Large sample statistical estimation of streamflow sensitivity to climate and land cover change

Bailey Joella Anderson

Hertford College

University of Oxford

Supervised by:

Prof. Louise Slater, University of Oxford

Prof. Simon Dadson, University of Oxford

Thesis submitted for the degree Doctor of Philosophy at the University of Oxford

October 2023

This thesis is dedicated to Bill Anderson, who would have read the whole thing in an afternoon.

Abstract

Some of the perennial questions in hydrology relate to how land cover change and precipitation affect streamflow change. These questions have been addressed in many ways over the years, but some common assumptions remain insufficiently challenged or addressed. The work presented in this thesis seeks to address some of these gaps using a large sample of observed hydrological, climatological, and land surface data. The increased availability of large-sample hydrological data in recent years and growing interest in the use of statistical models for causal inference with observed data have made it possible to conduct research which sheds light on patterns and drivers of hydrological non-stationarity, a topic that is a matter of increasing importance for the community.

I take advantage of large-sample hydrological methods and data availability to produce novel insights into hydrologic sensitivity and elasticity, defined throughout as the expected change in streamflow associated with a one percent change in another variable. I address the following interrelated questions: 1. What is the effect of tree cover change and urbanisation on streamflow in the United States? 2. How robust are single-site regression models relative to causally interpretable panel regression models in this space? 3. How does streamflow elasticity to precipitation vary spatially and across the streamflow distribution? and 4. How do climate and hydrological behaviour co-evolve in the context of streamflow elasticity?

In this thesis, I find small but statistically significant effects of urbanisation on median and high streamflow across the U.S. The effects of tree cover change on flow and urbanisation on low flows are not statistically significant. I interrogate the relationship between streamflow and precipitation across many segments of the flow distribution, and create a new approach for investigating these relationships. Termed "elasticity curves", this concept visually represents the elasticity of streamflow to average annual and seasonal precipitation across the entirety of the flow distribution. I cluster the curve according to their normalized shape and the resulting groups correspond, to some extent, with hydrologic signatures and catchment characteristics, including the baseflow index, slope of the flow duration curve, and the aridity index, among others. I posit that the shape of the curve corresponds to water storage capacity within a catchment. Thus, the normalized and clustered curve shapes might be used as a tool for understanding hydrologic behaviour relative to catchment storage. This work represents one of very few studies to investigate elasticity across different segments of the flow distribution simultaneously and offers new insights into hydrologic response to climate.

Finally, I investigate temporal variation in elasticity to precipitation for low, median, and high streamflow at a regional scale. This work shows high interannual and spatial variability with mean absolute year-to-year differences as high as 0.5 in some regions, relative to long-term averages typically ranging from about 1-2.5. Long term trends in regional-scale interannual elasticity were uncommon but present in some regions. Total absolute changes in elasticity based on Mann-Kendall trend test results in regions with significant trends range from 0.28 to 0.60 over the study period. The time period used to estimate elasticity may have an effect on the resultant value, because large year-to-year variability exists in some regions.

The results of this work have implications for hydrologic management and fundamental process understanding of hydrologic behaviour.

Acknowledgements

While it is true that everyone thanks their supervisors in their acknowledgements section, I really do consider myself to have been exceptionally lucky to have worked with my supervisors during my DPhil. To Simon, thank you for sharing your impressive depth of hydrological knowledge over the years, and for helping keep my work physically realistic. More than anything else, though, thank you for your encouragement and support, which has helped me get through some of the rougher moments of the DPhil. To Louise, I am so glad that I went to your lecture at Christ Church when you visited Oxford in 2017. Working with you throughout my masters and DPhil has been an absolute gift. I have grown immensely both as a scientist and person for having been your student, and I couldn't have asked for a better supervisor. A large part of my excitement to continue working in research comes from having had such amazing mentors throughout this process.

I am also grateful to the co-authors on my thesis chapters, many of whom have offered invaluable insight towards my work; my friends and colleagues in the water lab in the School of Geography and the Environment; and the hydrology group at the University of Freiburg for hosting me for a research visit.

To the communities of which I have been a part, including Sabali Pots studio, which has been a creative refuge, and the Oxford University Judo Club. One of the best things that has come out of my time doing the DPhil has been the hobbies I have been able to pursue and the friendships I have made. It is impossible to name all of the friends to whom I am grateful, but you know who you are. I am especially thankful for the many housemates who have made Oxford feel like home, particularly when social circles shrunk during the pandemic. I am grateful to friends who have provided emotional support, companionship in exercise, swimming, cycling (sometimes absurdly far), have fed me, celebrated, and commiserated with me throughout my time here. I am grateful to each of you and hope that I will tell you in person the specific reasons why I appreciate you so much.

Finally, to my family, and especially to my parents, Belinda and Arn Anderson, who have supported me in more ways than I could even describe, I am so grateful for you, thank you.

Table of Contents

List of Tables

List of Figures

1 Introduction

Hydrological hazards such as flooding and drought present challenges for water management, especially in a changing natural world. Understanding how environmental changes such as climate change, land cover changes, and human activity interact with the water cycle is especially important when considering non-stationarity of streamflow, defined at the simplest level as a process where the statistical properties of the flow distribution change significantly over time (Slater et al., 2020). Some of the most fundamental questions in hydrology relate to how the environment interacts with and affects the water cycle and the effects of climate change may be expressed through the hydrologic cycle (Douville et al., 2021) – with climatological shifts resulting in more extreme floods, droughts and elevated risks associated with water availability (Abbott et al., 2019). However, there is rarely a one-to-one relationship between climatological changes and streamflow change (Ivancic & Shaw, 2015; Sharma et al., 2018).

For these reasons, quantifying the relationship between environmental changes and streamflow is a perennial challenge in hydrology. In this thesis, I address some aspects of this challenge: namely, how land cover changes influence streamflow on average and how the responsiveness or sensitivity of streamflow to precipitation varies temporally and across the flow distribution, using statistical approaches applied to catchments in the United States. This body of work falls within the field of large-sample hydrology, and my use of observed data and large-sample statistical tools to infer relationships between hydrological variables allows for empirical hypothesis testing in a domain which is frequently dominated by simulation and modelling. This creates space to challenge existing theory and contribute positively to a growing data-intensive hydrological understanding (Peters-Lidard et al., 2017).

This thesis can be characterized as an assessment of the sensitivity of streamflow to environmental change. It is an exploration, at a large scale, of how water is partitioned through the components of the water balance at different rates, depending on the physical and climatological characteristics of hydrological catchments, and an acknowledgement of the non-constant and non-linear nature of these relationships.

The work presented here first offers a series of unique contributions regarding the roles which tree cover change and urbanisation play in historical changes in streamflow. Then, it presents a novel perspective on the sensitivity of streamflow to precipitation, characterised across the complete flow distribution. Finally, it challenges the assumption of a temporally-stationary relationship between precipitation and streamflow, which is present throughout much of the

literature. The scientific contributions offered by this work also include the application of contemporary methodological approaches.

1.1 Background and motivation

A common point of discussion in hydrology and climate change research relates to changing patterns in flood and low flow occurrence. This is often discussed in terms of non-stationarity, and it is true that there is some evidence of observed significant long-term changes in both streamflow and climate in the United States (U.S.), the geographic focus of this thesis. For instance, statistically significant trends in streamflow magnitude, as well as event frequency, have been detected in numerous locations across the U.S. (Archfield et al., 2016; Douglas et al., 2000; Rice et al., 2015; Sadri et al., 2016; Slