902J). GV acknowledges funding by the USACE Water Institute. USGS peak streamflow available at https://nwis.waterdata.usgs.gov/usa/nwis/peak. The USGS Streamgage NHDPlus

catchment shapes available at https://water.usgs.gov/GIS/metadata/usgswrd/XML/streamgagebasins.xml. Livneh data provided by the NOAA/OAR/ESRLPSLathttps://psl.noaa.gov. The GAGES-II Catchment attribute data is available at https://www.sciencebase.gov/catalog/item/59692a64e4b0d1f9f05fbd39.

2.A Appendix

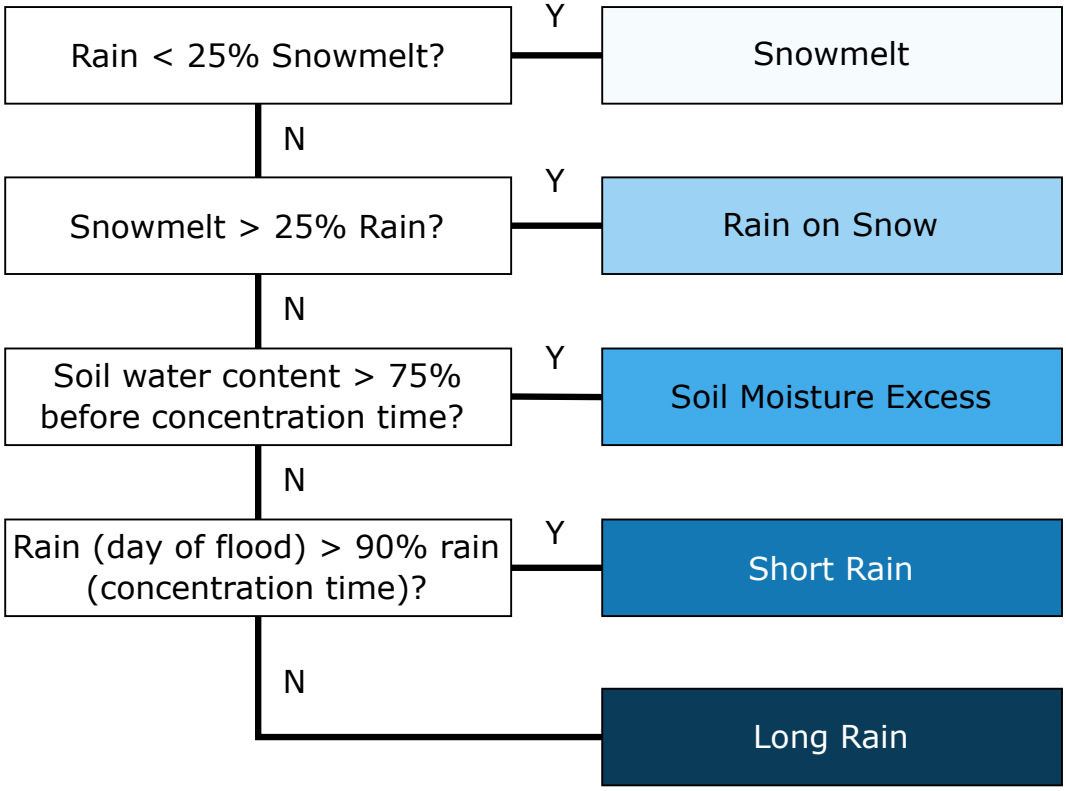

Figure 2.A1: Flood Process Classification. Based on antecedent catchment conditions, each flood is classified as one of the five flood types on the right.

Table 2.A1: Catchment-related Climate, Flood and Land Use Variables used as Potential Controls on Flood Trends. Adimensional quantities are identified as "1."

	Variable	Unit	Time Period	Data Source	Description
Climate	Mean Annual Rainfall	mm	1960-2010	Livneh	Catchment-averaged total annual rainfall averaged over 1960-2010. Rainfall (for all variables) is calculated as the difference between precipitation and snowfall (estimated from net daily gain in snow water equivalent)
	Annual Rainfall Trend	%/yr	1960-2010	Livneh	Trend of annual rainfall sum relative to the mean annual rainfall sum. Estimated using the Theil-Sen estimator (Sen 1968).
	Mean Annual Snowfall	mm	1960-2010	Livneh	Catchment-averaged total annual snowfall averaged over 1960-2010.
	Annual Snowfall Trend	mm/yr	1960-2010	Livneh	Trend of annual snowfall sum estimated using Theil-Sen estimator.
	Mean Annual Soil Moisture	%	1960-2010	Livneh	Catchment-averaged mean annual soil moisture relative to the range between maximum and minimum values of 1960-2010.
	Annual Soil Moisture Trend	%/yr	1960-2010	Livneh	Trend of annual mean soil moisture relative to the mean annual soil moisture of 1960-2010. Estimated using the Theil-Sen estimator.
	Precipitation Seasonality	1	1960-2010	Livneh	Seasonality index SI based on values from 1960-2010, with SI defined as (Walsh and Lawler 1981): $SI = \frac{1}{R}\sum_{n=1}^{12}\lvert X_n - \frac{\overline{R}}{12}\rvert$, where $\overline{R}$ is the mean annual precipitation and X_n is the mean monthly precipitation of month n. 0=all months receive the same amount of precipitation, 1.83=all precipitation occurs in one month.
	Precipitation Seasonality Trend	%/yr	1960-2010	Livneh	Trend of annual values for the precipitation seasonality estimated using Theil-Sen estimator. Relative to the average precipitation seasonality of 1960-2010.
	Aridity	1	1960-2010	Livneh	The ratio of mean annual total evapotranspiration and mean annual precipitation. 1=very arid, 0=very humid.

Topo	Catchment Area	km²	-	USGS AMS	Catchment area transformed to km². Base-ten logarithm used for visualization.
Flood Characteristics	Mean Flood Magnitude	m³/s	1960-2010	USGS AMS	Mean of AMS time series. Base-ten logarithm used for visualization.
	Flood Date Trend	days/yr	1960-2010	USGS AMS	Trend of AMS timing calculated with the Theil-Sen estimator adapted to circular date values (Blöschl et al. 2017).
	Flood Generating Rainfall Trend	%/dec	1960-2010	Livneh	Trend of the catchment-averaged rainfall sum in the concentration time t_c (see Section 2.3.1) before each flood, relative to the mean 1960-2010. Estimated using Theil-Sen estimator.
	Flood Generating Snowmelt Trend	%/dec	1960-2010	Livneh	Trend of the catchment-averaged snowmelt sum in the concentration time t_c before each flood, relative to the mean of 1960-2010. Estimated using Theil-Sen estimator.
	Flood Generating Soil Moisture Trend	%/dec	1960-2010	Livneh	Trend of the catchment-averaged soil moisture in the concentration time t_c before each flood, relative to the mean of 1960-2010. Estimated using Theil-Sen estimator.
Flood Type	Long Rain Relevance	%	1960-2010	Livneh	Fraction of floods classified as generated by long rain.
	Short Rain Relevance	%	1960-2010	Livneh	Fraction of floods classified as generated by short rain.
	Snowmelt Relevance	%	1960-2010	Livneh	Fraction of floods classified as generated by snowmelt.
	Rain on Snow Relevance	%	1960-2010	Livneh	Fraction of floods classified as generated by rain on snow.
	Soil Moisture Relevance	%	1960-2010	Livneh	Fraction of floods classified as generated by soil moisture.
	Trends of Type Relevancies	%/yr	1960-2010	Livneh	Trend of flood type relevance for the five types above for all stations with at least five occurrences of a given type. Estimated using Poisson regression.

Land Use				
Canopy Cover	%	2011	GAGES	Fraction of the catchment area covered with tree canopy derived from the National Land Cover Dataset (NLCD) which is based on Landsat5 remote sensing data and a Random Forest model (Homer et al. 2015).
Imperviousness	%	1974-2012	GAGES	Mean of the fraction of the catchment area covered with impervious surfaces for the years 1974, 1982, 1992, 2002, and 2012. Derived from the Wall-to-Wall Anthropogenic Land Use Trends (NWALT) dataset, based on land use classes with 60m resolution (Falcone 2017).
Imperviousness Trend	%/yr	1974-2012	GAGES	Trend of the five values of imperviousness estimated using Theil-Sen estimator.
Cropland	%	1964-2012	GAGES	Mean of the fraction of the catchment area covered with cropland for the years 1964, 1974, 1982, 1992, 2002, and 2012. Derived from the NWALT dataset, based on land use data at 1km resolution (Falcone and LaMotte 2016).
Cropland Trend	%/yr	1964-2012	GAGES	Trend of the six values of catchment-averaged cropland estimated using Theil-Sen estimator.
Regulation/ Diversion	1	1960-2010	GAGES	Direct human influence leading to reduced streamflow through regulation or diversion, based on National Water Information System peak flow qualification codes. Stations with a code #5 “Discharge affected to unknown degree by Regulation or Diversion” or #6 "Discharge affected by Regulation or Diversion" in at least ten years between 1960 and 2010 are marked as regulated (1), all others as unregulated (0).

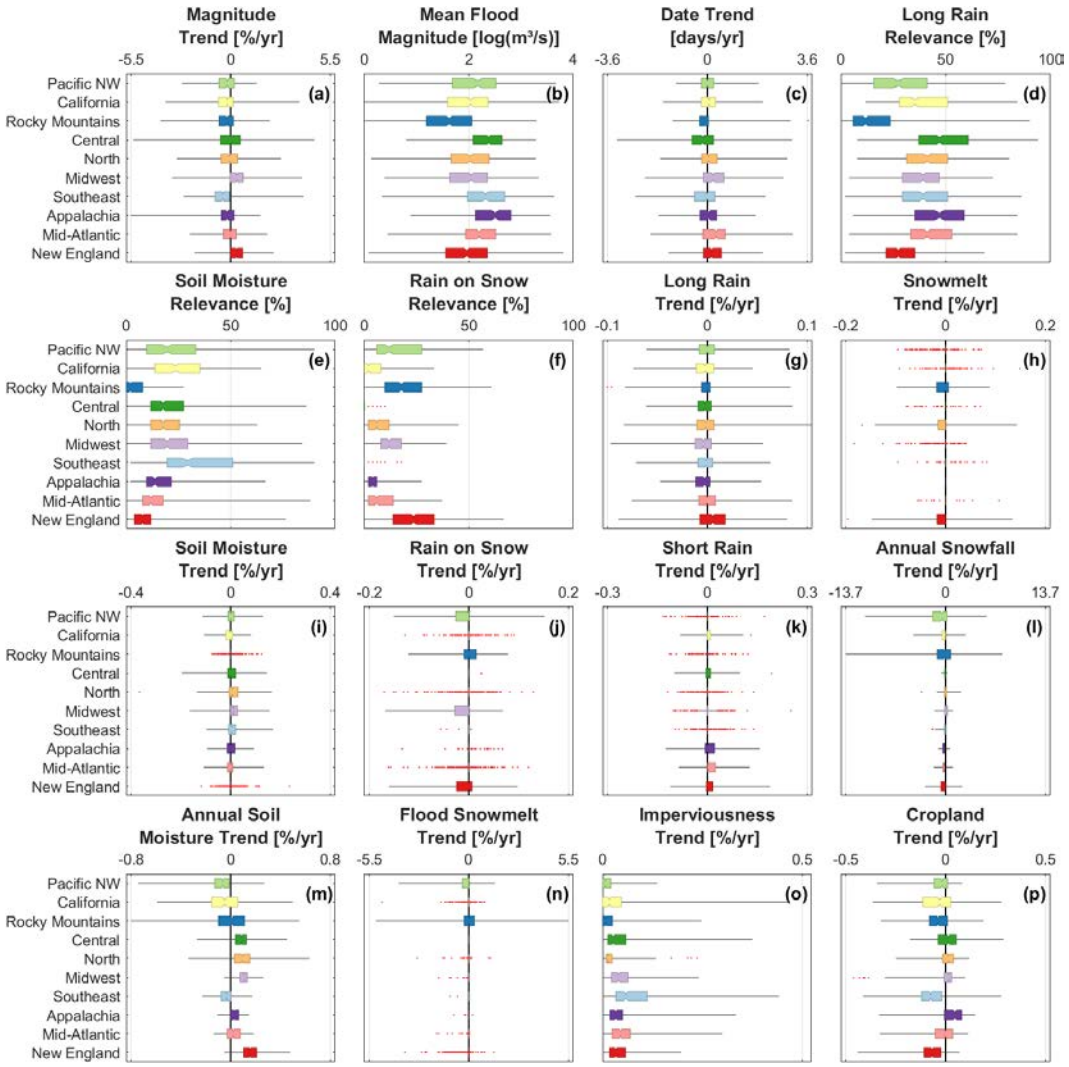

Figure 2.A2: Distribution of unselected variables. Distribution of flood magnitude trends (a) and the set of 15 variables, which were not selected in the predictor selection process grouped by clusters (p-t). Panel titles indicate the values depicted on the x-axes. Boxes range from the 25^{th} to the 75^{th} percentile with a notch indicating the median. Whiskers reach to ten times the interquartile range or to the furthest data point, if the latter is closer. Outliers marked in red.

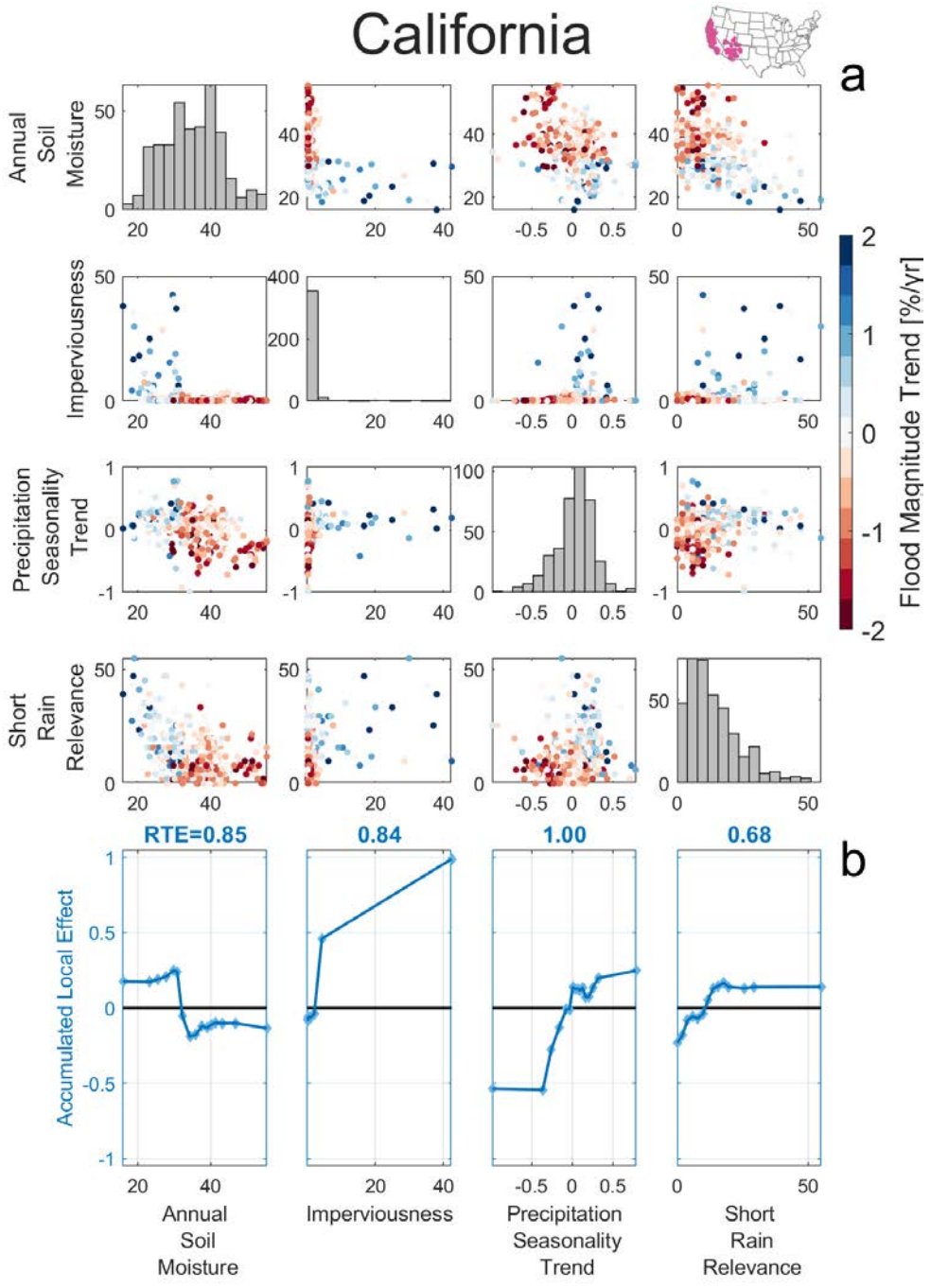

Figure 2.A3: Predictor Effects for the California Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges. Cluster color in the map changed to improve visibility on white background.

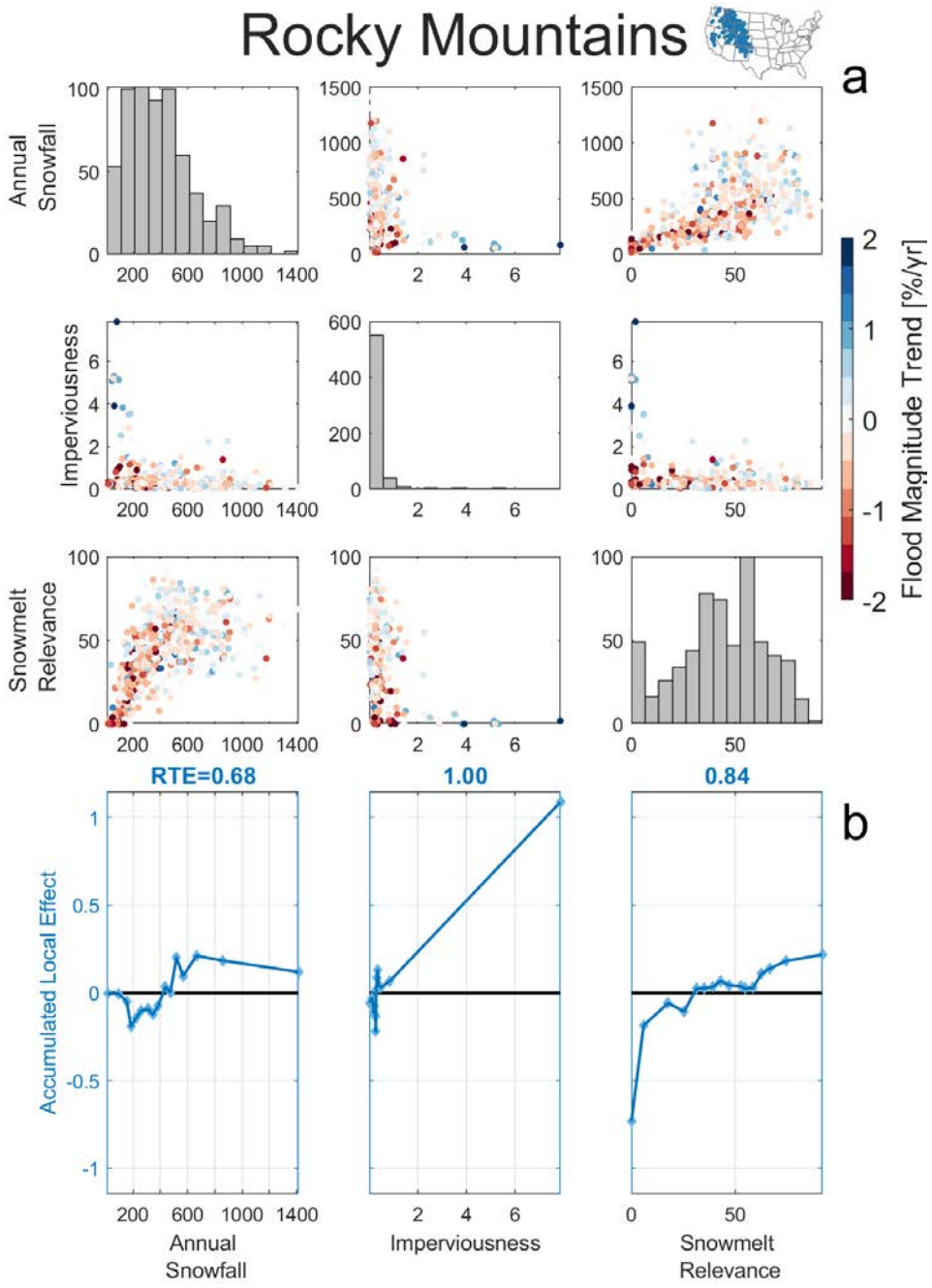

Figure 2.A4: Predictor Effects for the Rocky Mountains Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

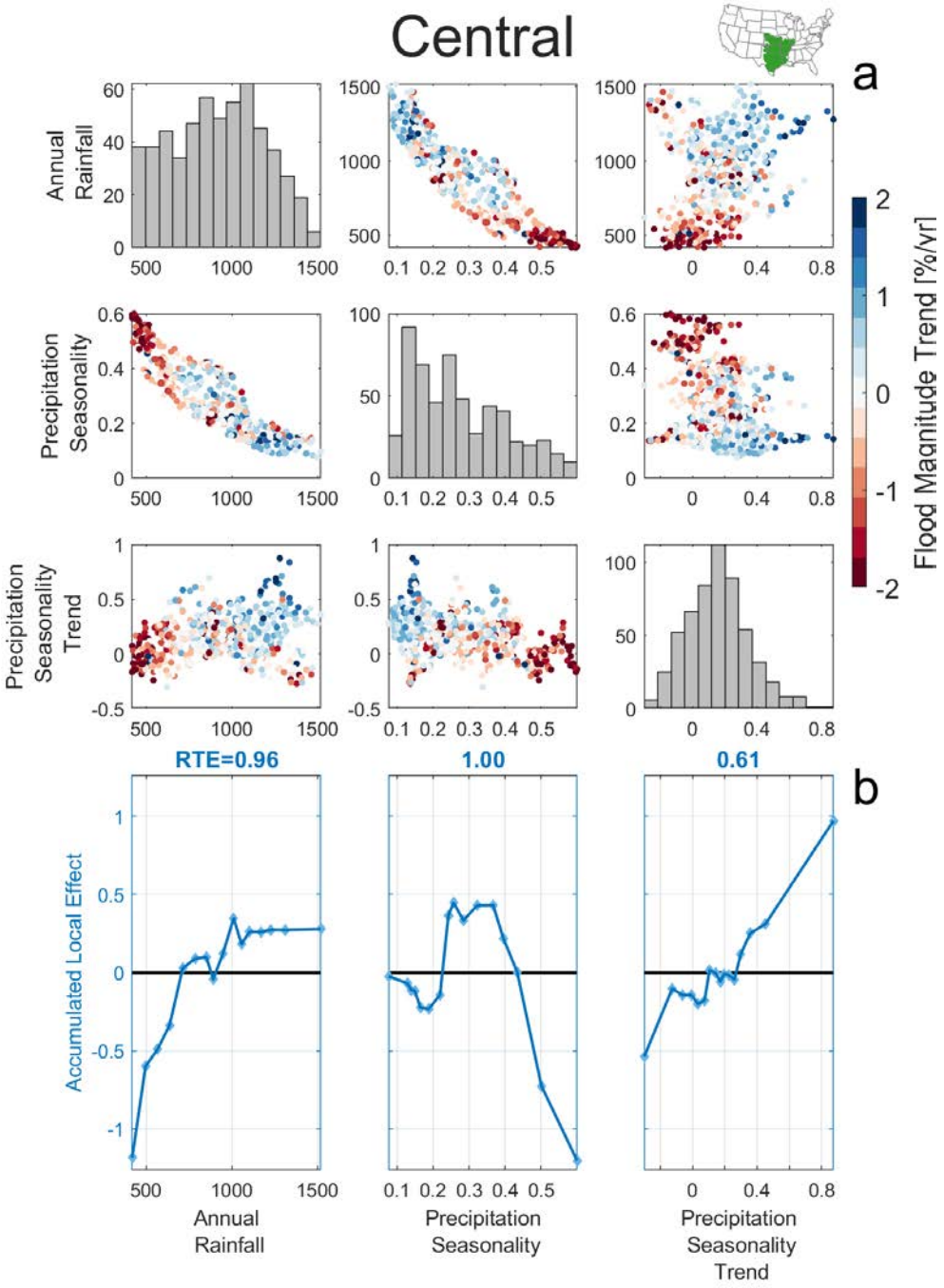

Figure 2.A5: Predictor Effects for the Central Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

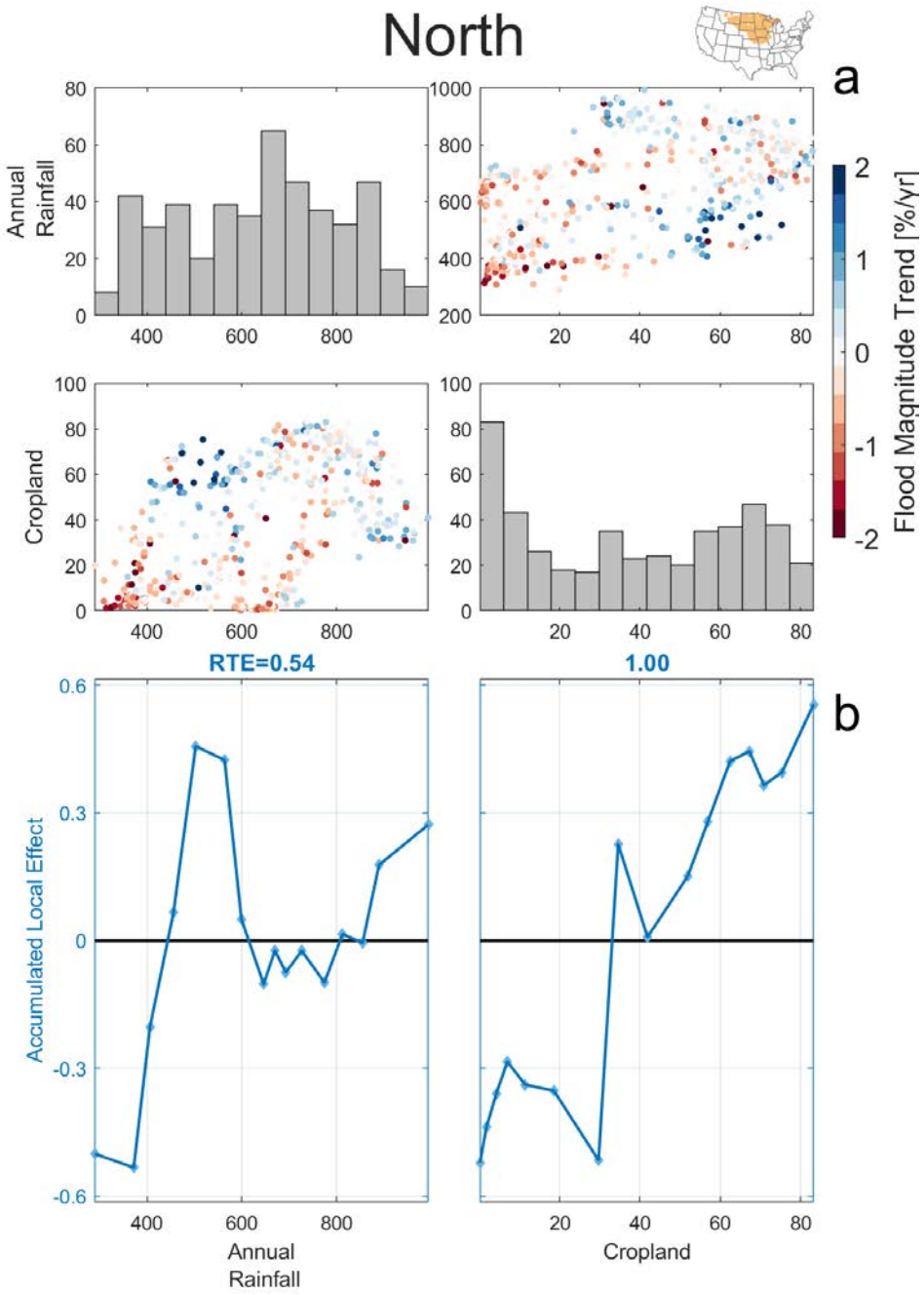

Figure 2.A6: Predictor Effects for the North Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

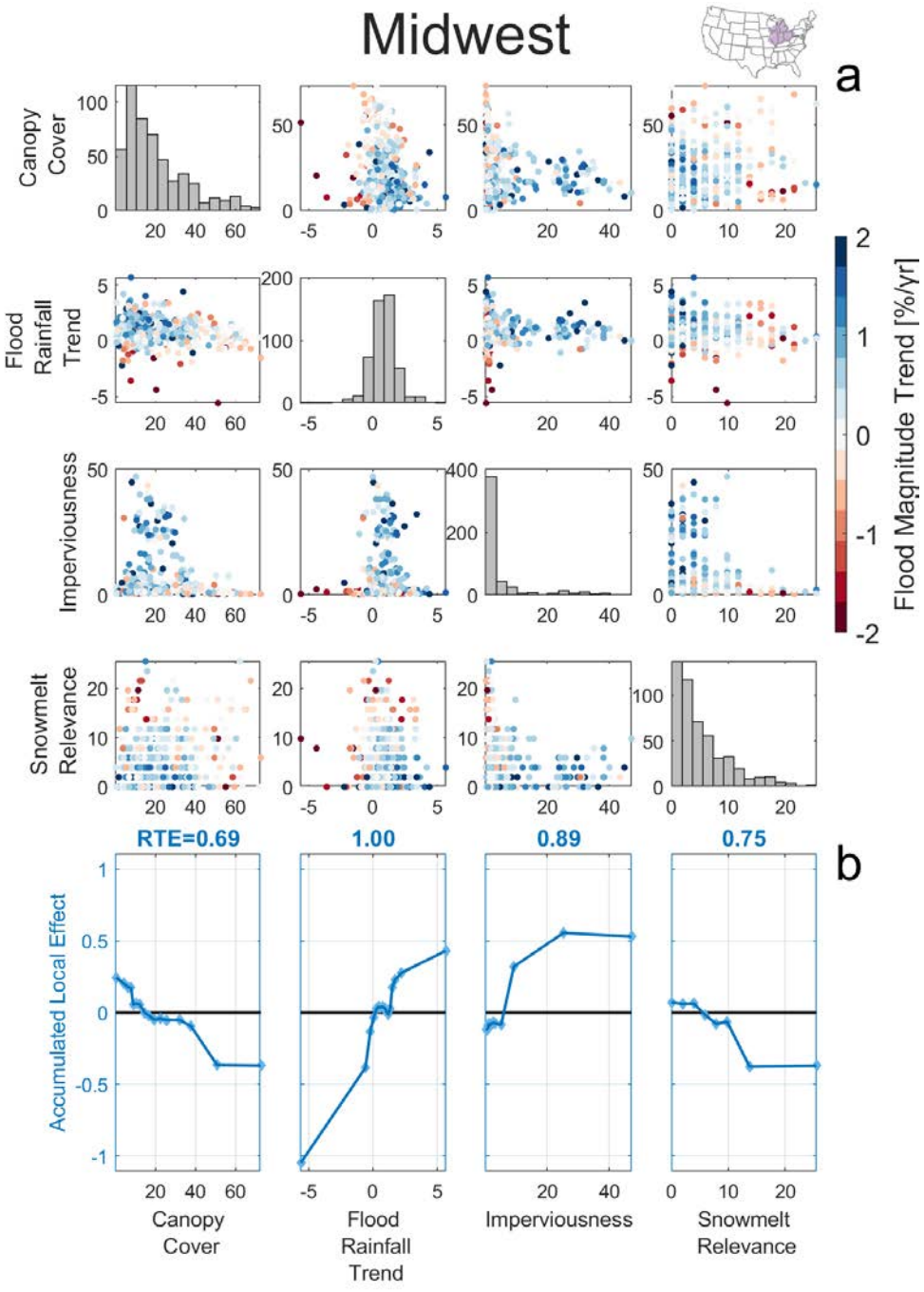

Figure 2.A7: Predictor Effects for the Midwest Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

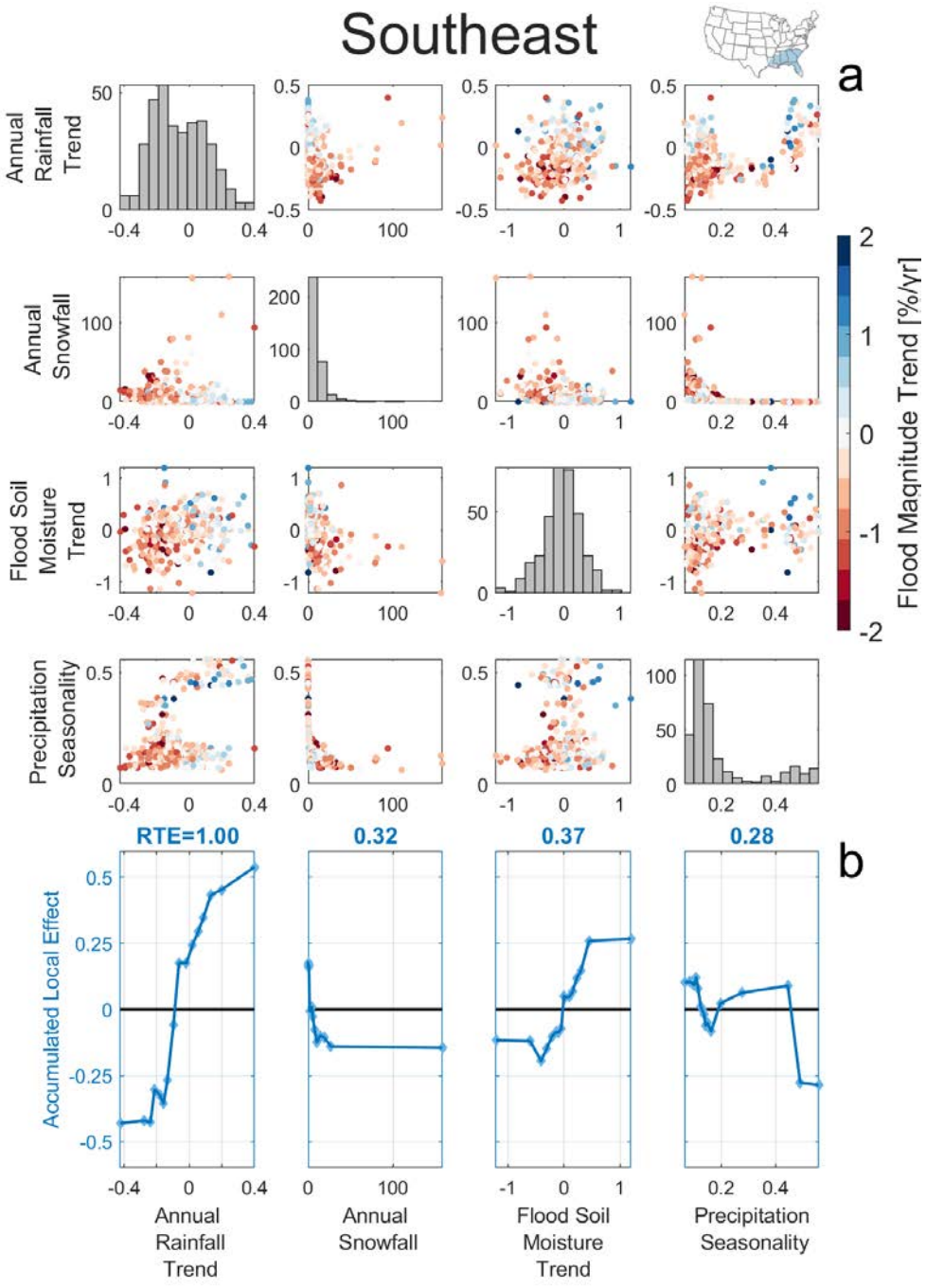

Figure 2.A8: Predictor Effects for the Southeast Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

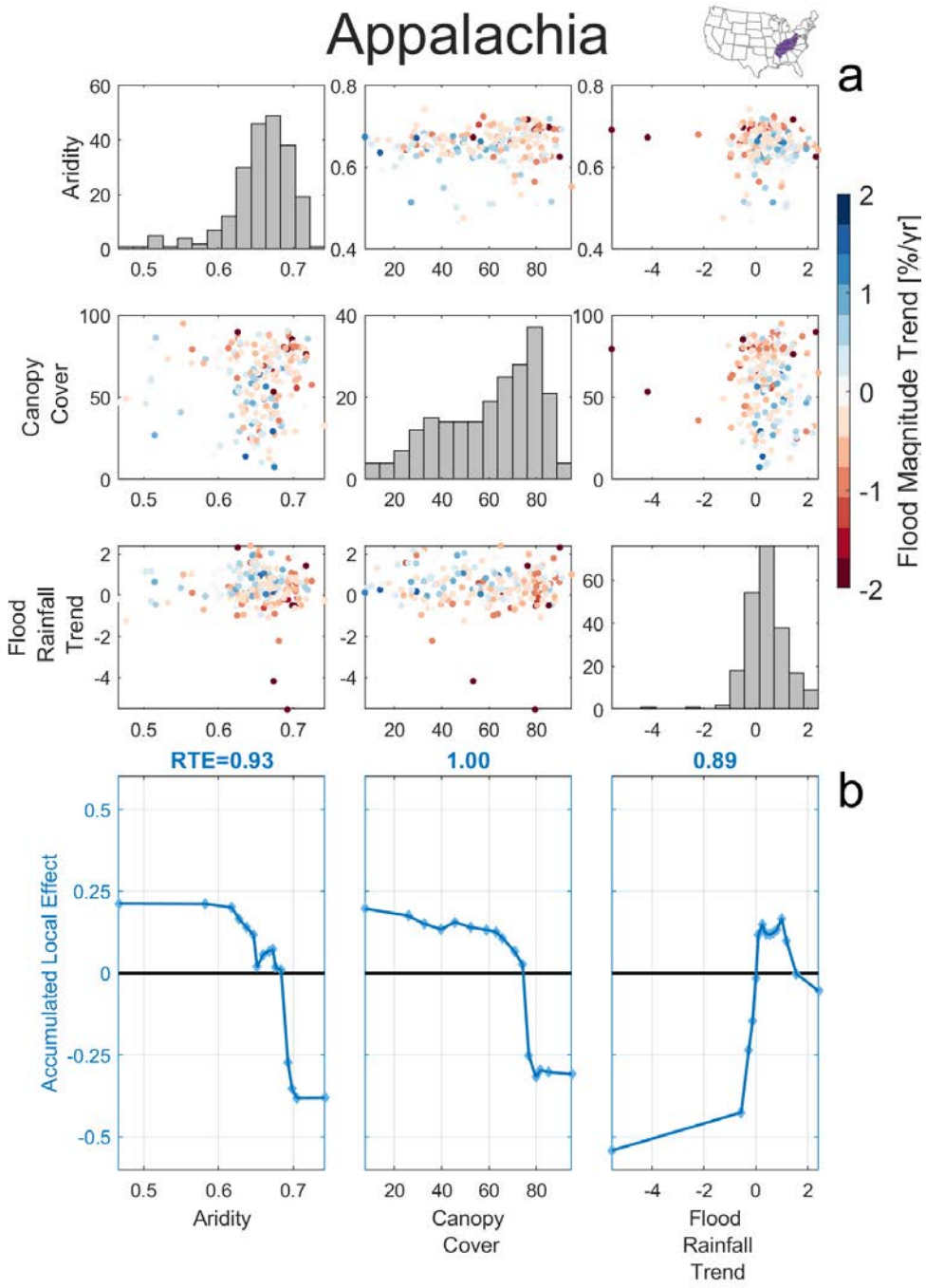

Figure 2.A9: Predictor Effects for the Appalachia Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

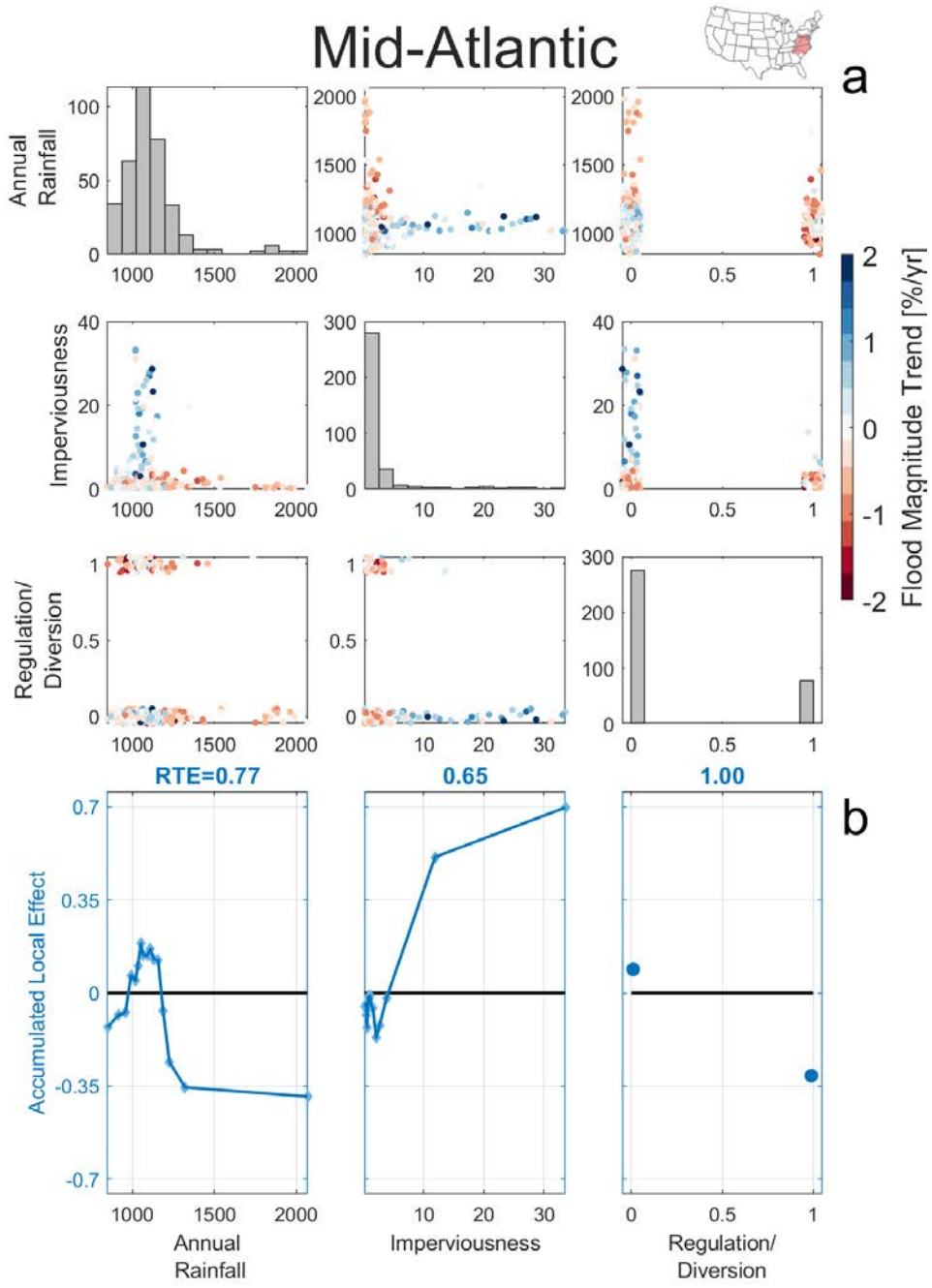

Figure 2.A10: Predictor Effects for the Mid-Atlantic Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

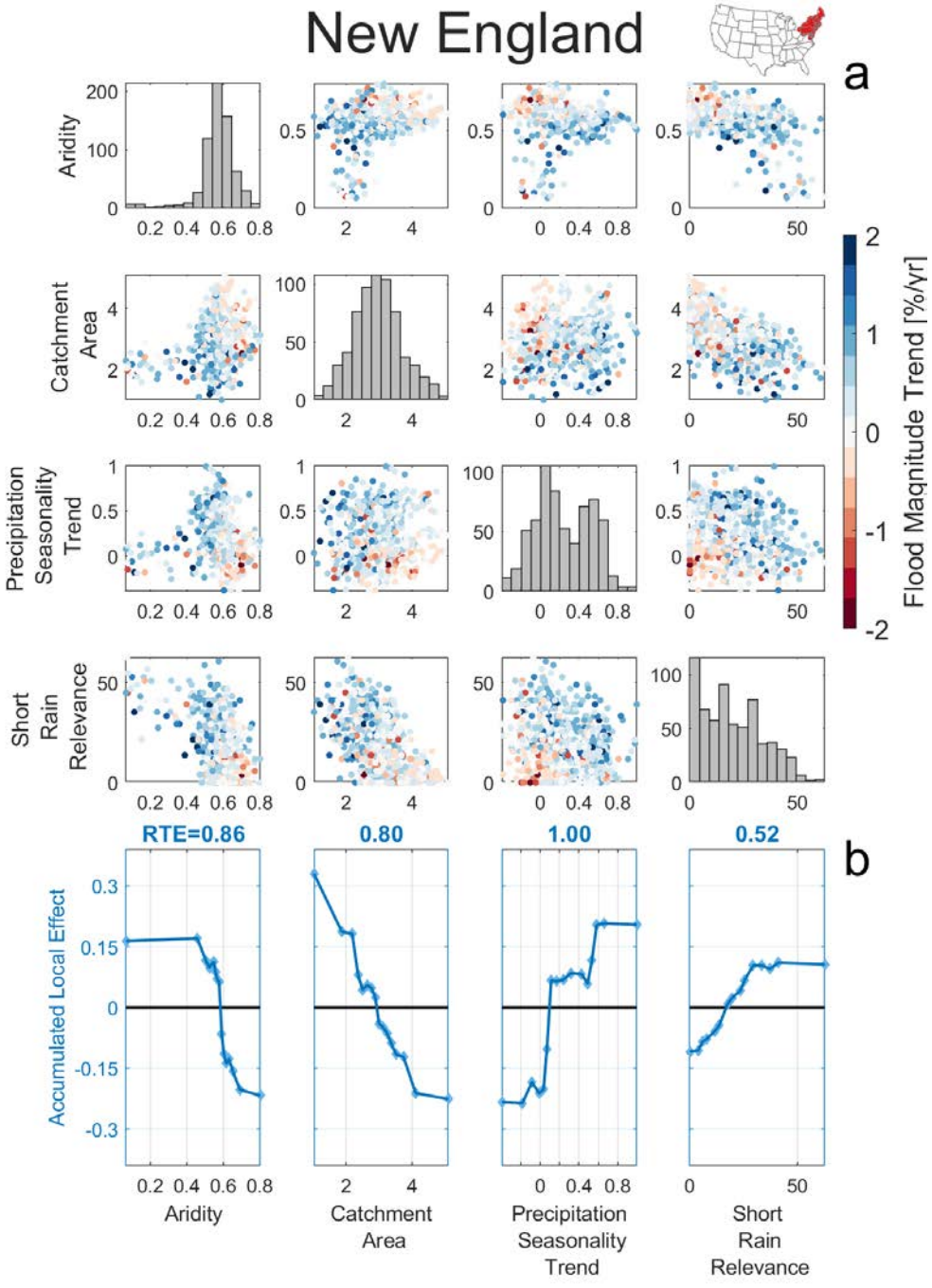

Figure 2.A11: Predictor Effects for the New England Cluster. a) Bivariate scatter plots of the selected predictors. Red (blue) dots indicate decreasing (increasing) flood magnitude trends of a given catchment. b) Accumulated Local Effect plots for the five predictors in blue. RTE=relative total effect, i.e. the mean absolute ALE divided by the maximum mean absolute ALE of the cluster. Diamonds mark the quantile edges for which ALEs were calculated, showing the variable quantile ranges.

3 Joint Trends in Flood Magnitudes and Spatial Extents across Europe

Published in Geophysical Research Letters: Kemter, M., Merz, B., Marwan, N., Vorogushyn, S. & Blöschl, G. (2020). Joint Trends in Flood Magnitudes and Spatial Extents across Europe. 46, e2020GL087464, https://doi.org/10.1029/2020GL087464

Abstract

The magnitudes of river floods in Europe have been observed to change but their alignment with changes in the spatial coverage or extent of individual floods has not been clear. We analyze flood magnitudes and extents for 3872 hydrometric stations across Europe over the past five decades and classify each flood based on antecedent weather conditions. We find positive correlations between flood magnitudes and extents for 95% of the stations. In Central Europe and the British Isles, the association of increasing trends in magnitudes and extents is due to a magnitude-extent correlation of precipitation and soil moisture along with a shift in the flood generating processes. The alignment of trends in flood magnitudes and extents highlights the increasing importance of transnational flood risk management.

Plain Language Summary

If multiple rivers flood at the same time because of large scale rainfall, the resulting damage can exceed the capacities of disaster recovery and insurance companies. We find that events with a large spatial coverage or extent tend to be associated with above average magnitudes of the flooding level. During 1960-2010 flood extents increased in Central Europe and the British Isles, but decreased in Eastern Europe. These trends are caused by changes in flood generation processes due to a changing climate. If these trends persist into the future, the combination of stronger floods and larger extents is likely to increase the flood risk substantially.

3.1 Introduction

River floods are among the most harmful natural hazards worldwide and their damages are expected to increase further as a consequence of climate change, population and economic growth and rising economic interdependence (Dottori et al. 2018; Field et al. 2012; Kundzewicz et al. 2014; UNDRR 2019). In Europe, trends in flood magnitudes have been identified (Blöschl et al. 2019; Jongman et al. 2014; Mangini et al. 2018). These trends vary in space because of differences in the flood generating processes. For example, increasing autumn and winter rainfall has resulted in increasing floods in North-Western Europe while decreasing snowmelt has led to decreasing floods in Eastern Europe in the past five decades (Blöschl et al. 2017; Mangini et al. 2018). If a flood event covers a large region, emergency response, disaster recovery and the insurance industry may be overtaxed, as resources and funds need to be provided at many locations at the same time (Jongman et al. 2014). In Europe, the flood extent, i.e. the area or distance over which flooding occurs simultaneously, has been found to change (Berghuijs et al. 2019), but the alignment of these changes with changes in the flood magnitudes has not been studied. An alignment of flood magnitude and flood extent trends has the potential of increasing the flood risk beyond the effects of the individual trends. If there are clear physical causes, the alignment may translate into the future.

3.2 Materials and Methods

3.2.1 Data

We use a flood dataset consisting of the timing and magnitude of annual maximum discharge for 5245 hydrometric stations in Europe (Blöschl et al. 2019). We choose a timeframe from 1960-2010 to keep the number of available stations for each year relatively constant over time and select only those stations with at least 30 years of data. The selection resulted in a total of 3872 stations with catchment sizes ranging from 1 to 800,000 km^2. The median catchment size is 312 km^2. In order to examine the process controls on floods, we use reanalysis data (Primo et al. 2019) with a spatial resolution of 0.11x0.11°. The variables are precipitation, snowfall, soil

moisture (46cm depth), soil pore space, snowmelt, convective available potential energy (CAPE) and convective inhibition (CIN). The original temporal resolution is 1 hour (precipitation, snowfall), 3 hours (CAPE, CIN) and one day (soil moisture, snowmelt). We aggregate all variables to daily totals. These data are used for two analyses, (i) pixel based magnitude-extent correlations, and (ii) catchment based identification of flood generation processes. For the latter we derive catchment boundaries, using the CCM (Vogt et al. 2007) and MERIT Hydro (Yamazaki et al. 2019) datasets to calculate catchment average time series of these variables. The daily time series are calculated by weighted averages of the pixels at each time step, where the weight of each pixel is set according to the fraction of its area covered by the catchment area.

3.2.2 Flood Synchrony Scale

We quantify the spatial extent of flood events by the flood synchrony scale (Berghuijs et al. 2019). It is defined as the maximum distance from a station within which at least 50% of the stations have the annual maximum flood discharge at the same time as the reference station. We allow for a time delay Δt of ±7 days in order to account for the travel time of weather patterns to move across Europe and flood routing in the river system. Therefore, the flood synchrony scale FS of a station i in year j is defined as:

$$FS(i,j) = max\{d[f(d) > 0.5]\}, \tag{3.1}$$

where f is the fraction of stations within distance d where the annual maxima occurred within the allowed time delay Δt:

$$f(d) = \frac{1}{n(d)} \sum_{i=1}^{n(d)} (t_{ref} - \Delta t < x_i < t_{ref} + \Delta t). \tag{3.2}$$

Here $n(d)$ is the number of stations within distance d from the reference station, t_{ref} is the day of the flood at the reference station and x_i is the day of the flood at the other stations. We estimate trends of the flood synchrony scale for each station using the Theil-Sen slope estimator (Sen 1968). Trends are averaged

following 30,000 random station resampling iterations, to minimize the effect of heterogeneous station density (Figure 3.A1). We then interpolate the trends spatially using the *autoKrige* function of the *R automap* package (Hiemstra et al. 2009) to obtain regional trends.

3.2.3 Magnitude Extent Correlation

We estimate the Spearman rank correlation coefficients between the annual series of flood magnitude and flood extent in terms of the flood synchrony scale (Equation 3.1) for each station. We test the significance of the correlation at the 5% level, adjusting the False Discovery Rate of multiple hypothesis testing using the Benjamini-Hochberg correction (Benjamini and Hochberg 1995). Additionally, we evaluate the analogous correlations for precipitation, soil moisture and snowmelt. In this case, we use the gridded dataset and calculate a time series of extent for each pixel by the same method as for the flood peaks. For precipitation, soil moisture and snowmelt a day is considered an event if it exceeds the 99% quantile. We compare pixels on the same day (Δt=0) because no river routing is involved. Furthermore, we repeat the analysis for measured satellite and station based precipitation data using PERSIANN (Ashouri et al. 2015) and ECA&D data (Klein Tank et al. 2002), respectively. Again, we evaluate Spearman rank correlation coefficients and apply a Benjamini-Hochberg correction for the significance tests.

3.2.4 Flood Classification

We classify the total of about 174,000 flood peaks by their dominant flood generating processes. Flood generation can be dominated by the hydrometeorological forcing (rainfall and snowmelt) as well as by the antecedent catchment state (soil moisture and snow cover). We therefore consider the following flood generating processes: convective precipitation, stratiform rainfall, soil moisture excess, snowmelt and rain-on-snow. We use the catchment average time series of the climate variables for the classification along with a simple decision tree (Figure 3.A2). As a first step, we subtract snowfall from precipitation to calculate rainfall. We estimate the concentration time t_c of each catchment by $t_c = \alpha A^{\beta}$ where $\alpha = 0.1$ and $\beta = 0.3$ with t_c in units days and A in units km² (CorCorradini et al. 1995; Robinson and

Sivapalan 1997). For 3,191 stations (82%) t_c = 1 day and for only 39 stations (1%) $t_c \geq 4$ days. We consider the climate variables on the day of the flood peak as well as t_c days prior to it for the classification. If snowmelt was greater than rainfall, we classify the flood generating process as snowmelt. If snowmelt was less than rainfall but most of the catchment (>66% of the area) was covered by any amount of snow, we classify the flood as rain-on-snow. If this was not the case, we check for convective conditions in the catchment. We detect these by using thresholds (Findell and Eltahir 2003) for CAPE (>400) and CIN (<5), and assign convective rainfall as a generating process, if at least 25% of the catchment area had convective conditions on any day during t_c. If none of the above conditions applied, we check whether the soil water content exceeded 70% of the available pore space on the day prior to t_c. If this was the case, we classify the flood as soil moisture excess related. All floods that did not meet any of the above criteria are considered to be mainly caused by stratiform rainfall. While floods are often caused by the interplay of different parameters, we only classify them by the dominant process for simplicity and clarity. The thresholds for CAPE and CIN we use here do not guarantee convective conditions, as high vertical wind shear values can also be necessary (Gilleland et al. 2016). As this parameter was not available to us, it is likely that the classification overestimates the frequency of convective rain.

While the classification concept used here is simple, it allows for a fast automatic classification of a large number of flood peaks based on widely available climate data. It considers flood generation processes beyond the timing of the flood within the year and therefore may provide more detailed information than timing-based classifications (Mediero et al. 2015; Tarasova et al. 2019). For instance, it never classifies a flood as snowmelt related unless there was a substantial amount of snow present in the catchment, whereas timing-based methods do not make this distinction. To check the plausibility of the classification results we examine the temporal distribution of the flood generation processes across the year.

For each catchment, we estimate the relative frequencies of flood peaks caused by a process (termed relevance), assigning 1 if a flood peak was associated with a process in a given year and 0 if it was not. We calculate their trends in the period 1960-2010 by the Theil-Sen estimator and spatially interpolate both the relevance

and its trend by kriging using *autoKrige* as before. Finally, we estimate the regional diversity (D_j) of flood generation processes by:

$$D_j = 1 - var(f_{m,j}), \tag{3.3}$$

where j is the year, m is the flood generation process (1...5) and $f_{m,j}$ is the relative frequency of process m in year j. We calculate D_j for two regions, one in Western Europe where flood magnitudes have increased and one in Eastern Europe where they have decreased (Blöschl et al. 2019). For comparison, we estimate the spatial variance T_j of the flood dates by the circular variance:

$$T_j = 1 - \sqrt{\overline{cos(\theta_{i,j})}^2 + \overline{sin(\theta_{i,j})}^2}, \tag{3.4}$$

where θ is the flood date converted into an angle between 0 (January 1) and 2π (December 31), i is the station and j is the year (Table 3.A1).

3.3 Results

Our data show significant positive correlations between annual flood magnitudes and their flood synchrony scale for 46.5% (N=1,802) of the stations (Figure 3.1). We find significant negative correlations for only 0.1% (N=4) of the stations. A total of 95.2% (N=3,685) and 4.8% (N=187) of the stations exhibit positive and negative correlations, respectively. The positive correlations are highly consistent across Europe. This is related to the similar correlations of three controls on floods (Figures 3.A3 to 3.A5). Specifically, the average correlations between magnitudes and the corresponding synchrony scales of precipitation, soil moisture excess and snowmelt are 0.44, 0.54 and 0.52, respectively (Figure 3.1 inset), while the average correlation for floods is 0.32. The presence of magnitude-extent relationships of the controls suggests that they propagate to the floods. The spatial extent of localized convective storms has been previously shown to increase with precipitation magnitudes (Lochbihler et al. 2017; Molnar et al. 2015) and this analysis suggests that this is also the case at the regional scale. For a given weather system velocity, large scale precipitation events tend to produce longer rainfall than short scale events, which appears to translate into higher daily precipitation (Skøien et al. 2003). Similarly,

the moisture content of soils close to saturation, which often occur in winter, tends to be more homogeneous over larger regions than that of soils if average moisture content (Pachepsky et al. 2003). Snowmelt (Figure 3.A5) also shows mostly positive correlations, because snowmelt will occur over large regions during events with high temperatures. The correlation values of precipitation extents and magnitudes for the PERSIANN and ECA&D data are similar to those presented above, which suggests that the correlations are not an artifact of the reanalysis dataset.

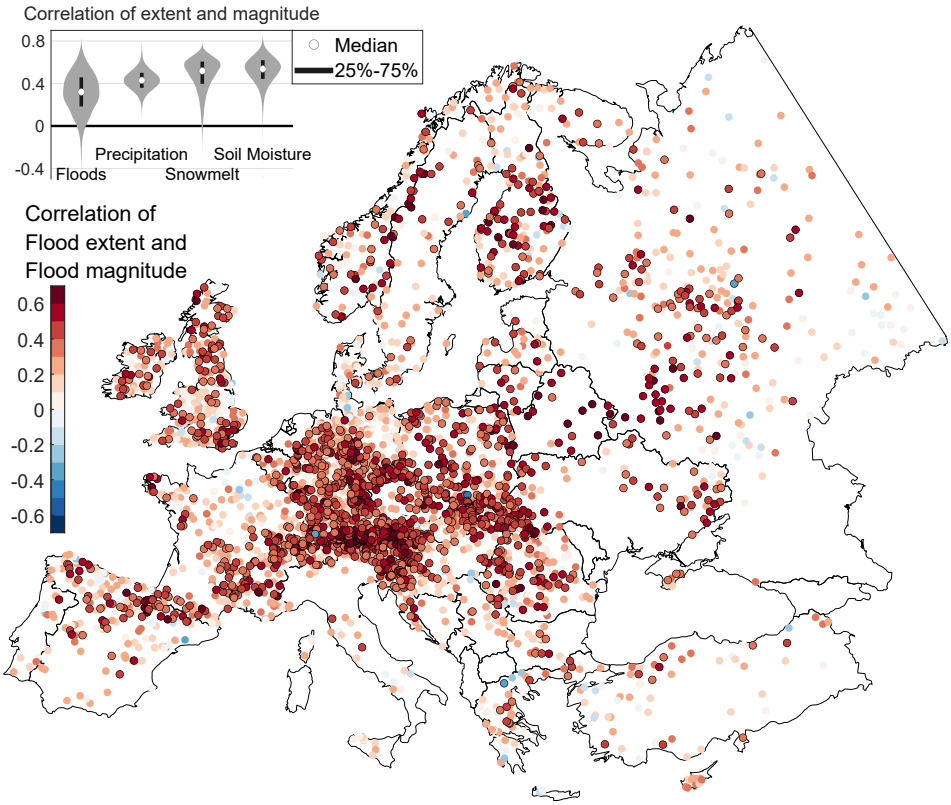

Figure 3.1: Spearman Rank Correlation of flood extent and flood magnitude. Annual series 1960-2010 (N=3872). Positive correlations are shown in red and negative in blue. Stations with significant correlations are indicated by black edges. The consistent positive correlation implies that high magnitude floods tend to be associated with large spatial extents all over Europe. The inset in the upper left corner shows the distribution of this correlation for floods as well as for three flood controls.

The flood extent, in terms of the flood synchrony scale, averages 140 km across Europe, but varies regionally (Berghuijs et al. 2019). The trends in the flood extent show clear spatial patterns (Figure 3.2). Relative to the mean flood extent over

1960–2010, regional trends range from an increase of +19% to a decrease of -20% per decade (Figure 3.2). On the British Isles, in Central Europe and at the Atlantic coast, the flood synchrony scales have increased by about 9% per decade. In Eastern Europe, the flood synchrony scales have decreased by about -11%. The spatial patterns of the trends in extent are closely aligned with those in flood magnitudes, i.e. increasing trends in North-Western and parts of Central Europe and mostly decreasing trends in the rest of the continent (Figure 3.A6). The correlation between the trends of flood magnitude and trends of flood extent of individual stations across Europe is 0.31 (N=3872) but the regional trends are more correlated (r=0.59), as some of the estimation uncertainty is removed (Blöschl et al. 2019).

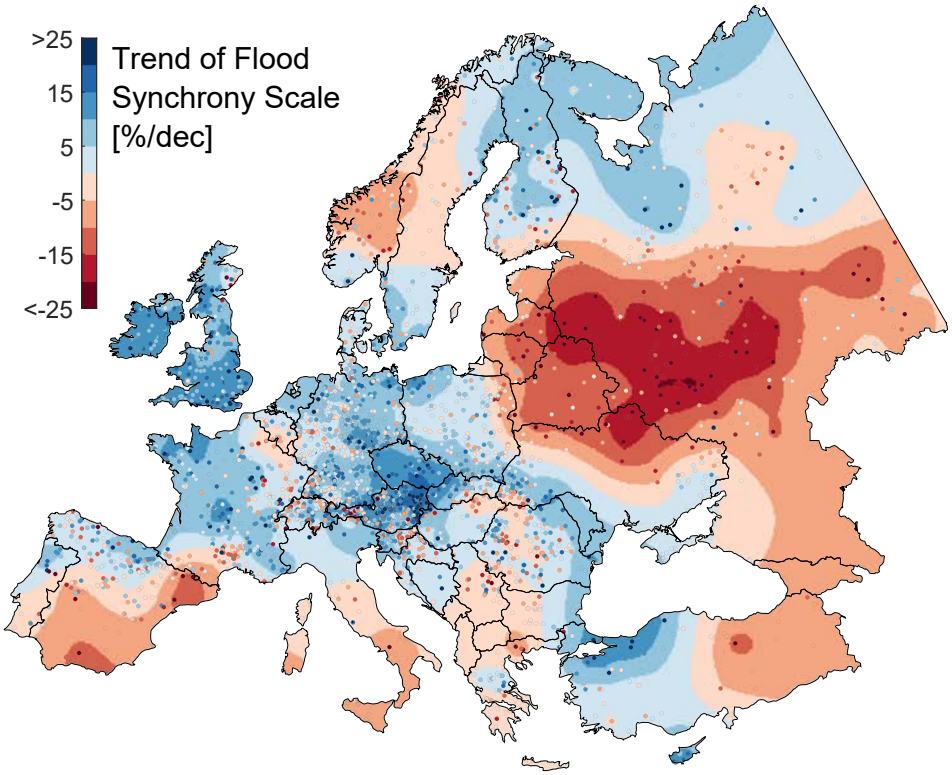

Figure 3.2: Observed trends of flood extent in Europe, 1960-2010. Blue indicates increasing flood synchrony scales and red denotes decreasing flood synchrony scales (in percent change of the mean scale per decade). The station-based trends (shown as dots) are spatially interpolated to obtain the regional trends (background color).

Our classification shows that the relevance of the flood generation processes varies across Europe (Figure 3.3a). Stratiform rainfall is particularly important in the Alps and the Carpathians, soil moisture excess in the Atlantic climate of Western Europe, and snowmelt in the North and East of Europe. Rain-on-snow has some relevance in the mid mountain ranges of Central Europe. The spatial patterns of stratiform rainfall, soil moisture excess and snowmelt controls on floods are in agreement with a previous study (Berghuijs et al. 2019). The low frequency of convective floods in this classification is because such floods usually occur in catchments of a few square kilometers (Merz and Blöschl 2003). The median catchment size of the flood data set used here is 312 km², so most convective floods are not captured in the data (Blöschl et al. 2019). As would be expected, soil moisture and rain-on-snow generated floods mainly occur in winter while convective and stratiform rainfall floods mainly occur in summer (Figure 3.A7).

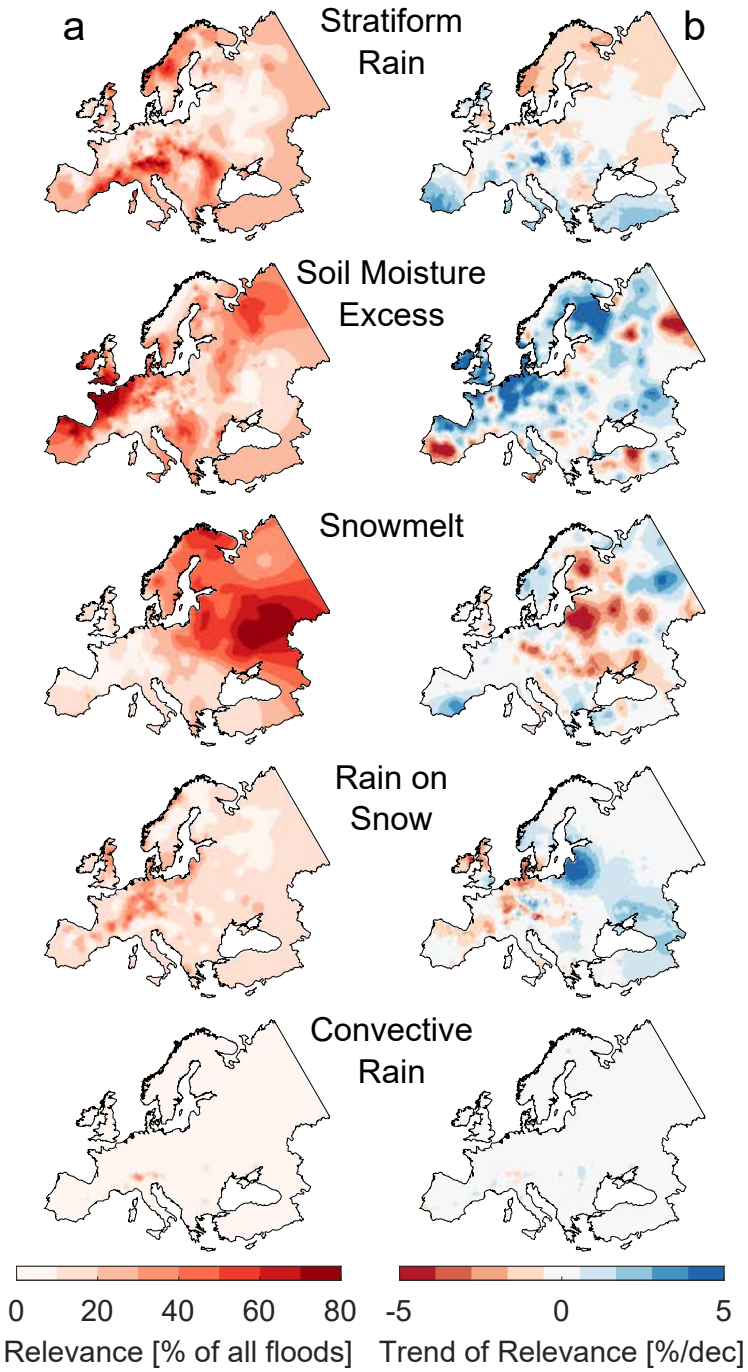

Figure 3.3: Relevance of flood generating processes and corresponding trends. (a) Relevance of each process quantified by the relative frequency of floods caused by that process in the period 1960-2010. (b) Trends of the annual relevance (in percent change of the mean relevance per decade). Only stations with at least five floods caused by the respective process are considered in the trend analysis.

In the various regions of Europe, the relevance of flood generation processes has shifted during 1960-2010 (Figure 3.3b). The relevance of stratiform rainfall has decreased in the North, but increased along the Mediterranean coast. The relevance of soil moisture excess has increased on the British Isles and in Central and Northern Europe. The relevance of snowmelt has decreased in Eastern Europe, where it is the most important process. The relevance of rain-on-snow has decreased in Western Europe but increased in parts of Eastern Europe.

The process analysis (Figure 3.3, Figures 3.A3 to 3.A5, Tables 3.A1 and 3.A2) explains why the flood extents have changed in Europe and why these extent changes are aligned with changes in the flood magnitudes. The increased flood extent in Central Europe and the British Isles is related to the increases in precipitation and soil moisture (Blöschl et al. 2019) along with the significant correlation between the magnitudes and extents of these two variables (Figures 3.A3 and 3.A4) which propagates to the floods. Additionally, there is a shift towards soil moisture excess related floods that possess larger flood synchrony scales (Table 3.1). In Eastern Europe, the opposite is the case. The decreasing flood extent is related to decreasing snowmelt along with a significant correlation between the magnitude and extent of snowmelt in this part of Europe (Figure 3.A5). Additionally, there is a shift towards a larger diversity of flood generation processes, measured by their variance, which increases by 3.2% per decade (Table 3.A1), thus reducing the flood synchrony scale. This change is aligned with an increasing variance of the timing of the annual floods by 7.3% per decade, as other processes than snowmelt increase in frequency (Table 3.A1). Finally, the shift from snowmelt towards rain-on-snow and soil moisture excess related floods increases the frequency of floods with smaller flood synchrony scales in this region (Table 3.A2).

Our findings are consistent with observed shifts towards later floods in the North Sea region and parts of the Mediterranean coast, and shifts towards earlier floods at the Atlantic coast and the continental North-East (Blöschl et al. 2017) associated with changes in the timing of snowmelt, winter storms and soil moisture excess maxima.

Table 3.1: Statistics of the flood generating processes. Floods that are generated by different processes have significantly different spatial extents (synchrony scales). The large-extent soil moisture related floods have increased in frequency. The average scaled flood magnitudes are almost independent of the flood generation process. Asterisk (*) indicates synchrony scales and flood magnitudes significantly different from the overall mean (p=0.001), dagger (†) indicates no significance.

Process	Stratiform Rainfall	Soil Moisture Excess	Snowmelt	Rain on Snow	Convective Rainfall
Overall relevance (% of all floods)	31.0	26.0	21.5	18.1	3.4
Average trend of relevance (% per decade)	0.49	1.55	−1.65	−0.41	−0.06
Average flood synchrony scale (distance as % of station average)	92.0*	113.2*	93.6*	107.2*	73.9*
Average flood magnitude (% of station average)	100.7†	103.2*	97.0*	98.0*	98.8†

3.4 Conclusions

We present a clear alignment of flood magnitudes and extents, both in terms of absolute values and trends. Additionally, we determine climatic magnitude-extent correlations and shifts in flood generation that explain these trends. The processes explaining the observed alignment of increasing trends in parts of Europe and the consistency with climate projections (Thober et al. 2018) emphasize the role of climate change in flood changes and the possibility that these changes may persist into the future. For example, in Central Europe and the British Isles, flood extents of about 43 km have increased to about 110 km during 1960-2010. If these trends continue, the alignment of magnitude and extent trends may pose more serious challenges to flood management than expected, highlighting the importance of transnational cooperation in emergency response, disaster recovery and flood risk management.

3.5 Acknowledgments, Samples, and Data

The flood data are available in the TU Vienna git-repository (https://github.com/tuwhydro/europe_floods). The reanalysis data averaged by catchment will be available at the GFZ Data Services repository (http://doi.org/10.5880/GFZ.4.4.2020.002). We furthermore used the MERIT DEM (http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM), the PERSIANN precipitation dataset (https://www.ncdc.noaa.gov/cdr/atmospheric/precipitation-persiann-cdr) and the ECA&D station-based rainfall dataset (https://www.ecad.eu/dailydata). We thank C. Primo and B. Guse for their help in acquiring the reanalysis dataset.

3.A Appendix

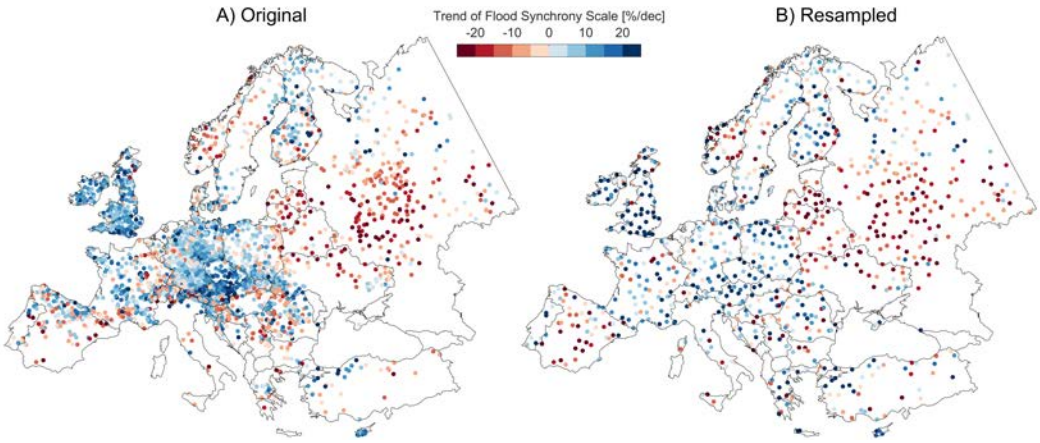

Figure 3.A1: Sensitivity of trend in flood synchrony scale to station density. Sensitivity of trend in flood synchrony scale to station density. In the complete dataset, (A) central Europe and the UK are much more densely filled with stations than the rest of Europe. By randomly sampling one station per 1x1° grid cell, we achieve a more homogeneous density of stations (B). The regional patterns in flood extent trends do not change as a result of the resampling.

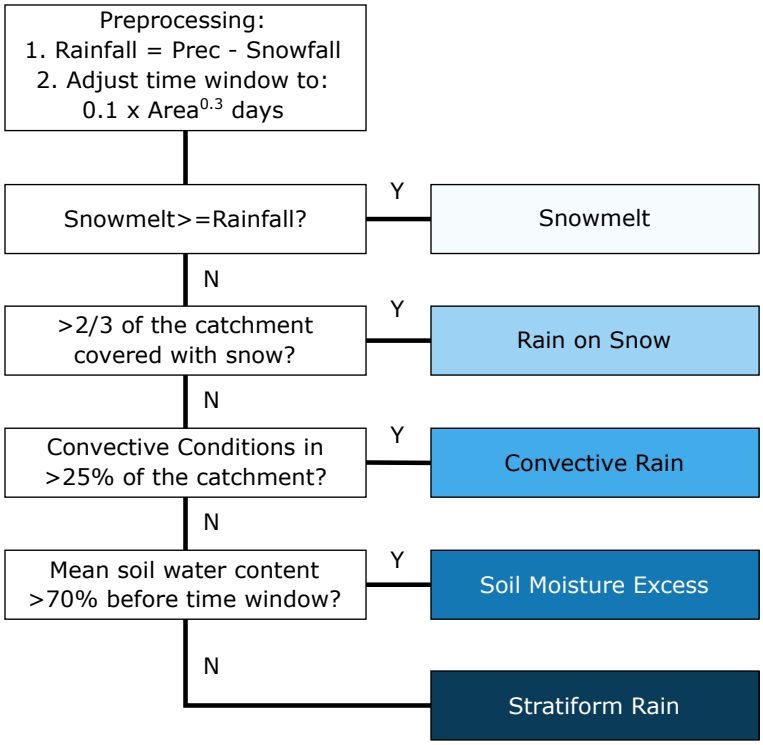

Figure 3.A2: Classification workflow of the flood generating processes. The time window is set according to the catchment area. Each flood peak is classified according to the catchment conditions by five consecutive binary decisions.

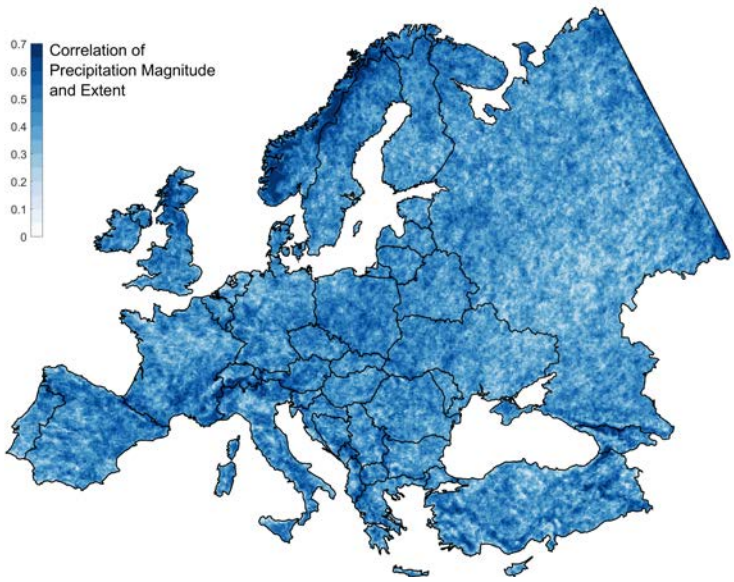

Figure 3.A3: Correlations between annual series of the magnitude of 99th percentile precipitation and associated extents. We find positive correlation in virtually all of Europe. For 99.7% of pixels, the correlation is significant at the 95% level.

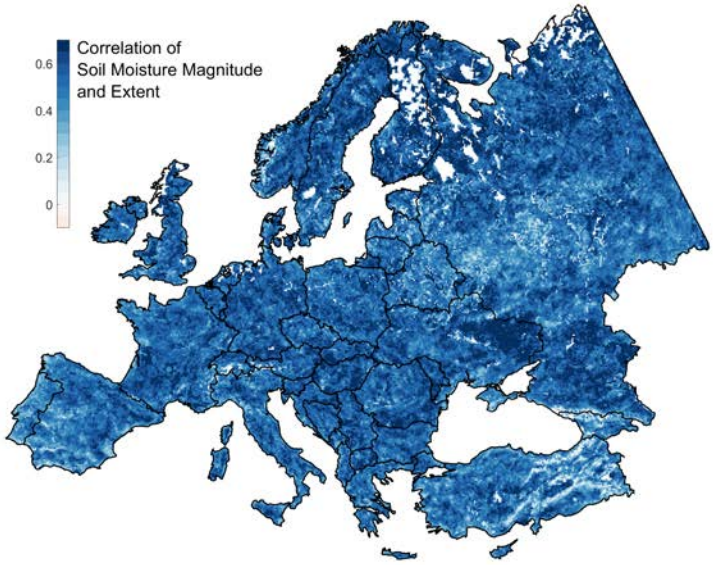

Figure 3.A4: Correlations between annual series of the magnitude of 99th percentile soil moisture and associated extents. We find positive correlation in most parts of Europe. For 99.1% of pixels, the correlation is significant at the 95% level.

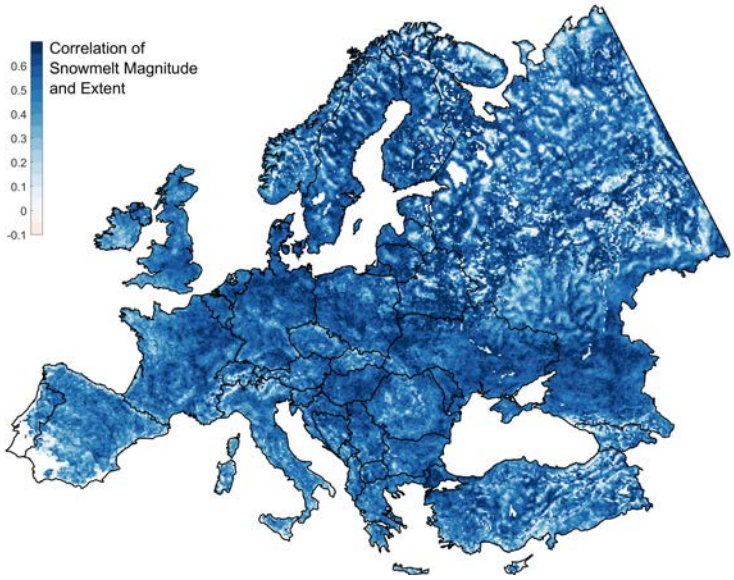

Figure 3.A5: Correlation between annual series of the magnitude of 99th percentile snowmelt and associated extents. Correlation between annual series of the magnitude of 99th percentile snowmelt and associated extent. Only pixels with at least 10 snowmelt events in the period 1960-2010 were considered. We find positive correlation in most parts of Europe. For 95.3% of pixels, the correlation is significant at the 95% level.

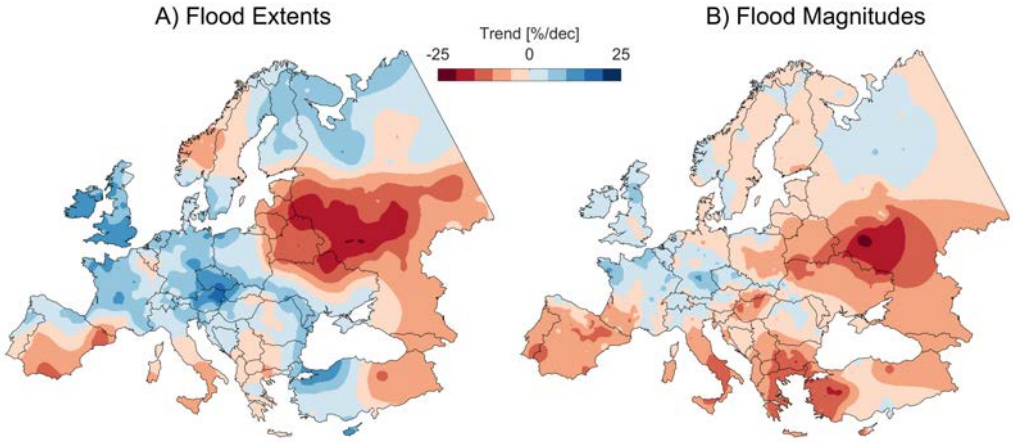

Figure 3.A6: European flood trends. (A) Regional trends in flood extent and (B) Regional trends in flood discharges in Europe (1960–2010) (Blöschl et al., 2019). Blue indicates increasing extents and discharges and red denotes decreasing extents and discharges (in percent change of the mean per decade).

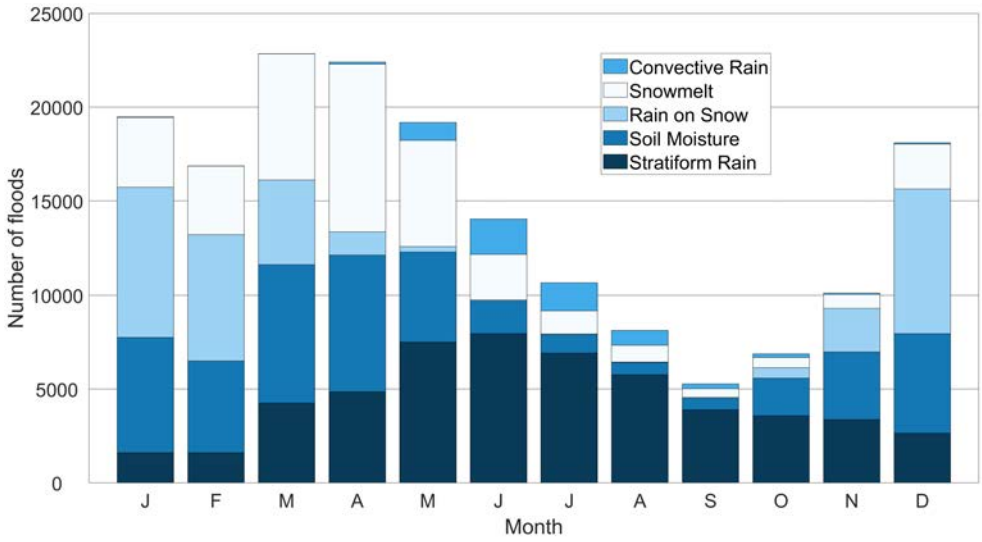

Figure 3.A7: Distribution of the flood generation processes across the year. All processes occur predominantly in the expected seasons (rain-on-snow in winter, convective rainfall in summer, etc.).

Table 3.A1: Regional flood trends. Trends of flood synchrony scales (extents) and magnitudes, as well as trends in the variance of flood dates and the diversity of flood generation processes in Europe and in two regions. Region West consists of Austria, Belgium, Czech Republic, Denmark, France, Germany, Ireland, Luxembourg, the Netherlands, Poland, Slovakia, Switzerland and the United Kingdom. Region East consists of Belarus, Estonia, Latvia, Lithuania, Russia (south of 60°N) and the Ukraine (east of 27.5°E). Significance of the trends was tested by a Mann Kendall test (p=0.05). Asterisk (*) indicates significant trends, dagger (†) indicates no significance. In the East, the timing of flood occurrence gets more variable and the floods generation processes get more diverse, while in the West these changes are very small.

Region	Europe	Region West	Region East
Number of stations	3872	2394	308
Flood synchrony scale trend (% per decade) [station based]	6.8	8.5	−10.7
Percentage of stations with significant synchrony scale trends (%, positive/negative)	37.8/15.9	49.2/6.7	4.2/62
Trend of flood synchrony scale (% per decade) [interpolated pixels]	−1.2	5.9	−8.5
Flood magnitude trend (% per decade) [station based]	0.2	1.9	−6.9
Percentage of stations with significant flood magnitude trends (%, positive/negative)	8.7/11.9	11.7/4.7	0.9/36.7
Trend of flood date variance (% of average per decade)	0.43†	0.75†	7.3*
Trend of flood type diversity (% of average per decade)	−0.35†	−0.79*	3.2*

Table 3.A2: Flood extents for different generating processes. Average synchrony scales (extents) for floods with different generation processes in Europe and in two regions (see caption Table 3.A1). Asterisk (*) indicates synchrony scales significantly different (p=0.001) from the overall mean of all generation processes in the respective region, dagger (†) indicates no significance.

Average flood synchrony scale for each generation process (% of station average)	Europe	Region West	Region East
Stratiform Rain	92.0*	90.2*	77.9*
Snowmelt	93.6*	85.3*	108.4*
Soil Moisture	113.2*	114.9*	102.3†
Rain on Snow	107.2*	112.2*	73.4*
Convective Rain	73.9*	72.3*	32.1*

4 Cascading Hazards in the Aftermath of Australia's 2019/2020 Black Summer Wildfires

Published in Earth's Future: Kemter, M., Fischer, M., Luna, L. V., Schönfeldt, E., Vogel, J., Banerjee, A., Korup, O. & Thonicke, K. (2021). Cascading hazards in the aftermath of Australia's 2019/2020 Black Summer wildfires. 9:3, e2020EF001884. https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020EF001884

Abstract

Following an unprecedented drought, Australia's 2019/2020 "Black Summer" fire season caused severe damage, gravely impacting both humans and ecosystems, and increasing susceptibility to other hazards. Heavy precipitation in early 2020 led to flooding and runoff that entrained ash and soil in burned areas, increasing sediment concentration in rivers, and reducing water quality. We exemplify this hazard cascade in a catchment in New South Wales by mapping burn severity, flood, and rainfall recurrence; estimating changes in soil erosion; and comparing them with river turbidity data. We show that following the extreme drought and wildfires, even moderate rain and floods led to undue increases in soil erosion and reductions in water quality. While natural risk analysis and planning commonly focuses on a single hazard, we emphasize the need to consider the entire hazard cascade, and highlight the impacts of ongoing climate change beyond its direct effect on wildfires.

Plain Language Summary

In 2019/20, a chain of natural hazards impacted Australia's East Coast. Following the severest drought since weather records began, record-breaking wildfires known as the "Black Summer" ravaged the region for months. In early 2020, the rainfall that extinguished the last of these fires caused further damage, as the burned soils repelled much of the rain. Water took the exposed soil and charred vegetation with it on its way to the rivers, flooding streets and polluting drinking water. We show

an example of this cascade of hazards in a single river catchment. We found that after the wildfires, even moderate rainfall caused floods, increased soil erosion, and reduced water quality drastically. Natural risk analyses mostly focus on single types of events in isolation. However, this hazard cascade shows that, especially in the face of ongoing climate change, scientists and decision makers need to consider events not just by themselves, but connected with each other.

4.1 Introduction

Australia's 2019/2020 "Black Summer" fire season was exceptional in terms of the number of fires, burned area, and fire severity (Baldwin and Ross 2020; Deb et al. 2020; Hughes et al. 2020). The fires followed an unprecedented drought; 2019 was the driest year on record (Hughes et al. 2020; Van Oldenborgh et al. 2021). Throughout the continent, the fires caused direct damages to humans and ecosystems, including at least 33 directly fire-related deaths, 3100 homes lost, an area of at least 24 million hectares burned – the size of the United Kingdom –, and never before seen air pollution levels in major cities (Davey and Sarre 2020; Hughes et al. 2020; Royal Commission into National Natural Disaster Arrangements 2020; Vardoulakis et al. 20200). The wildfires led to the formation of a record number of pyrocumulonimbus clouds that reached the lower stratosphere over southeastern Australia (Kablick III et al. 2020).

Wildfires cause hydrometeorological and geomorphic changes that can heighten the susceptibility of burned areas to other hazards; for example, raised soil water repellency after a fire can lead to increased runoff (Shakesby and Doerr 2006). This was the case with the 2019/2020 fires: following an extreme drought, the fires were the second step in an entire cascade of adverse processes (Figure 4.1). Next, rainfall in February 2020 triggered increased surface runoff and eroded ash and soil. Entrained ash, plant, and soil deposits enhanced sediment concentration in rivers, damaging infrastructure and compromising water quality (Alexandra and Finlayson 2020). In some cases, the ash-laden water contaminated water bodies such as the Lake Burragorang reservoir, Sydney's main drinking water supply (**??**).

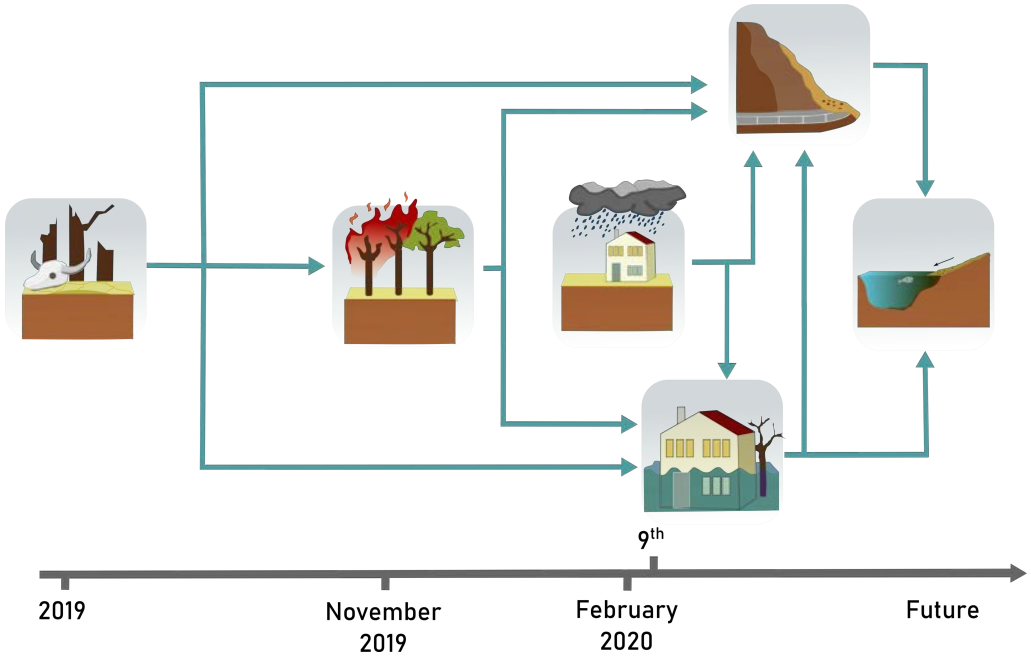

Figure 4.1: Australia's 2019/2020 hazard cascade. Drought increased the likelihood of wildfires, which burned vegetation and raised the likelihood of increased surface runoff, soil erosion and hillslope failures. When heavy rain fell in early 2020, runoff from burned areas led to flooding and entrained ash, soil, and organic matter, increasing sediment concentrations in rivers and negatively impacting water quality.

Extreme impacts, like those observed in Australia in early 2020, are often caused by a combination of several drivers (Figure 4.1). Their linkage can lead to a so-called cascading event characterized by an initial impact that triggers other, partly unexpected, effects of potentially destructive magnitudes (Pescaroli and Alexander 2015). However, the underlying drivers are mostly studied separately and without considering their potential interactions (AghaKouchak et al. 2018; Zscheischler et al. 2018). Appraisals of flood risk in Australia, for example, may underestimate the actual risk, if neglecting the impacts of an antecedent fire in the upstream catchment. When extreme impacts are combined, their effect can be greater than the sum of their parts, making a holistic approach crucial to analyzing event sequences (AghaKouchak et al. 2018; Gill and Malamud 2016; Hegerl et al. 2011; Zscheischler, Martius, et al. 2020). The analysis of cascading events remains challenging because completely documented cascades are scarce, suitable indices and methods for their quantification are limited, and bulk uncertainties are often

much higher than for single events (Kappes et al. 2012; Zscheischler, Martius, et al. 2020). Here, we illustrate the stages of a hazard cascade in a catchment in New South Wales (NSW), Australia (Figure 4.2A). We argue that considering the hazards separately may lead to serious misestimates of magnitudes, intensities, and durations of the processes involved, all of which may reverberate on hazard and risk appraisals.

During 2019/2020, the Manning River catchment was affected by drought, fires, heavy rainfall, and high sediment fluxes. Three of its tributaries experienced different degrees of burn severity (Figure 4.2B) and rainfall amounts (Figure 4.2C), allowing us to compare the post-fire impacts on streamflow and soil erosion (Figure 4.2D). By moving through the sequence of hazards, we explore how certain events triggered and influenced each other, changing their susceptibility as the event chain developed and its effects propagated throughout the catchment.

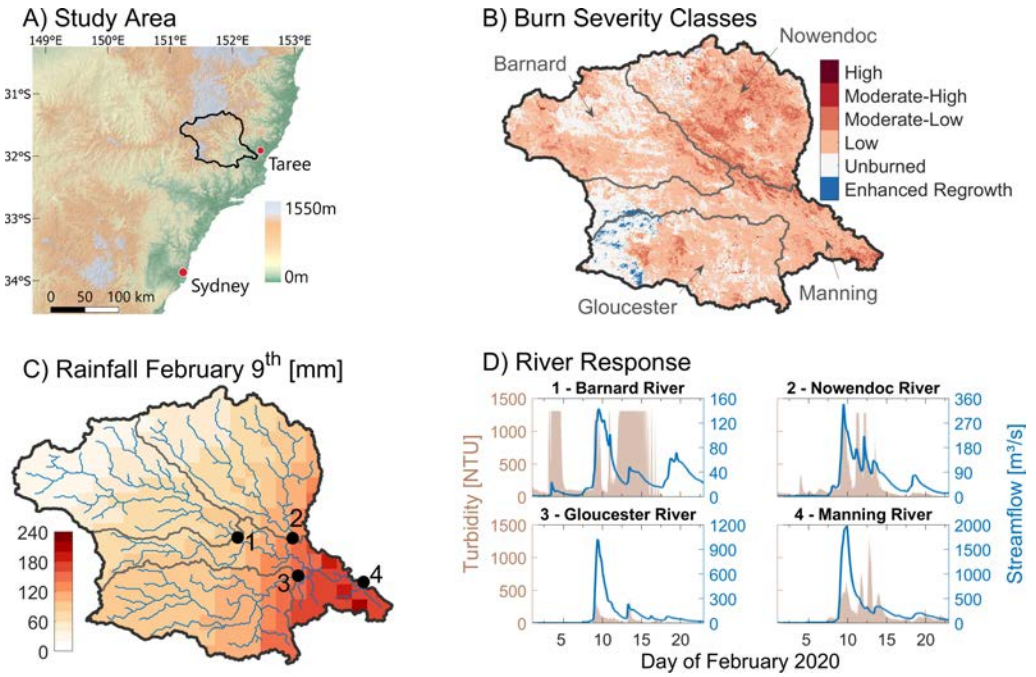

Figure 4.2: Study area. A) The Manning River catchment is located 250 km north of Sydney in one of the steepest regions of New South Wales, Australia. B) Fires affected the tributaries of the Manning River differently, with the highest burn severities occurring in the Nowendoc catchment. C) Gridded rainfall data for February 9th, 2020, show increasing rainfall totals towards the coast. 1-Barnard River (Mackay), 2-Nowendoc River (Rock's Crossing), 3-Gloucester River (Doon Ayre), 4-Manning River (Killawarra). D) Turbidity in brown and discharge in blue for Manning River and its tributaries between February 1st and 22nd.

4.2 Cascade onset: drought and heat

2019 was the driest year on record in Australia (Van Oldenborgh et al. 2021), with the lowest rainfall on record from July to December in many parts of southeastern Australia (Nolan et al. 2020; data accessible from http://www.bom.gov.au/climate/history/rainfall/). Neutral El Niño-Southern Oscillation conditions and a positive Indian Ocean dipole were the main causes for the drought (King et al. 2020; Van Oldenborgh et al. 2021). In summer 2019, this event was accompanied by the highest mean maximum temperatures since recording began in 1910, with the highest anomalies in December 2019 surpassing those of the "Angry Summer" of 2012/2013 (Van Oldenborgh et al. 2021). This extraordinary drought was a key

driver of the wildfires, whereas the role of fuel accumulation due to fire suppression is still disputed (Bradstock et al. 2020).

Based on gridded rainfall data (Jones et al. 2009), see Section 4.A) we find that 2019 was the driest year in the Manning River catchment since at least 1970 with a catchment average of only 440 mm of rainfall, or 42% of the average annual rainfall of 1040 mm from 1970 to 2018. In December 2019, the river ran completely dry at Killawarra (Figure 4.2D) for the first time on record (since 1945), where it has a daily average streamflow of 55 m^3/s.

4.3 Initial impact: extreme wildfire

Wildfires are a frequent natural hazard in Australia and have caused substantial economic and environmental impacts in the past. Yet the 2019/2020 fires were exceptional in scale, and likely linked to anomalous weather conditions driven by climate change (Bowman et al. 2020; Deb et al. 2020; Van Oldenborgh et al. 2021). They burned the largest continental fraction of any forest biome in at least two decades (Boer et al. 2020). Insurance claims from these fires totaled $2.34 billion AUD, making up 44% of all natural disaster claims for the entire fire season (Whelan 2020). In comparison, wildfires accounted for 12% of normalized insurance losses from natural hazards between 1966 and 2017 (McAneney et al. 2019). The total loss also far exceeds that incurred by the 2009 "Black Saturday" fires, when insurance claims totaled $1.2 billion AUD (Parliament of Victoria 2010). In NSW the fires caused the largest area burned and highest property loss ever recorded (Hughes et al. 2020).

The 2019/2020 fires also had detrimental health effects. Most prominently, smoke-related air pollution had an unprecedented burden on public health, with 417 total pollution-related excess deaths in eastern Australia (Queensland, NSW, Australian Capital Territory, Victoria) of which 219 were recorded in NSW (Borchers Arriagada et al. 2020). Smoke-related hospital admissions for cardiovascular and respiratory conditions totaled 3151, with 1627 cases in NSW (Borchers Arriagada et al. 2020).

To assess the overall scope of burning in the Manning River catchment, we classified burn severity by calculating the differential Normalized Burned Ratio (dNBR)

from pre- and post-fire satellite imagery from February 2019 and January 2020 respectively (Figure 4.2B) (Key and Benson 2002; Key and Benson 2006); methods are described in Section 4.A (Alleaume et al. 2005; Barrett 2006; French et al. 2008; Kinnell 2010; Lentile et al. 2006; Soverel et al. 2010; Walz et al. 2007). While dNBR-derived burn severity levels solely define burn-induced magnitude of radiometric change, Chafer (2008) conducted field studies in NSW to provide a calibration to fire effects on vegetation community strata observed on the ground. They reported that low severities signify burned grass and herbs; moderate severities imply consumed shrubs; high severities indicate scorching of the lower canopy; and very high severities denote the consumption of stems with diameters <10 mm (Chafer 2008). We found that wildfires in the Manning River catchment, which occurred from mid-November to mid-December 2019 (Data.NWS NPWS, https://data.nsw.gov.au/data/dataset/fire-history-wildfires-and-prescribed-burns-1e8b6), burned (dNBR > 0.1) a total area of 4765 km² or some 72% of the catchment (Figure 4.2B). Moderate to high burn severities (dNBR > 0.27) mostly occurred in the Nowendoc tributary, where 57% (463 km²) of the catchment area burned with this intensity at least (**??**).

4.4 Subsequent effects: floods, soil erosion, and water quality

Heavy rainfall eventually extinguished fires throughout NSW in February 2020. The rain replenished depleted water reservoirs, but also led to the next hazard in the cascade. The resulting runoff flooded parts of Sydney and other cities in NSW, caused mass movements which disrupted infrastructure, and washed soil, ash, and debris into water bodies (**??**). Insurance claims of $896 million AUD were lodged in response to the rainstorms and associated floods (Whelan 2020).

According to gridded rainfall data between 1970-2018 (see Section 4.A), the Manning river catchment averaged 78 mm of rainfall on February 9th alone (Figure 4.2C), which is about 58% of an average February rainfall total in one day. On the scale of the entire catchment, such rainfall totals occur once in $5.6^{+17.3}_{-2.4}$ years on average (**????**). Rainfall was most intense in the southern part of the catchment

(**??**), where two rain gauges measured their second highest values in records of at least 43 years (see Section 4.A).

Although parts of the Manning River catchment witnessed heavy rainfall in February 2020, the resulting floods, which we define here as the peak streamflow following the February 9th rainfall event, were only minor. The return periods of the February 9th floods range from $1.8^{+0.6}_{-0.3}$ years (Nowendoc catchment) to $4.7^{+9.8}_{-1.7}$ years (Gloucester catchment), and are thus lower than those of the preceding rainfall (**????**). We hypothesize that low soil moisture in the catchment following the drought led to decreased streamflow (Sharma et al. 2018; Wasko and Nathan 2019). The hydrographs (Figure 4.2D) show no signs of extensive surface runoff, which would form a narrow sharp spike minutes to a few hours before the main flood peak (Shakesby and Doerr 2006).

Water quality was drastically affected by this flood. In the Manning, Barnard and Nowendoc Rivers, turbidity data logged in February 2020 show sharp peaks with no precedence in the 5-7 years on record (Figure 4.2D). In some cases, the turbidity exceeded the sensor measurement scale. The uncalibrated turbidity values only allow a relative comparison of sediment loads in the tributaries. In the six years of shared record prior to the 2019 fire season, synchronous turbidity peaks for the Gloucester and Nowendoc River were of almost equal magnitude (see Section 4.A). In the more severely burned Nowendoc catchment the magnitude of the turbidity peak associated with the February 2020 flood was six times higher than in the less severely burned Gloucester catchment.

We apply the RUSLE model (Kinnell 2010; Renard et al. 1991) to estimate first order the pre- and post-fire soil erosion rates within the Manning River catchment based on rainfall erosivity, soil erodibility, steepness, land cover and management, using input parameters from pre-existing datasets (Yang 2015; Yang et al. 2017) (see Section 4.A). The dNBR burn severity is included by adjusting the post-fire land cover-factor accordingly (Blake et al. 2020; Larsen and MacDonald 2007) based on satellite data from February 2019 and 2020. The estimated post-fire soil erosion rates range from 11-27 t ha^{-1} y^{-1} (**??**), reflecting an increase of over 200%. The absolute values and relative changes are consistent with field measurements from severely burned catchments in NSW (Atkinson 2012; Blake et al. 2020; Shakesby and Doerr 2006). The increases in estimated soil erosion in the three tributaries range

from 88% in the Gloucester catchment to 358% in the Nowendoc catchment (**??** and **??**). The difference in the increase of erosion rates between these two tributaries is consistent with the respective increase in turbidity values, and likely linked to commensurate differences in burn severity.

4.5 Conclusions and outlook

The 2019/2020 hazard cascade observed in the Manning River catchment in south-east Australia highlights how the impact of ongoing climate change on wildfires affects the likelihood and magnitude of adverse consequences from other hazards that are in parts physically linked to each other. We show that following extreme drought and wildfires, moderate rainfall and flood events were sufficient to increase estimated soil erosion and reduce water quality far beyond expected levels in the absence of fires. These amplifying effects of individual impacts within hazard cascades are still insufficiently considered in risk analysis. It is crucial to fill this knowledge gap in hazard and risk appraisals, as moderate processes in hazard cascades can incur much more damage than when they occur on their own.

Climate change is projected to increase the frequency of compounding extreme warm and dry periods in Australia and beyond (Kharin and Zwiers 2005; Zscheischler and Seneviratne 2017), which could lead to further event cascades like the one in 2019/2020 (Zscheischler, Van Den Hurk, et al. 2020). Indeed, in 2020, following Australia's "Black Summer," the western United States experienced its most-extensive fire season in 70 years, while extensive fires burned across Siberia (Irannezhad et al. 2020; Pickrell and Pennisi 2020). So far, however, we can draw on only few examples of thoroughly studied hazard cascades. Mitigating the effects of climate change will require investigating these complex interactions, including these events in risk analysis and planning, establishing consistent monitoring systems to be better prepared for future hazard cascades (Bowman et al. 2020; Royal Commission into National Natural Disaster Arrangements 2020), and increasing adaptive capacity in affected regions.

4.6 Acknowledgments, Samples, and Data

This research was funded by the DFG Research Training Group "Natural Hazards and Risks in a Changing World" (NatRiskChange GRK 2043) and the DFG International Research Training Group "Surface processes, Tectonics and Georesources: The Andean foreland basin of Argentina" (StRATEGy IGK 2018) at the University of Potsdam (https://www.natriskchange.de, https://irtg-strategy.de). Gridded rainfall data was obtained from the Australian Bureau of Meteorology and is available at http://www.bom.gov.au/climate/maps/rainfall. Station rainfall data was obtained from the Australian Bureau of Meteorology and is available at http://www.bom.gov.au/climate/data. Discharge and turbidity data was obtained from the State of NSW (Lands and Water) and is available at http://www.bom.gov.au/waterdata. The soil erosion data from the State Government of NSW and Department of Planning, Industry and Environment is available at https://datasets.seed.nsw.gov.au/dataset/modelled-hillslope-erosion-over-new-south-wales. Landsat OLI imagery is available from the US Geological Survey (https://www.earthexplorer.usgs.gov). Images from Sentinel are from Sentinel Playground, https://apps.sentinel-hub.com/sentinel-playground, Sinergise Ltd.

4.A Appendix

4.A.1 Introduction

This supplementary information provides additional detail on the methods used to estimate and the results of 1. Burn Severity, 2. Rainfall return periods, 3. Flood return periods, and 4. Erosion and water quality.

4.A.2 Burn severity

We quantified the extent of burned area in the Manning River catchment with remote sensing techniques described by Key and Benson (2002; 2006) in the context of the United States Forest Service's FIREMON Landscape assessment methods. Although initially described for the North American boreal region, these methods have been applied to forests of different vegetation zones worldwide in a number of

studies (French et al. 2008). In order to quantify burn severity, which is defined as a scaled index of the magnitude of fire-induced change in an ecosystem, we calculated the Normalized Burned Ratio (NBR) for pre- and post-fire satellite imagery. We derived NBR-values from Landsat OLI scenes acquired in February 2019 (2019/02/07 and 2019/02/17), before major fires took place in the study area, and in December 2019 and January 2020 (2019/12/31 and 2020/01/09) after the peak of the 2019/20 fire season. Changes in vegetation growth are especially evident at near-infrared (NIR) wavelengths; burned wood as well as bare soil are detectable in the short-wave infrared (SWIR) spectrum. Thus the NBR is calculated using the NIR and SWIR bandwidths as:

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} \tag{4.1}$$

Translated to the multispectral Landsat OLI band's wavelength ranges this is:

$$NBR = \frac{Band5 - Band7}{Band5 + Band7} \tag{4.2}$$

To assess the changes in vegetation caused by a fire, we calculated the differential Normalized Burned Ratio (dNBR) (Key and Benson 2002; Key and Benson 2006; Lentile et al. 2006; Soverel et al. 2010):

$$dNBR = preNBR - postNBR \tag{4.3}$$

The possible range of dNBR values is from -−2.0 to + 2.0, with dNBR values close to 0 indicating insignificant vegetative change (unburned background); strongly positive values mark intensively burned biomass in a given Landsat OLI pixel (Key and Benson 2006). Pixels with strongly negative values, however, can be interpreted as having enhanced regrowth of vegetation during post-fire recovery of some herbaceous communities (Key and Benson 2006). Based on this metric, Key and Benson (Key and Benson 2002; Key and Benson 2006) originally classified seven burn severity levels (**??**). These dNBR thresholds and burn severity classes have been used in numerous studies globally (e.g. for Namibia see Alleaume et al. 2005; for Australia see Barrett 2006; Walz et al. 2007) and adopted here to extract areas of moderate-low, moderate-high, and high burn severity. Lentile et al. (2006)

pointed out that the dNBR thresholds were defined based on a local study and were, hence, not intended to constitute a common standard.

4.A.3 Rainfall return periods

We use a spatially interpolated 5x5 km grid of daily rainfall values from the Australian Bureau of Meteorology (Jones et al. 2009, http://www.bom.gov.au/climate/maps/rainfall). We obtained data from 1970-2020. To obtain catchment averages, daily means of all pixels in the catchment were calculated and weighted by the fraction of pixel area within the catchment. We chose those eight rainfall gauges within or close to the boundaries of the Manning River catchment, which have at least 40 years of daily rainfall sums on record and data at least as recent as March 2020. Return periods were calculated by fitting a Generalized Extreme Value (GEV) distribution to the annual maxima of the rainfall time series. We chose this type of distribution to minimize model selection bias. Uncertainty intervals were estimated at the 95% confidence level.

4.A.4 Flood return periods

We use discharge data with a resolution of up to 15 minutes from the NSW Department of Industry – Lands and Water. For flood frequency analysis the discharge values were averaged to daily values to allow the comparison with earlier values for which only daily resolution is available. Station locations are shown in Figure 4.2C and station details are listed in **??**. Return periods were calculated by fitting a GEV distribution to the annual maxima of the discharge time series. Uncertainty intervals were estimated at the 95% confidence level.

4.A.5 Erosion and water quality

We use turbidity data with a resolution of 15 minutes from the NSW Department of Industry – Lands and Water. Turbidity data for the Gloucester River is unreliable from February 23rd and is flagged as low-quality data after this date by the provider NSW Department of Industry - Lands and Water. Hence, we did not use any turbidity data beyond this date for the analysis. For the comparison of turbidity

peak values between the Nowendoc and Gloucester rivers, we first downsampled the time series to daily resolution. We then defined a set of peaks for each river so that peaks are separated by at least 7 days and that they exceed a turbidity value of 25 NTU. We defined the common set of peaks as those that occur at both rivers within one day. We define the turbidity peak magnitude as the area under the curve of the 7-day window starting one day prior to the peak. The magnitudes averaged 256 and 329 NTU before the fires and 540 and 3399 NTU afterwards, for the Gloucester and Nowendoc River, respectively. The difference in post-fire turbidity values is shown in Figure 4.2D. The Revised Universal Soil Loss Equation (RUSLE) is a widely used empirical model for predicting the average annual soil loss from rain splash, sheetwash and rill erosion at a hillslope (Kinnell 2010). Soil loss (A) [t ha^{-1} yr^{-1}], is estimated as:

$$A = R * K * LS * C, \tag{4.4}$$

where R is rainfall-runoff erosivity (MJ mm ha^{-1} hr^{-1} yr^{-1}), K is soil erodibility (t ha h ha^{-1} MJ^{-1} mm^{-1}), LS is slope length and steepness factor (unitless) and C is a land-cover and management factor (unitless). For the R-, K-, LS- and pre-fire C-factors existing datasets were used (Yang 2015; Yang et al. 2017). The RUSLE equation can be modified by substituting the C-factor according to the burn severity classified from dNBR values (Larsen and MacDonald 2007). Post-fire C-factor were set to 0.01 for low severity, 0.05 for moderate low severity, 0.1 for moderate high severity and 0.2 for high severity (Blake et al. 2020).

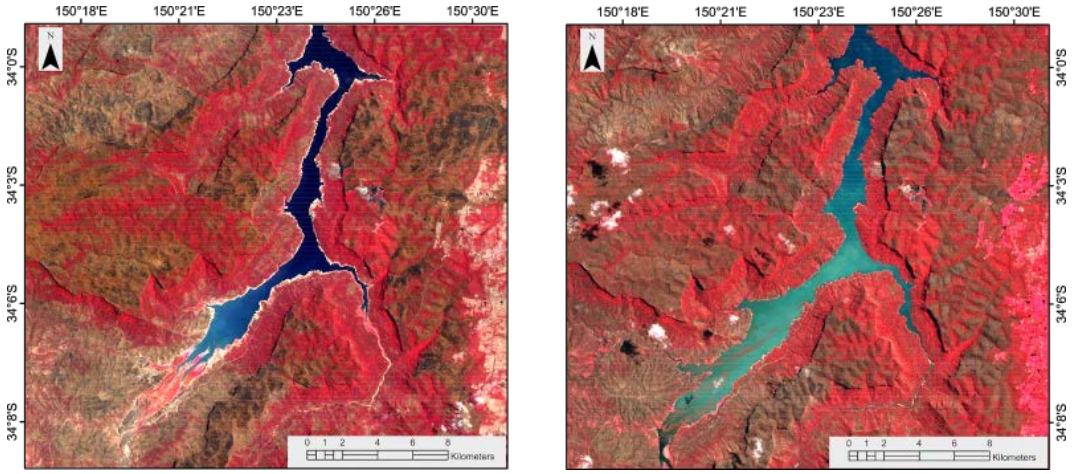

Figure 4.A1: Fire-related suspended sediment loads in the Burragorang reservoir. The lake-level drop during the last years is well visible at the left picture due to the white fringe around the lake. Sediment input from soils and sediments eroded by post-fire erosion in the right picture is mainly from the southern inlet and marked by a distinct change in water color. Left picture 2020/01/10; right picture 2020/02/19 (infrared color, Sentinel, DLR).

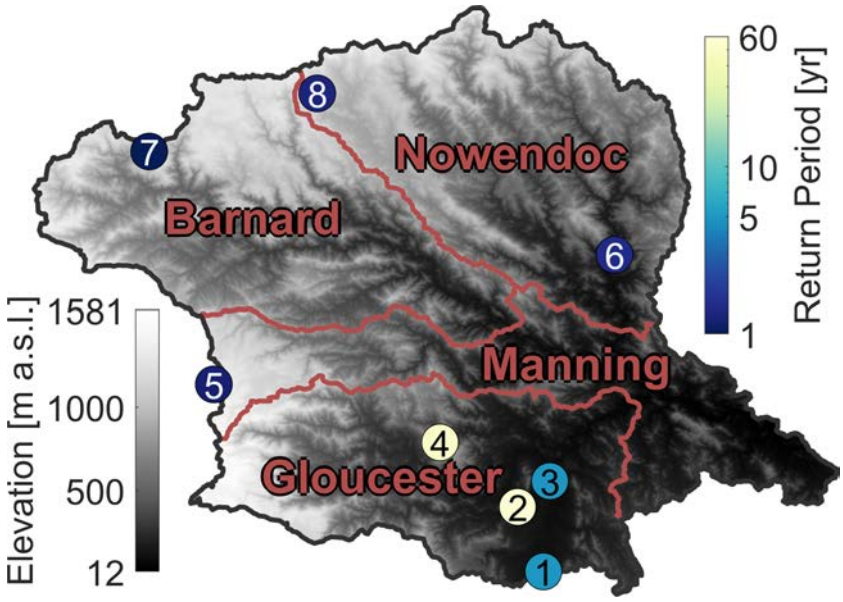

Figure 4.A2: February 9th rainfall return periods, Manning River catchment, NSW. The map shows the location of the eight rainfall gauges we considered. Gauges are colored on a logarithmic scale expressing the return period of rainfall total on February 9th 2020 or one day before or after if rainfall total was higher on that day (**??**). The large contrast in return periods for the neighboring stations 2 and 3 is due to different record length and station elevation.

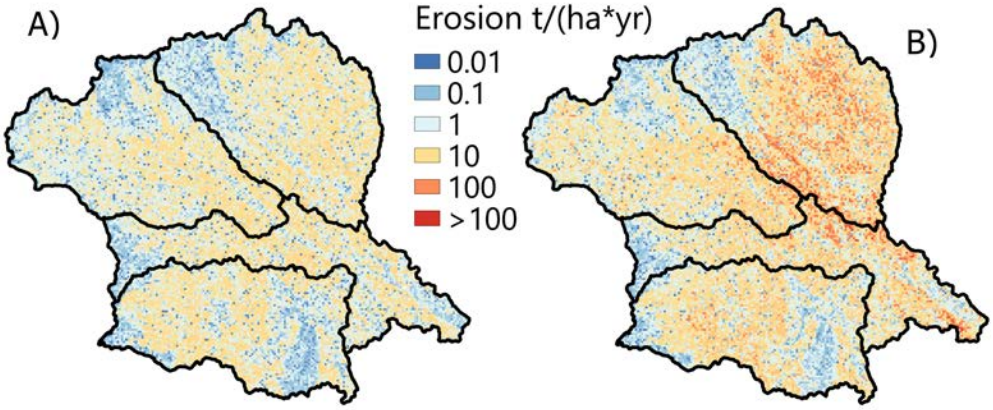

Figure 4.A3: Fire-induced changes in soil erosion rate. Estimated pre-fire (A) and post-fire (B) soil erosion rates, Manning River catchment, NSW. The output of the RUSLE model shows a severe impact of wildfires on the soil erosion rates. The most severely burned areas in the Nowendoc tributary show the most drastic increase in soil erosion. Estimated post-fire soil erosion rates are in agreement with field measurements in severely burned catchments in NSW (Atkinson 2012; Blake et al. 2020; Shakesby and Doerr 2006). Note the logarithmic color scale.

Table 4.A1: Catchment Comparison. The three tributaries of the Manning River we study here are very similar in size and pre-fire erosion rate but have been affected very differently by the wildfires of 2019/20 and the extreme rainfall of February 2020. Post-fire erosion rates are estimates and depend strongly on burn severity classes and the chosen severity specific *C*-factors. [1] – Area with a burn severity of moderate low or higher is considered as burned. [2] – Catchment averages.

River	Area [km^2]	Burned Area [%][1]	Burned Forest Area [%][1]	Pre-Fire Erosion [t/ha*y][2]	Post-Fire Erosion [t/ha*y][2]	Erosion Increase [%][2]
Barnard	1815	5.9	3.9	5.1	11.2	121
Nowendoc	1892	24.5	22.7	6.0	27.3	358
Gloucester	1630	4.4	1.0	6.0	11.3	88
Manning	6633	12.1	9.9	5.8	17.7	202

Table 4.A2: Estimated return periods of recorded post-fire rainfall, Manning River catchment, NSW. The rainstorms of early February 2020 produced rare rainfall totals in parts of the Gloucester tributary. Rainfall decreased towards the north of the Manning river catchment.

#	Station Name	Event Rainfall [mm]	Day of February 2020	Return Period [yr] (95% Conf. Interval)	Record Length [yr]	Elevation [m a.s.l.]
1	**Craven (Longview)**	113.2	9th	5.1 (3.1-11.0)	46	130
2	**Gloucester (Hiawatha)**	146.0	10th	59.9 (17.7-336.4)	43	125
3	**Gloucester Post Office**	116.2	9th	5.1 (3.6-8.3)	94	105
4	**Gloucester (Upper Bowman)**	169.0	9th	54.8 (19.0-231.6)	54	280
5	**Hunter Springs (Wondecla)**	48.0	9th	1.2 (1.1-1.3)	48	1235
6	**Number One (Murrays Creek)**	69.4	9th	1.5 (1.3-1.8)	51	120
7	**Nundle (Olsland)**	29.6	9th	1.0 (1.0-1.1)	40	1240
8	**Nowendoc (Green Hills)**	52.0	8th	1.3 (1.2-1.5)	49	1130

Table 4.A3: Estimated rainfall return periods extrapolated for the Manning River and its tributary catchments. The 5km gridded daily rainfall values were averaged for each catchment. The record length is 50 years in every case.

#	Station Name	Event Rainfall [mm]	Return Period [yr] (95% Conf. Interval)
1	**Barnard (Mackay)**	51	2.15 (1.71-3.52)
2	**Nowendoc (Rocks Crossing)**	76	2.82 (2.03-6.03)
3	**Gloucester (Doon Ayre)**	102	8.21 (4.03-59.49)
4	**Manning (Killawarra)**	85	5.63 (3.19-22.98)

Table 4.A4: Frequency analysis of post-fire floods, Manning River, NSW. Even the heavy rainfall in the Gloucester tributary only led to moderate levels of discharge in the river.

#	Station Name	Peak Discharge [m^3/sec]	Flood Peak Time on February 9th	Return Period [yr] (95% Conf. Interval)	Record Length [yr]
1	**Barnard (Mackay)**	143	03:30	2.09 (1.71-3.16)	58
2	**Nowendoc (Rocks Crossing)**	339	03:00	1.75 (1.49-2.41)	75
3	**Gloucester (Doon Ayre)**	1024	00:00	4.69 (2.97-14.50)	75
4	**Manning (Killawarra)**	1976	11:30	2.86 (2.17-5.03)	75

Table 4.A5: Estimated burn severity levels based on dNBR values. As defined by Key & Benson (Key and Benson 2002; Key and Benson 2006).

Burn Severity Levels	dNBR Range
Enhanced Regrowth, High	−0.5 to − 0.251
Enhanced Regrowth, Low	−0.25 to − 0.101
Unburned	−0.10 to +0.099
Low Severity	+0.1 to +0.27
Moderate – Low Severity	+0.27 to + 0.439
Moderate – High Severity	+0.44 to + 0.659
High Severity	+0.66 to +1.300

5.1 Discussion

In the introduction of this thesis, three questions were raised regarding the effects of human activity on river floods. The three respective studies present approaches to study these questions and provide several answers to them. As global data availability is highly heterogeneous and biased towards Europe, North America, and Australia, the results presented here are limited to these regions. Future studies – based on global hydrological models or newly measured datasets with greater coverage – will need to show, how the presented results translate to other continents. The methods presented here can be easily applied to different regions once the necessary data becomes available.

5.1.1 Controls on Flood Magnitude Trends

Rainfall is the most important factor influencing flood magnitude trends in the CONUS. Changes in rainfall magnitude and extent are expected given the intensification of the water cycle due to climate change. However, regionally these trends do not always lead to increasing flood magnitudes (Sharma et al. 2018). In many parts of the CONUS, flood magnitude trends are negative despite the increasing rainfall amounts. The reasons for this differ regionally. In catchments, in which snowmelt is the main cause of floods, a change in rainfall amounts will have a very limited effect. Instead, decreasing snowpacks due to warmer temperatures contribute less meltwater to the peak flows (Rood et al. 2016). For now, this effect is most common in mid altitudes (Rood et al. 2016; Rottler et al. 2019) but is likely to propagate to higher parts of mountainous catchments with increasing global temperatures, as research in Europe indicate (Rottler et al. 2021). Another aspect mitigating the increase of flood magnitudes regionally is land use. This can be due to aimed regulation or diversion of flood water using reservoirs. It can also be

caused by the presence or absence of certain land cover types. Forests can mitigate the effects of increasing rainfall volume by storing excess water (Blöschl et al. 2007; Marks et al. 1998). On the other hand, agricultural fields (Schilling et al. 2014) and impervious surfaces increase surface runoff, reducing flood concentration time and increasing flood magnitudes. The results of Chapter 2 indicate that the relevance of these land cover types follows a non-linear process, only affecting flood magnitude trends after exceeding a certain threshold. It will be crucial to detect such threshold processes as they could have important implications for flood-risk and land-use management. This will require the use of suitable methods, and the combination of Random Forests and ALE plots presented here are a promising option.

Additionally, in many regions of the CONUS static variables were found to be the most important controls on flood magnitude trends. As trends cannot be caused by stationarity, this indicates that the selection of these variables is due to a lack of further information. In all likelihood, the static variables serve as proxies for dynamic processes, which are not reflected in any of the trend variables, but which correlate with these static values to some extent. Furthermore, there is substantial uncertainty attached to the trend variables, especially those of the land cover types (see Section 2.5.4). This issue is due to data scarcity and can be resolved in future studies using decades of remote sensing data. The presented methodology can be easily expanded with additional or improved variables.

5.1.2 Flood Extents

In Europe, a strong positive correlation between flood magnitudes and flood extents exists. This can be explained by an even stronger correlation of the same variables for the flood generating processes (rainfall, soil moisture, snowmelt). This relationship poses a severe challenge for flood risk management, as it indicates that those floods that have the highest peaks, are also those that are most likely to be accompanied by simultaneous flooding in surrounding catchments. At the same time, the trends of flood magnitudes and flood extents are aligned as well. To study whether these trends will persist in the future, the physical processes driving them need to be investigated in more detail and future trends need to be projected for relevant climate change scenarios and for different parts of the world. Additionally,

extreme scenarios – comparable to the ARkStorm in California (Porter et al. 2011) – with severe simultaneous flooding along multiple major rivers need to be developed for different regions to estimate whether or how they would overtax infrastructure, disaster relief, and insurers.

The same relationship between flood extents (**??**a) and flood magnitudes exists in the CONUS (**??**b). Similar to the European dataset the average Spearman correlation lies at 0.32, with 47.5% of catchments showing significant positive correlation (46.5% in Europe). However, the trends of flood extents (**??**c) and flood magnitudes (Figure 2.2b) are less correlated in the CONUS with a Spearman correlation of 0.19, compared to 0.31 in Europe. This suggests, that in the CONUS the alignment of increasing flood magnitudes and extents is of lesser concern than in Europe so far, but the drivers of this discrepancy will need to be studied further. A recent study about flood extents using a different measure of spatial dependence indicates that, like in Europe, flood extents are seasonally variable and process dependent in the CONUS (Brunner et al. 2020). Therefore, it is likely that, similar to flood magnitude trends, flood extent trends in the CONUS are controlled by spatially varying hydro-meteorological variables.

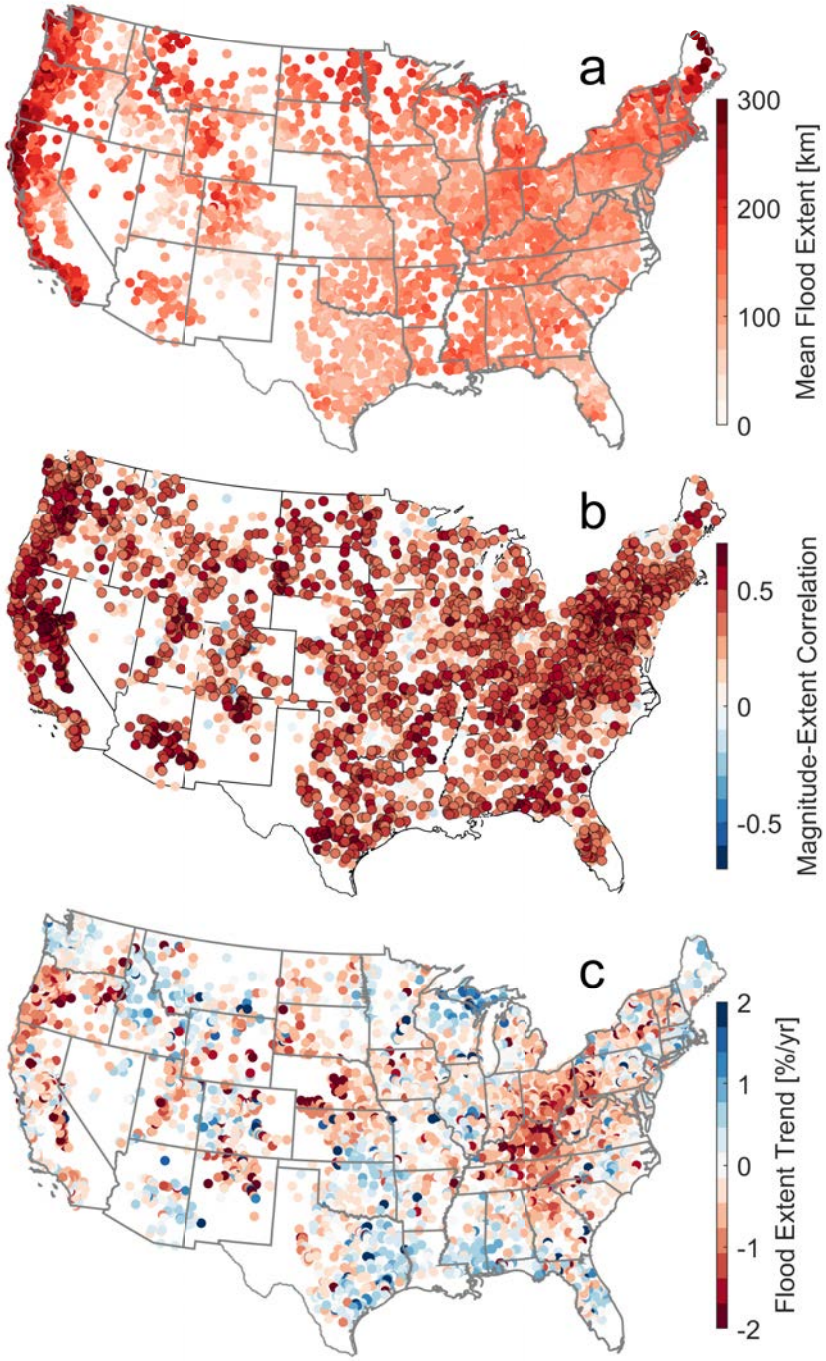

Figure 5.1: Spatial Flood Extents in the CONUS. Based on the same 4390 catchments as in Chapter 2, using the methods from Chapter 3. a) Average flood synchrony scale of AMS from 1960-2010. b) Spearman Correlation between annual values of AMS magnitude and flood synchrony scale. Dark-bordered dots show significant values ($p<0.05$, multiple testing correction). c) Trends in flood synchrony scale estimated using Theil-Sen.

5.1.3 Floods in Hazard Cascades

The analysis of the 2019/2020 hazard cascade in the Manning River highlights several previously neglected aspects of flood generation within hazard cascades. Firstly, it shows that flood return levels can be both increased and decreased by preceding hazards in the catchment. In the case of the Manning River, the persisting dry conditions from the 2019 drought led to a higher available water storage capacity than usual, halving the return period of the 2020 flood compared to the return period of the flood generating rainfall.

Secondly, cascades can alter the properties of floods outside of their magnitudes, affecting their potential damage. The measured turbidity values in the tributaries of the Manning River correspond directly to the local burn severity of the preceding wildfires. This high turbidity is a sign of flood water contamination, which is an underestimated aspect of flood hazard and can exacerbate flood damages, particularly to river ecosystems (Robinne et al. 2020). Through such flood property changes, even medium magnitude floods can cause substantial damages in unexpected ways.

The locally large differences in the effects of hazard cascades on floods that we see in the Manning River tributaries, show that there are many challenges for flood modelling in the future. Large-scale or even global models are far away from having the spatial resolution to represent this local variability. Even small-scale models are generally focused on single hazards only (UNDRR 2019). Therefore, very little is known about the interactions and likelihood of hazard cascades or how they might change in the future. The example of the Manning River shows that these open questions need to be investigated.

5.2 Outlook

This thesis is limited to analysis of the past, aiming to give a guideline of the aspects that coming studies into the future of flood hazard will need to consider in order to capture all possible changes. It shows that flood generation has already changed in many parts of the Earth and that the factors controlling these changes can be of different importance or even opposite effect regionally. Given this spatial

heterogeneity, future flood research and risk management will need to adapt its methods and assumptions to the previously underrepresented regions of the world (Africa, South America, Asia). It is likely that different variables control flood trends in these areas, than in the well-studied parts of Europe and North America.

Controls on changing flood generation are not restricted to the direct impacts of climate change on the water cycle (precipitation, evapotranspiration). Studies into future river flood hazard should not only consider these factors, but also possible interactions with changes in land cover. Investigating the role of land cover in flood generation could reveal pathways for the mitigation of increasing flood hazard. Barren or agricultural land can be reforested. Even impervious urban areas can become "sponge cities", increasing their water storage capacity and thereby reducing flood magnitudes, a practice that the Chinese government considers as a key solution to pluvial flooding (Li et al. 2017), but which could also attenuate downstream river floods.

Furthermore, these changes do not solely lead to the trends in flood magnitudes that are most commonly associated with a changing climate. The changing spatial extents of river floods (and of other natural hazards) should be studied globally. Their future trends need to be projected, to find regions of growing flood extents. In these regions, flood risk assessment will need to become transnational, to correctly forecast large-scale flood events with impacts on multiple national economies and international supply chains and to coordinate disaster relief. The Global Flood Partnership is an example for a recently established international cooperation to improve flood forecasting, monitoring and damage assessment through shared expertise (UNDRR 2019).

The management of hazard cascades needs to be integrated and improved, to allow decision makers to minimize the combined risk of the entire cascade. In many countries different natural hazards are still viewed as distinct events and handled by entirely separate government entities, with little to no exchange between them. Instead, future risk management will require models and frameworks that are capable of handling complete hazard cascades and of projecting trends in their likelihood. In the recent decade, many methods have been established and tested – from copulas to multi-layer complex networks – allowing for the analysis of physically or socially connected extreme events (Raymond et al. 2020). However,

so far most of them do not account for temporal or spatial non-stationarity of the different hazards (Gallina et al. 2016), which will be indispensable in the face of climate change.

The effects of human activity on river flood hazard are not limited to those studied in this thesis. River course alterations (Munoz et al. 2018) and groundwater withdrawal (Rasmussen and Perry 2001) are just two more examples of direct and indirect human influences on flood generation. Additionally, the hazard is only one component of flood risk, which is defined as the expected loss of material and intangible assets due to floods in a certain area. This factors in two more risk-components. Firstly, exposure to the hazard, meaning the assets and humans at risk from flooding. Secondly, vulnerability of these elements, i.e. the expected damage to them in case of exposure to flood. Just as flood hazard can no longer be seen as stationary, all three interdependent components of flood risk change with time and their projection and risk assessment are becoming more complex. These changes and interdependencies lead to an increasing uncertainty in flood loss predictions (Merz et al. 2010). It will be crucial, to find ways to estimate and communicate this uncertainty properly – for example through Bayesian statistics (Sieg et al. 2019) – to allow stakeholders to make informed decisions. Possible approaches to handle these uncertainties from a decision-making perspective could be adaptive policy pathways (Haasnoot et al. 2013), which allow for dynamic reaction to emerging changes, or storylines (Shepherd et al. 2018) exploring high-impact scenarios to prepare for the worst case. Such non-deterministic approaches can complement the rapidly improving global hydrological models to successfully assess future flood risk. All of them will need to consider the regional variability, temporal component, and interconnectivity of flood hazard change highlighted in this thesis.

5.3 Synthesis

This thesis shows how river flood generation is affected by climate change and human activity and what the resulting changes mean for different aspects of flood hazard across the world. Chapter 2 and 4 show that weather pattern changes, land cover conditions, and hazard cascades can influence flood magnitudes substantially, both positively and negatively. All of these factors are influenced by humanity

directly or indirectly through anthropogenic climate change. Furthermore, all three studies show that human impacts on flood hazard can be regionally different and even lead to opposing trends depending on the initial conditions. This spatial variability can occur on the continental or even the sub-catchment scale, as the example of the hazard cascade in the Manning River basin shows.

Chapter 3 highlights the fact that different aspects of flood hazard can be correlated and even show aligned trends. Such an alignment can lead to consequences that are greater than just the sum of its parts, similar to the amplifying effects within a hazard cascade. However, **??** shows that these alignments are not necessarily globally universal. They might even be non-stationary themselves and suspect to future changes. The three studies show that a vast variety of methods already exists, that allows for the detailed investigation of human influences on flood hazard, even in the presence of collinearities, non-linear relationships, or spatial heterogeneity in flood generation. These methods can be applied to different regions and newly available data.

All the discussed aspects need to be considered to reliably predict future changes in flood generation, flood hazard, and ultimately flood risk. Transnational cooperation will be necessary to successfully mitigate and adapt to these changes.

Agarwal, A., N. Marwan, M. Rathinasamy, B. Merz, and J. Kurths (2017). **Multi-scale event synchronization analysis for unravelling climate processes: A wavelet-based approach**. *Nonlinear Processes in Geophysics* 24:4, 1607–7946. (see page 18).

AghaKouchak, A., L. S. Huning, F. Chiang, M. Sadegh, F. Vahedifard, O. Mazdiyasni, H. Moftakhari, and I. Mallakpour (2018). **How do natural hazards cascade to cause disasters?** *Nature* 561, 458–460. (see page 79).

Alexandra, J. and C. M. Finlayson (2020). **Floods after bushfires: rapid responses for reducing impacts of sediment, ash, and nutrient slugs**. *Australasian Journal of Water Resources* 24:1, 9–11. (see page 78).

Allan, R. P. and B. J. Soden (2007). **Large discrepancy between observed and simulated precipitation trends in the ascending and descending branches of the tropical circulation**. *Geophysical Research Letters* 34:18, 1–6. (see page 4).

Alleaume, S., C. Hely, J. Le Roux, S. Korontzi, R. Swap, H. Shugart, and C. Justice (2005). **Using MODIS to evaluate heterogeneity of biomass burning in southern African savannahs: a case study in Etosha**. *International Journal of Remote Sensing* 26:19, 4219–4237. (see pages 83, 87).

Apley, D. W. and J. Zhu (2020). **Visualizing the effects of predictor variables in black box supervised learning models**. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 82:4, 1059–1086. (see pages 13, 22).

Archfield, S. A., R. M. Hirsch, A. Viglione, and G. Blöschl (2016). **Fragmented patterns of flood change across the United States**. *Geophysical Research Letters* 43:19, 232–239. (see page 12).

Armstrong, W. H., M. J. Collins, and N. P. Snyder (2014). **Hydroclimatic flood trends in the Northeastern United States and linkages with large-scale atmospheric circulation patterns**. *Hydrological Sciences Journal* 59:9, 1636–1655. (see pages 4, 12).

Arnell, N. W. and S. N. Gosling (2016). **The impacts of climate change on river flood risk at the global scale**. *Climatic Change* 134:3, 387–401. (see page 1).

Aryal, Y. N., G. Villarini, W. Zhang, and G. A. Vecchi (2018). **Long term changes in flooding and heavy rainfall associated with North Atlantic tropical cyclones:**

Roles of the North Atlantic Oscillation and El Niño-Southern Oscillation. *Journal of Hydrology* 559, 698–710. (see page 32).

Ashouri, H., K. L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat (2015). **PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies**. *Bulletin of the American Meteorological Society* 96:1, 69–83. (see page 60).

Atkinson, G. (2012). **Soil erosion following wildfire in Royal National Park, NSW**. In: *Proceedings of the Linnean Society of New South Wales*. Vol. 134 (see page 84).

Baldwin, C. and H. Ross (2020). **Beyond a tragic fire season: a window of opportunity to address climate change?** *Australasian journal of environmental management* 27:1, 1–5. (see page 78).

Barrett, T. (2006). **Modelling burn severity for the 2003 NSW/ACT unplanned fires using Landsat imagery**. In: *Bushfire conference* (see pages 83, 87).

Barth, N. A., G. Villarini, M. A. Nayak, and K. White (2017). **Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers**. *Water Resources Research* 53:1, 257–269. (see page 29).

Benjamini, Y. and Y. Hochberg (1995). **Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing**. *Journal of the Royal statistical society: series B (Methodological)* 57:1, 289–300. (see page 60).

Berghuijs, W. R., S. T. Allen, S. Harrigan, and J. W. Kirchner (2019). **Growing Spatial Scales of Synchronous River Flooding in Europe**. *Geophysical Research Letters* 46:3, 1423–1428. (see pages 2, 5, 58, 59, 63, 65).

Berghuijs, W. R., R. A. Woods, C. J. Hutton, and M. Sivapalan (2016). **Dominant flood generating mechanisms across the United States**. *Geophysical Research Letters* 43:9, 4382–4390. (see pages 4, 17, 40).

Bertola, M., A. Viglione, and G. Blöschl (2019). **Informed attribution of flood changes to decadal variation of atmospheric, catchment and river drivers in Upper Austria**. *Journal of Hydrology* 577. (see page 4).

Bertola, M., A. Viglione, D. Lun, J. Hall, and G. Blöschl (2020). **Flood trends in Europe: Are changes in small and big floods different?** *Hydrology and Earth System Sciences* 24:4, 1805–1822. (see page 17).

Best, J. (2019). **Anthropogenic stresses on the world's big rivers**. *Nature Geoscience* 12:1, 7–21. (see page 2).

Blake, D., P. Nyman, H. Nice, F. M. D'souza, C. R. Kavazos, and P. Horwitz (2020). **Assessment of post-wildfire erosion risk and effects on water quality in southwestern Australia**. *International Journal of Wildland Fire* 29:3, 240–257. (see page 84).

Blöschl, G., S. Ardoin-Bardin, M. Bonell, M. Dorninger, D. Goodrich, D. Gutknecht, D. Matamoros, B. Merz, P. Shand, and J. Szolgay (2007). **At what scales do climate variability and land cover change impact on flooding and low flows?** *Hydrological Processes* 21:9, 1241–1247. (see page 16).

Blöschl, G., J. Hall, J. Parajka, R. A. Perdigão, B. Merz, B. Arheimer, G. T. Aronica, A. Bilibashi, O. Bonacci, M. Borga, I. Čanjevac, A. Castellarin, G. B. Chirico, P. Claps, K. Fiala, N. Frolova, L. Gorbachova, A. Gül, J. Hannaford, S. Harrigan, M. Kireeva, A. Kiss, T. R. Kjeldsen, S. Kohnová, J. J. Koskela, O. Ledvinka, N. Macdonald, M. Mavrova-Guirguinova, L. Mediero, R. Merz, P. Molnar, A. Montanari, C. Murphy, M. Osuch, V. Ovcharuk, I. Radevski, M. Rogger, J. L. Salinas, E. Sauquet, M. Šraj, J. Szolgay, A. Viglione, E. Volpi, D. Wilson, K. Zaimi, and N. Živković (2017). **Changing climate shifts timing of European floods**. *Science* 357:6351, 588–590. (see pages 15, 45, 58, 67).

Blöschl, G., J. Hall, A. Viglione, R. Perdigão, J. Parajka, B. Merz, D. Lun, B. Arheimer, G. Aronica, A. Bilibashi, M. Boháč, O. Bonacci, M. Borga, I. Čanjevac, A. Castellarin, G. Chirico, P. Claps, N. Frolova, D. Ganora, L. Gorbachova, A. Gül, J. Hannaford, S. Harrigan, M. Kireeva, A. Kiss, T. Kjeldsen, S. Kohnová, J. Koskela, O. Ledvinka, N. Macdonald, M. Mavrova-Guirguinova, L. Mediero, R. Merz, P. Molnar, A. Montanari, C. Murphy, M. Osuch, V. Ovcharuk, I. Radevski, J. Salinas, E. Sauquet, M. Šraj, J. Szolgay, E. Volpi, D. Wilson, K. Zaimi, and N. Živković (2019). **Changing climate both increases and decreases European floods**. *Nature* 573, 108–111. (see pages 4, 16, 58, 62, 64, 65, 67).

Blum, A. G., P. J. Ferraro, S. A. Archfield, and K. R. Ryberg (2020). **Causal Effect of Impervious Cover on Annual Flood Magnitude for the United States**. *Geophysical Research Letters* 47:5. (see pages 2, 12, 16).

Boccaletti, S., V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang (2006). **Complex networks: Structure and dynamics**. *Physics Reports* 424:4-5, 175–308. (see page 19).

Boer, M. M., V. R. de Dios, and R. A. Bradstock (2020). **Unprecedented burn area of Australian mega forest fires**. *Nature Climate Change* 10:3, 171–172. (see page 82).

Borchers Arriagada, N., A. J. Palmer, D. M. Bowman, G. G. Morgan, B. B. Jalaludin, and F. H. Johnston (2020). **Unprecedented smoke-related health burden associated**

with the 2019–20 bushfires in eastern Australia. *Medical Journal of Australia* 213:6, 282–283. (see page 82).

Bowman, D., G. Williamson, M. Yebra, J. Lizundia-Loiola, M. L. Pettinari, S. Shah, R. Bradstock, and E. Chuvieco (2020). **Wildfires: Australia needs national monitoring agency**. *Nature* 584, 188–191. (see pages 82, 85).

Bradstock, R. A., R. Nolan, L. Collins, V. Resco de Dios, H. Clarke, M. E. Jenkins, B. Kenny, and M. M. Boer (2020). **A broader perspective on the causes and consequences of eastern Australia's 2019-20 season of mega-fires: A response to Adams et al.** *Global Change Biology* 26:7. (see page 82).

Breiman, L. (2001). **Random forests**. *Machine Learning* 45, 5–32. (see pages 13, 20).

Brunner, M. I. and E. Gilleland (2021). **Complex High- and Low-Flow Networks Differ in Their Spatial Correlation Characteristics, Drivers, and Changes**. *Water Resources Research* 57:9. (see page 2).

Brunner, M. I., E. Gilleland, A. Wood, D. L. Swain, and M. Clark (2020). **Spatial dependence of floods shaped by spatiotemporal variations in meteorological and land-surface processes**. *Geophysical Research Letters* 47:13, 1–13. (see page 35).

Chafer, C. J. (2008). **A comparison of fire severity measures: an Australian example and implications for predicting major areas of soil erosion**. *Catena* 74:3, 235–245. (see page 83).

Chou, C., J. C. Chiang, C. W. Lan, C. H. Chung, Y. C. Liao, and C. J. Lee (2013). **Increase in the range between wet and dry season precipitation**. *Nature Geoscience* 6:4, 263–267. (see page 37).

Clauset, A., M. E. Newman, and C. Moore (2004). **Finding community structure in very large networks**. *Physical Review E* 70:6, 066111. (see page 19).

Conticello, F. R., F. Cioffi, U. Lall, and B. Merz (2020). **Synchronization and Delay Between Circulation Patterns and High Streamflow Events in Germany**. *Water Resources Research* 56:4, 1–16. (see page 19).

Conticello, F. R., F. Cioffi, B. Merz, and U. Lall (2018). **An event synchronization method to link heavy rainfall events and large-scale atmospheric circulation features**. *International Journal of Climatology* 38:3, 1421–1437. (see page 19).

Corradini, C., F. Melone, and V. P. Singh (1995). **Some remarks on the use of GIUH in the hydrological practice**. *Hydrology Research* 26:4-5, 297–312. (see pages 15, 60).

Davenport, F. V., J. E. Herrera-Estrada, M. Burke, and N. S. Diffenbaugh (2020). **Flood Size Increases Nonlinearly Across the Western United States in Response to Lower Snow-Precipitation Ratios**. *Water Resources Research* 56:1, 1–19. (see page 31).

Davey, S. M. and A. Sarre (2020). **Editorial: the 2019/20 Black Summer bushfires**. *Australian Forestry* 83:2, 47–51. (see page 78).

Deb, P., H. Moradkhani, P. Abbaszadeh, A. S. Kiem, J. Engström, D. Keellings, and A. Sharma (2020). **Causes of the widespread 2019–2020 Australian bushfire season**. *Earth's Future* 8:11, e2020EF001671. (see pages 78, 82).

Debano, L. F. and J. S. Krammes (1966). **Water repellent soils and their relation to wildfire temperatures**. *International Association of Scientific Hydrology. Bulletin* 11:2, 14–19. (see page 6).

Desai, B., A. Maskrey, P. Peduzzi, A. De Bono, and C. Herold (2015). **Making development sustainable: the future of disaster risk management, global assessment report on disaster risk reduction**. Tech. rep. (see page 1).

Do, H. X., Y. Mei, and A. D. Gronewold (2020). **To What Extent Are Changes in Flood Magnitude Related to Changes in Precipitation Extremes?** *Geophysical Research Letters* 47:18, 1–10. (see page 12).

Doocy, S., A. Daniels, S. Murray, and T. D. Kirsch (2013). **The human impact of floods: a historical review of events 1980-2009 and systematic literature review**. *PLoS currents* 5, ecurrents.dis.f4deb457904936b07c09daa98ee8171a. (see page 1).

Dottori, F., W. Szewczyk, J. C. Ciscar, F. Zhao, L. Alfieri, Y. Hirabayashi, A. Bianchi, I. Mongelli, K. Frieler, R. A. Betts, and L. Feyen (2018). **Increased human and economic losses from river flooding with anthropogenic warming**. *Nature Climate Change* 8:9, 781–786. (see page 58).

Falcone, J. A. (2017). **Coefficient-based consistent mapping of imperviousness in the conterminous U.S. at 60-m resolution for 1974, 1982, 1992, 2002, and 2012** (see pages 14, 46).

Falcone, J. A. and A. E. LaMotte (2016). **National 1-kilometer rasters of selected Census of Agriculture statistics allocated to land use for the time period 1950 to 2012** (see pages 15, 46).

Farquharson, F. A., J. R. Meigh, and J. V. Sutcliffe (1992). **Regional flood frequency analysis in arid and semi-arid areas**. *Journal of Hydrology* 138:3, 487–501. (see page 16).

Field, C. B., V. Barros, T. F. Stocker, Q. Dahe, D. Jon Dokken, K. L. Ebi, M. D. Mastrandrea, K. J. Mach, G. K. Plattner, S. K. Allen, M. Tignor, and P. M. Midgley (2012). **Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change**. Ed. by C. B.

Field, V. Barros, T. F. Stocker, and Q. Dahe. Cambridge: Cambridge University Press, 1–582 (see page 58).

Findell, K. L. and E. A. Eltahir (2003). **Atmospheric controls on soil moisture-boundary layer interactions. Part I: Framework development**. *Journal of Hydrometeorology* 4:3, 552–569. (see page 61).

French, N. H., E. S. Kasischke, R. J. Hall, K. A. Murphy, D. L. Verbyla, E. E. Hoy, and J. L. Allen (2008). **Using Landsat data to assess fire and burn severity in the North American boreal forest region: an overview and summary of results**. *International Journal of Wildland Fire* 17:4, 443–462. (see pages 83, 87).

Gallina, V., S. Torresan, A. Critto, A. Sperotto, T. Glade, and A. Marcomini (2016). **A review of multi-risk methodologies for natural hazards: Consequences and challenges for a climate change impact assessment**. *Journal of environmental management* 168, 123–132.

Gill, J. C. and B. D. Malamud (2016). **Hazard interactions and interaction networks (cascades) within multi-hazard methodologies**. *Earth System Dynamics* 7:3, 659–679. (see page 79).

Gilleland, E., M. Bukovsky, C. L. Williams, S. McGinnis, C. M. Ammann, B. G. Brown, and L. O. Mearns (2016). **Evaluating NARCCAP model performance for frequencies of severe-storm environments**. *Advances in Statistical Climatology, Meteorology and Oceanography* 2, 137–153. (see page 61).

Gilmour, D. A. (1968). **Water repellence of soils related to surface dryness**. *Australian Forestry* 32:3, 143–148. (see page 6).

Gregorutti, B., B. Michel, and P. Saint-Pierre (2017). **Correlation and variable importance in random forests**. *Statistics and Computing* 27:3, 659–678. (see page 21).

Greve, P., B. Orlowsky, B. Mueller, J. Sheffield, M. Reichstein, and S. I. Seneviratne (2014). **Global assessment of trends in wetting and drying over land**. *Nature Geoscience* 7:10, 716–721. (see page 37).

Guo, J., H.-Y. Li, L. R. Leung, S. Guo, P. Liu, M. Sivapalan, Y. Cha, S. S. Park, K. Kim, M. Byeon, and C. A. Stow (2014). **Links between flood frequency and annual water balance behaviors: A basis for similarity and regionalization**. *Water Resources Research* 50:2, 937–953. (see page 16).

Haasnoot, M., J. H. Kwakkel, W. E. Walker, and J. Ter Maat (2013). **Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world**. *Global Environmental Change* 23:2, 485–498.

Hatchett, B. J. (2018). **Snow level characteristics and impacts of a spring typhoon-originating atmospheric river in the Sierra Nevada, USA**. *Atmosphere* 9:6, 1–13. (see page 29).

Hawcroft, M. K., L. C. Shaffrey, K. I. Hodges, and H. F. Dacre (2012). **How much Northern Hemisphere precipitation is associated with extratropical cyclones?** *Geophysical Research Letters* 39:24, 1–7. (see page 34).

Hegerl, G. C., H. Hanlon, and C. Beierkuhnlein (2011). **Elusive extremes**. *Nature Geoscience* 4:3, 142–143. (see page 79).

Hiemstra, P. H., E. J. Pebesma, C. J. Twenhöfel, and G. B. Heuvelink (2009). **Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network**. *Computers and Geosciences* 35:8, 1711–1721. (see page 60).

Hirabayashi, Y., R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim, and S. Kanae (2013). **Global flood risk under climate change**. *Nature Climate Change* 3:9, 816–821. (see page 1).

Hirsch, R. M. and K. R. Ryberg (2012). **Has the magnitude of floods across the USA changed with global CO 2 levels?** *Hydrological Sciences Journal* 57:1, 1–9. (see page 12).

Ho, M., U. Lall, M. Allaire, I. Pal, D. Raff, D. Wegner, N. Devineni, and H. H. Kwon (2017). **The future role of dams in the United States of America**. *Water Resources Research* 53, 982–998. (see page 2).

Hodgkins, G. A., R. W. Dudley, S. A. Archfield, and B. Renard (2019). **Effects of climate, regulation, and urbanization on historical flood trends in the United States**. *Journal of Hydrology* 573, 697–709. (see pages 12, 36).

Hodgkins, G. A., P. H. Whitfield, D. H. Burn, J. Hannaford, B. Renard, K. Stahl, A. K. Fleig, H. Madsen, L. Mediero, J. Korhonen, C. Murphy, and D. Wilson (2017). **Climate-driven variability in the occurrence of major floods across North America and Europe**. *Journal of Hydrology* 552, 704–717. (see pages 4, 12).

Homer, C., J. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown (2015). **Completion of the 2011 national land cover database for the conterminous United States – Representing a decade of land cover change information**. *Photogrammetric Engineering and Remote Sensing* 81:5, 345–354. (see pages 15, 46).

Hothorn, T., K. Hornik, and A. Zeileis (2006). **Unbiased recursive partitioning: A conditional inference framework**. *Journal of Computational and Graphical Statistics* 15:3, 651–674. (see page 20).

Hu, H., F. Dominguez, Z. Wang, D. A. Lavers, G. Zhang, and F. M. Ralph (2017). **Linking atmospheric river hydrological impacts on the U.S. West Coast to Rossby wave breaking**. *Journal of Climate* 30:9, 3381–3399. (see page 29).

Hughes, L., W. Steffen, G. Mullins, A. Dean, E. Weisbrot, and M. Rice (2020). **Summer of crisis**. (see pages 78, 82).

Irannezhad, M., J. Liu, B. Ahmadi, and D. Chen (2020). **The dangers of Arctic zombie wildfires**. *Science* 369:6508, 1171–1171. (see page 85).

Ivancic, T. J. and S. B. Shaw (2015). **Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge**. *Climatic Change* 133:4, 681–693. (see pages 3, 12).

John, V. O., R. P. Allan, and B. J. Soden (2009). **How robust are observed and simulated precipitation responses to tropical ocean warming?** *Geophysical Research Letters* 36:14, 1–5. (see page 4).

Jones, D. A., W. Wang, and R. Fawcett (2009). **High-quality spatial climate data-sets for Australia**. *Australian Meteorological and Oceanographic Journal* 58:4, 233–248. (see pages 82, 88).

Jongman, B., S. Hochrainer-Stigler, L. Feyen, J. C. Aerts, R. Mechler, W. J. Botzen, L. M. Bouwer, G. Pflug, R. Rojas, and P. J. Ward (2014). **Increasing stress on disaster-risk finance due to large floods**. *Nature Climate Change* 4, 264–268. (see page 58).

Jongman, B., H. C. Winsemius, J. C. Aerts, E. Coughlan De Perez, M. K. Van Aalst, W. Kron, and P. J. Ward (2015). **Declining vulnerability to river floods and the global benefits of adaptation**. *Proceedings of the National Academy of Sciences of the United States of America* 112:18, E2271–E2280. (see page 1).

Kablick III, G., D. R. Allen, M. D. Fromm, and G. E. Nedoluha (2020). **Australian pyroCb smoke generates synoptic-scale stratospheric anticyclones**. *Geophysical Research Letters* 47:13, e2020GL088101. (see page 78).

Kappes, M. S., M. Keiler, K. von Elverfeldt, and T. Glade (2012). **Challenges of analyzing multi-hazard risk: a review**. *Natural hazards* 64:2, 1925–1958. (see page 80).

Kemter, M., B. Merz, N. Marwan, S. Vorogushyn, and G. Blöschl (2020). **Joint Trends in Flood Magnitudes and Spatial Extents Across Europe**. *Geophysical Research Letters* 47:7, 1–8. (see pages 16, 17, 40).

Key, C. H. and N. C. Benson (2002). **Remote sensing measure of severity, the normalized burn ratio**. *In J. L. Coffelt & R. K. Livingston (Eds.), Fire effects monitoring and inventory protocol, landscape assessment (p. 55).* (see pages 83, 86, 87).

Key, C. H. and N. C. Benson (2006). **Landscape assessment (LA)**. *In: Lutes, Duncan C.; Keane, Robert E.; Caratti, John F.; Key, Carl H.; Benson, Nathan C.; Sutherland, Steve; Gangi, Larry J. 2006. FIREMON: Fire effects monitoring and inventory system. Gen. Tech. Rep. RMRS-GTR-164-CD. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.* 164, 1–51. (see pages 83, 86, 87).

Kharin, V. V. and F. W. Zwiers (2005). **Estimating extremes in transient climate change simulations**. *Journal of Climate* 18:8, 1156–1173. (see page 85).

King, A. D., A. J. Pitman, B. J. Henley, A. M. Ukkola, and J. R. Brown (2020). **The role of climate variability in Australian drought**. *Nature Climate Change* 10:3, 177–179. (see page 81).

Kinnell, P. (2010). **Event soil loss, runoff and the Universal Soil Loss Equation family of models: A review**. *Journal of hydrology* 385:1-4, 384–397. (see pages 83, 84).

Klein Tank, A. M., J. B. Wijngaard, G. P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A. F. Van Engelen, E. Forland, M. Mietus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar, J. Antonio López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L. V. Alexander, and P. Petrovic (2002). **Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment**. *International Journal of Climatology* 22:12, 1441–1453. (see page 60).

Kochanek, K., W. G. Strupczewski, and E. Bogdanowicz (2012). **On seasonal approach to flood frequency modelling. Part II: Flood frequency analysis of Polish rivers**. *Hydrological Processes* 26:5, 717–730. (see page 4).

Kormann, C., T. Francke, M. Renner, and A. Bronstert (2015). **Attribution of high resolution streamflow trends in Western Austria - An approach based on climate and discharge station data**. *Hydrology and Earth System Sciences* 19:3, 1225–1245. (see page 20).

Kron, W., M. Steuer, P. Löw, and A. Wirtz (2012). **How to deal properly with a natural catastrophe database - Analysis of flood losses**. *Natural Hazards and Earth System Science* 12:3, 535–550. (see page 1).

Kundzewicz, Z. W., S. Kanae, S. I. Seneviratne, J. Handmer, N. Nicholls, P. Peduzzi, R. Mechler, L. M. Bouwer, N. Arnell, K. Mach, R. Muir-Wood, G. R. Brakenridge, W. Kron, G. Benito, Y. Honda, K. Takahashi, and B. Sherstyukov (2014). **Flood risk and climate change: global and regional perspectives**. *Hydrological Sciences Journal* 59:1, 1–28. (see page 58).

Larsen, I. J. and L. H. MacDonald (2007). **Predicting postfire sediment yields at the hillslope scale: Testing RUSLE and Disturbed WEPP**. *Water Resources Research* 43:11. (see page 84).

Lavers, D. A. and G. Villarini (2013). **Atmospheric rivers and flooding over the central United States**. *Journal of Climate* 26:20, 7829–7836. (see page 32).

Lentile, L. B., Z. A. Holden, A. M. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. A. Lewis, P. E. Gessler, and N. C. Benson (2006). **Remote sensing techniques to assess active fire characteristics and post-fire effects**. *International Journal of Wildland Fire* 15:3, 319–345. (see pages 83, 87, 88).

Li, H., L. Ding, M. Ren, C. Li, and H. Wang (2017). **Sponge city construction in China: A survey of the challenges and opportunities**. *Water (Switzerland)* 9:9, 1–17.

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier (2013). **A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions**. *Journal of Climate* 26:23, 9384–9392. (see pages 14, 39).

Lochbihler, K., G. Lenderink, and A. P. Siebesma (2017). **The spatial extent of rainfall events and its relation to precipitation scaling**. *Geophysical Research Letters* 44:16, 8629–8636. (see pages 5, 62).

Magilligan, F. J. and K. H. Nislow (2005). **Changes in hydrologic regime by dams**. *Geomorphology* 71:1-2, 61–78. (see page 17).

Malik, N., N. Marwan, and J. Kurths (2010). **Spatial structures and directionalities in Monsoonal precipitation over South Asia**. *Nonlinear Processes in Geophysics* 17:5, 371–381. (see page 18).

Mallakpour, I. and G. Villarini (2015). **The changing nature of flooding across the central United States**. *Nature Climate Change* 5:3, 250–254. (see pages 4, 12).

Mangini, W., A. Viglione, J. Hall, Y. Hundecha, S. Ceola, A. Montanari, M. Rogger, J. L. Salinas, I. Borzì, and J. Parajka (2018). **Detection of trends in magnitude and frequency of flood peaks across Europe**. *Hydrological Sciences Journal* 63:4, 493–512. (see page 58).

Mann, H. B. (1945). **Nonparametric Tests Against Trend**. *Econometrica* 13:3, 245–259. (see page 22).

Mann, H. B. and D. R. Whitney (1947). **On a test of whether one of two random variables is stochastically larger than the other**. *The annals of mathematical statistics* 18:1, 50–60. (see page 21).

Marks, D., J. Kimball, D. Tingey, and T. Link (1998). **The sensitivity of snowmelt processes to climate conditions and forest cover during rain-on-snow: a case study of the 1996 Pacific Northwest flood**. *Hydrological Processes* 12:10-11, 1569–1587. (see page 16).

McAneney, J., B. Sandercock, R. Crompton, T. Mortlock, R. Musulin, R. Pielke Jr, and A. Gissing (2019). **Normalised insurance losses from Australian natural disasters: 1966–2017**. *Environmental Hazards* 18:5, 414–433. (see page 82).

McGuire, V. L. (2017). **Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2015 and 2013–15**. Tech. rep. USGS, 1–14. URL: https://pubs.usgs.gov/sir/2017/5040/sir20175040.pdf (see page 32).

Mediero, L., T. R. Kjeldsen, N. Macdonald, S. Kohnova, B. Merz, S. Vorogushyn, D. Wilson, T. Alburquerque, G. Blöschl, E. Bogdanowicz, A. Castellarin, J. Hall, M. Kobold, J. Kriauciuniene, M. Lang, H. Madsen, G. Onuşluel Gül, R. A. Perdigão, L. A. Roald, J. L. Salinas, A. D. Toumazis, N. Veijalainen, and Ó. órarinsson (2015). **Identification of coherent flood regions across Europe by using the longest streamflow records**. *Journal of Hydrology* 528, 341–360. (see page 61).

Merz, B., G. Blöschl, S. Vorogushyn, F. Dottori, J. C. Aerts, P. Bates, M. Bertola, M. Kemter, H. Kreibich, U. Lall, and E. Macdonald (2021). **Causes, impacts and patterns of disastrous river floods**. *Nature Reviews Earth and Environment* 2:9, 592–609. (see pages 1, 2).

Merz, B., J. Hall, M. Disse, and A. Schumann (2010). **Fluvial flood risk management in a changing world**. *Natural Hazards and Earth System Science* 10:3, 509–527.

Merz, R. and G. Blöschl (2003). **A process typology of regional floods**. *Water Resources Research* 39:12, 1–20. (see pages 4, 65).

Metzger, A., F. Marra, J. A. Smith, and E. Morin (2020). **Flood frequency estimation and uncertainty in arid/semi-arid regions**. *Journal of Hydrology* 590, 125254. (see page 16).

Milly, P., J. Betancourt, M. Falkenmark, R. Hirsch, Z. Kundzewicz, D. Lettenmaier, and R. Stouffer (2008). **Stationarity is dead: whither water management?** *Science* 319, 573–574. (see page 2).

Molnar, P., S. Fatichi, L. Gaál, J. Szolgay, and P. Burlando (2015). **Storm type effects on super Clausius-Clapeyron scaling of intense rainstorm properties with air temperature**. *Hydrology and Earth System Sciences* 19:4, 1753–1766. (see pages 5, 62).

Munich Re (2021). **NatCatService**. Tech. rep. URL: www.munichre.com (visited on 01/15/2022) (see page 12).

Munoz, S. E., L. Giosan, M. D. Therrell, J. W. Remo, Z. Shen, R. M. Sullivan, C. Wiman, M. O'Donnell, and J. P. Donnelly (2018). **Climatic control of Mississippi River flood hazard amplified by river engineering**. *Nature* 556, 95–98.

Nayak, M. A. and G. Villarini (2017). **A long-term perspective of the hydroclimatological impacts of atmospheric rivers over the central United States**. *Water Resources Research* 53:2, 1144–1166. (see pages 32–34).

Neiman, P. J., L. J. Schick, F. Martin Ralph, M. Hughes, and G. A. Wick (2011). **Flooding in western washington: The connection to atmospheric rivers**. *Journal of Hydrometeorology* 12:6, 1337–1358. (see page 29).

Neri, A., G. Villarini, L. J. Slater, and F. Napolitano (2019). **On the statistical attribution of the frequency of flood events across the U.S. Midwest**. *Advances in Water Resources* 127:February, 225–236. (see page 33).

Newman, M. E. (2003). **The structure and function of complex networks**. *SIAM Review* 45:2, 167–256. (see page 19).

Nolan, R. H., M. M. Boer, L. Collins, V. Resco de Dios, H. G. Clarke, M. Jenkins, B. Kenny, and R. A. Bradstock (2020). **Causes and consequences of eastern Australia's 2019-20 season of mega-fires**. *Global change biology* 26, 1039–1041. (see page 81).

Oudin, L., B. Salavati, C. Furusho-Percot, P. Ribstein, and M. Saadi (2018). **Hydrological impacts of urbanization at the catchment scale**. *Journal of Hydrology* 559, 774–786. (see pages 2, 39).

Pachepsky, Y., D. E. Radcliffe, and H. M. Selim (2003). **Scaling methods in soil physics**. CRC press (see page 63).

Parliament of Victoria (2010). **Final Report**. 2009 Victorian Bushfires Royal Commission. URL: http://royalcommission.vic.gov.au/Commission-Reports/Final-Report/Volume-1/High-Resolution-Version.html (visited on 01/15/2022) (see page 82).

Pescaroli, G. and D. Alexander (2015). **A definition of cascading disasters and cascading effects: Going beyond the "toppling dominos" metaphor**. *Planet@ risk* 3:1, 58–67. (see page 79).

Petheram, C., N. Potter, J. Vaze, F. Chiew, and L. Zhang (2011). **Towards better understanding of changes in rainfall-runoff relationships during the recent drought**

in south-eastern Australia. *MODSIM 2011 - 19th International Congress on Modelling and Simulation - Sustaining Our Future: Understanding and Living with Uncertainty*: December, 3622–3628. (see page 6).

Pickrell, J. and E. Pennisi (2020). **Record US and Australian fires raise fears for many species**. Tech. rep. (see page 85).

Porter, K., A. Wein, C. Alpers, A. Baez, P. Barnard, J. Carter, A. Corsi, J. Costner, D. Cox, T. Das, M. Dettinger, J. Done, C. Eadie, M. Eymann, J. Ferris, P. Gunturi, M. Hughes, R. Jarrett, L. Johnson, H. Dam Le-Griffin, D. Mitchell, S. Morman, P. Neiman, A. Olsen, S. Perry, G. Plumlee, M. Ralph, D. Reynolds, A. Rose, K. Schaefer, J. Serakos, W. Siembieda, J. Stock, D. Strong, I. Sue Wing, A. Tang, P. Thomas, K. Topping, C. Wills, L. Jones, and D. Cox (2011). **Overview of the ARkStorm scenario**. Tech. rep. URL: https://pubs.usgs.gov/of/2010/1312/ (visited on 01/15/2022).

Primo, C., F. D. Kelemen, H. Feldmann, and B. Ahrens (2019). **A regional atmosphere-ocean climate system model (CCLMv5.0clm7-NEMOv3.3-NEMOv3.6) over Europe including three marginal seas: on its stability and performance**. *Geoscientific Model Development Discussions*: June, 1–33. (see page 58).

Rasmussen, T. J. and C. A. Perry (2001). **Trends in peak flows of selected streams in Kansas**. Tech. rep., 62 (see page 32).

Raymond, C., R. M. Horton, J. Zscheischler, O. Martius, A. AghaKouchak, J. Balch, S. G. Bowen, S. J. Camargo, J. Hess, K. Kornhuber, M. Oppenheimer, A. C. Ruane, T. Wahl, and K. White (2020). **Understanding and managing connected extreme events**. *Nature climate change* 10:7, 611–621.

Renard, K. G., G. R. Foster, G. A. Weesies, and J. P. Porter (1991). **RUSLE: Revised universal soil loss equation**. *Journal of soil and Water Conservation* 46:1, 30–33. (see page 84).

Robinne, F. N., D. W. Hallema, K. D. Bladon, and J. M. Buttle (2020). **Wildfire impacts on hydrologic ecosystem services in North American high-latitude forests: A scoping review**. *Journal of Hydrology* 581.

Robinson, J. S. and M. Sivapalan (1997). **An investigation into the physical causes of scaling and heterogeneity of regional flood frequency**. *Water Resources Research* 33:5, 1045–1059. (see pages 15, 60).

Rood, S. B., S. G. Foster, E. J. Hillman, A. Luek, and K. P. Zanewich (2016). **Flood moderation: Declining peak flows along some Rocky Mountain rivers and the underlying mechanism**. *Journal of Hydrology* 536, 174–182. (see pages 31, 37).

Rottler, E., A. Bronstert, G. Bürger, and O. Rakovec (2021). **Projected changes in Rhine River flood seasonality under global warming**. *Hydrology and Earth System Sciences* 25:5, 2353–2371.

Rottler, E., C. Kormann, T. Francke, and A. Bronstert (2019). **Elevation-dependent warming in the Swiss Alps 1981–2017: Features, forcings and feedbacks**. *International Journal of Climatology* 39:5, 2556–2568.

Royal Commission into National Natural Disaster Arrangements (2020). **Report**. Tech. rep. URL: https://naturaldisaster.royalcommission.gov.au/publications/royal-commission-national-natural-disaster-arrangements-report (visited on 01/15/2022) (see pages 78, 85).

Rubinov, M. and O. Sporns (2010). **Complex network measures of brain connectivity: Uses and interpretations**. *NeuroImage* 52:3, 1059–1069. (see page 19).

Sauer, I., R. Reese, C. Otto, T. Geiger, S. Willner, B. Guillod, D. Bresch, and K. Frieler (2021). **Climate Signals in River Flood Damages Emerge under Sound Regional Disaggregation**. *Nature communications* 12:1, 1–11. (see page 12).

Schilling, K. E., P. W. Gassman, C. L. Kling, T. Campbell, M. K. Jha, C. F. Wolter, and J. G. Arnold (2014). **The potential for agricultural land use change to reduce flood risk in a large watershed**. *Hydrological Processes* 28:8, 3314–3325. (see page 17).

Sen, P. K. (1968). **Estimates of the Regression Coefficient Based on Kendall's Tau**. *Journal of the American Statistical Association* 63:324, 1379–1389. (see pages 15, 44, 59).

Shakesby, R. A. and S. H. Doerr (2006). **Wildfire as a hydrological and geomorphological agent**. *Earth-Science Reviews* 74:3-4, 269–307. (see pages 6, 78, 84).

Sharma, A., C. Wasko, and D. P. Lettenmaier (2018). **If Precipitation Extremes Are Increasing, Why Aren't Floods?** *Water Resources Research* 54:11, 8545–8551. (see pages 3, 12, 84).

Shepherd, T. G., E. Boyd, R. A. Calel, S. C. Chapman, S. Dessai, I. M. Dima-West, H. J. Fowler, R. James, D. Maraun, O. Martius, et al. (2018). **Storylines: an alternative approach to representing uncertainty in physical aspects of climate change**. *Climatic change* 151:3, 555–571.

Sieg, T., K. Vogel, B. Merz, and H. Kreibich (2019). **Seamless Estimation of Hydrometeorological Risk Across Spatial Scales**. *Earth's Future* 7:5, 574–581.

Skøien, J. O., G. Blöschl, and A. W. Western (2003). **Characteristic space scales and timescales in hydrology**. *Water Resources Research* 39:10. (see pages 5, 62).

Slater, L. J. and G. Villarini (2016). **Recent trends in U.S. flood risk**. *Geophysical Research Letters* 43:24, 428–436. (see pages 4, 12, 32).

Smith, H. G., G. J. Sheridan, P. N. Lane, P. Nyman, and S. Haydon (2011). **Wildfire effects on water quality in forest catchments: A review with implications for water supply**. *Journal of Hydrology* 396:1-2, 170–192. (see page 6).

Smith, J. A., M. L. Baeck, Y. Zhang, and C. A. Doswell (2001). **Extreme rainfall and flooding from supercell thunderstorms**. *Journal of Hydrometeorology* 2:5, 469–489. (see page 32).

Soverel, N. O., D. D. Perrakis, and N. C. Coops (2010). **Estimating burn severity from Landsat dNBR and RdNBR indices across western Canada**. *Remote Sensing of Environment* 114:9, 1896–1909. (see pages 83, 87).

Stein, L., M. P. Clark, W. J. M. Knoben, F. Pianosi, and R. Woods (2021). **Process oriented insights from interpretable machine learning - what influences flood generating processes?** Under review in Water Resources Research: https://doi.org/10.1002/essoar.10503704.1 (see pages 22, 23).

Stein, L., F. Pianosi, and R. Woods (2020). **Event-based classification for global study of river flood generating processes**. *Hydrological Processes* 34:7, 1514–1529. (see pages 4, 40).

Strobl, C., A. L. Boulesteix, A. Zeileis, and T. Hothorn (2007). **Bias in random forest variable importance measures: Illustrations, sources and a solution**. *BMC Bioinformatics* 8:25. (see page 20).

Tang, Q., E. R. Vivoni, F. Muñz-Arriola, and D. P. Lettenmaier (2012). **Predictability of evapotranspiration patterns using remotely sensed vegetation dynamics during the North American Monsoon**. *Journal of Hydrometeorology* 13:1, 103–121. (see page 14).

Tarasova, L., S. Basso, D. Wendi, A. Viglione, R. Kumar, and R. Merz (2020). **A Process-Based Framework to Characterize and Classify Runoff Events: The Event Typology of Germany**. *Water Resources Research* 56:5. (see pages 4, 40).

Tarasova, L., R. Merz, A. Kiss, S. Basso, G. Blöschl, B. Merz, A. Viglione, S. Plötner, B. Guse, A. Schumann, S. Fischer, B. Ahrens, F. Anwar, A. Bárdossy, P. Bühler, U. Haberlandt, H. Kreibich, A. Krug, D. Lun, H. Müller-Thomy, R. Pidoto, C. Primo, J. Seidel, S. Vorogushyn, and L. Wietzke (2019). **Causative classification of river flood events**. *Wiley Interdisciplinary Reviews: Water* 6:4, e1353. (see pages 4, 16, 40, 61).

Thober, S., R. Kumar, N. Wanders, A. Marx, M. Pan, O. Rakovec, L. Samaniego, J. Sheffield, E. F. Wood, and M. Zink (2018). **Multi-model ensemble projections of European river floods and high flows at 1.5, 2, and 3 degrees global warming**. *Environmental Research Letters* 13:1. (see page 68).

Trenberth, K. E. (2011). **Changes in precipitation with climate change**. *Climate Research* 47:1-2, 123–138. (see page 3).

UNDRR (2019). **Global Assessment Report on Disaster Risk Reduction**. Geneva, Switzerland: United Nations Office for Disaster Risk Reduction. URL: https://www.unisdr.org/we/inform/terminology%5C#letter-c (visited on 01/15/2022) (see pages 3, 5, 58).

Van Oldenborgh, G. J., F. Krikken, S. Lewis, N. J. Leach, F. Lehner, K. R. Saunders, M. van Weele, K. Haustein, S. Li, D. Wallom, et al. (2021). **Attribution of the Australian bushfire risk to anthropogenic climate change**. *Natural Hazards and Earth System Sciences* 21:3, 941–960. (see pages 78, 81, 82).

Vardoulakis, S., B. B. Jalaludin, G. G. Morgan, I. C. Hanigan, and F. H. Johnston (2020). **Bushfire smoke: urgent need for a national health protection strategy**. *The Medical Journal of Australia* 212:8, 349. (see page 78).

Villarini, G. and L. J. Slater (2017). **Climatology of Flooding in the United States**. Ed. by S. Cutter. Oxford: Oxford University Press (see page 12).

Villarini, G. and L. J. Slater (2018). **Examination of Changes in Annual Maximum Gauge Height in the Continental United States Using Quantile Regression**. *Journal of Hydrologic Engineering* 23:3, 06017010. (see page 12).

Villarini, G., J. A. Smith, M. L. Baeck, B. K. Smith, and P. Sturdevant-Rees (2013). **Hydrologic analyses of the July 17–18, 1996, flood in Chicago and the role of urbanization**. *Journal of hydrologic engineering* 18:2, 250–259. (see page 33).

Vogel, R. M., C. Yaindl, and M. Walter (2011). **Nonstationarity: Flood magnification and recurrence reduction factors in the united states**. *Journal of the American Water Resources Association* 47:3, 464–474. (see pages 12, 36).

Vogt, J., P. Soille, A. De Jager, E. Rimavičiūtė, W. Mehl, S. Foisneau, K. Bódis, J. Dusart, M. Paracchini, P. Haastrup, and C. Bamps (2007). **A pan-European River and Catchment Database**, 121–144. In: *Digital Terrain Modelling*. Ed. by R. J. Peckham and J. Gyozo. Heidelberg: Springer (see page 59).

Walsh, R. P. D. and D. M. Lawler (1981). **Rainfall seasonality: description, spatial patterns and change through time**. *Weather* 36:7, 201–208. (see pages 16, 44).

Walz, Y., S. W. Maier, S. W. Dech, C. Conrad, and R. R. Colditz (2007). **Classification of burn severity using Moderate Resolution Imaging Spectroradiometer (MODIS): A case study in the jarrah-marri forest of southwest Western Australia**. *Journal of Geophysical Research: Biogeosciences* 112:G2. (see pages 83, 87).

Wang, X. L., Y. Feng, G. P. Compo, V. R. Swail, F. W. Zwiers, R. J. Allan, and P. D. Sardeshmukh (2013). **Trends and low frequency variability of extra-tropical cyclone activity in the ensemble of twentieth century reanalysis**. *Climate Dynamics* 40:11-12, 2775–2800. (see page 34).

Wasko, C. and R. Nathan (2019). **Influence of changes in rainfall and soil moisture on trends in flooding**. *Journal of Hydrology* 575:November 2018, 432–441. (see page 84).

Whelan, R. (2020). **Lessons to be learned in relation to the Australian bushfire season 2019-20**. Insurance Council of Australia. URL: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/Bushfirerecovery/Interim_Report (visited on 01/15/2022) (see pages 82, 83).

Yamazaki, D., D. Ikeshima, J. Sosa, P. D. Bates, G. Allen, and T. Pavelsky (2019). **MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets**. *Water Resources Research* 55:6, 5053–5073. (see page 59).

Yan, H., N. Sun, M. Wigmosta, R. Skaggs, L. R. Leung, A. Coleman, and Z. Hou (2019). **Observed Spatiotemporal Changes in the Mechanisms of Extreme Water Available for Runoff in the Western United States**. *Geophysical Research Letters* 46:2, 767–775. (see pages 3, 12, 29).

Yang, L., Y. Xu, L. Han, S. Song, X. Deng, and Y. Wang (2016). **River networks system changes and its impact on storage and flood control capacity under rapid urbanization**. *Hydrological Processes* 30:13, 2401–2412. (see page 2).

Yang, W., H. Yang, and D. Yang (2019). **Identification of homogeneous regions in terms of flood seasonality using a complex network approach**. *Journal of Hydrology* 576, 726–735. (see pages 19, 35).

Yang, X. (2015). **Digital mapping of RUSLE slope length and steepness factor across New South Wales, Australia**. *Soil Research* 53:2, 216–225. (see page 84).

Yang, X., J. Gray, G. Chapman, Q. Zhu, M. Tulau, and S. McInnes-Clarke (2017). **Digital mapping of soil erodibility for water erosion in New South Wales, Australia**. *Soil Research* 56:2, 158–170. (see page 84).

Ye, S., H. Y. Li, L. R. Leung, J. Guo, Q. Ran, Y. Demissie, and M. Sivapalan (2017). **Understanding flood seasonality and its temporal shifts within the contiguous United States**. *Journal of Hydrometeorology* 18:7, 1997–2009. (see page 4).

Yeung, J. K., J. A. Smith, G. Villarini, A. A. Ntelekos, M. L. Baeck, and W. F. Krajewski (2011). **Analyses of the warm season rainfall climatology of the northeastern US using regional climate model simulations and radar rainfall fields**. *Advances in Water Resources* 34:2, 184–204. (see page 35).

Yin, J., P. Gentine, S. Zhou, S. C. Sullivan, R. Wang, Y. Zhang, and S. Guo (2018). **Large increase in global storm runoff extremes driven by climate and anthropogenic changes**. *Nature Communications* 9:1. (see pages 12, 37).

Zhao, G., H. Gao, and L. Cuo (2016). **Effects of urbanization and climate change on peak flows over the San Antonio River basin, Texas**. *Journal of Hydrometeorology* 17:9, 2371–2389. (see page 39).

Zscheischler, J., O. Martius, S. Westra, E. Bevacqua, C. Raymond, R. M. Horton, B. van den Hurk, A. AghaKouchak, A. Jézéquel, M. D. Mahecha, D. Maraun, A. M. Ramos, N. N. Ridder, W. Thiery, and E. Vignotto (2020). **A typology of compound weather and climate events**. *Nature Reviews Earth and Environment* 1:7, 333–347. (see pages 3, 5, 79, 80).

Zscheischler, J. and S. I. Seneviratne (2017). **Dependence of drivers affects risks associated with compound events**. *Science advances* 3:6, e1700263. (see page 85).

Zscheischler, J., B. Van Den Hurk, P. J. Ward, and S. Westra (2020). **Multivariate extremes and compound events**, 59–76. In: *Climate extremes and their implications for impact and risk assessment*. Ed. by J. Sillmann, S. Sippel, and S. Russo. Elsevier (see page 85).

Zscheischler, J., S. Westra, B. J. Van Den Hurk, S. I. Seneviratne, P. J. Ward, A. Pitman, A. Aghakouchak, D. N. Bresch, M. Leonard, T. Wahl, and X. Zhang (2018). **Future climate risk from compound events**. *Nature Climate Change* 8:6, 469–477. (see pages 3, 79).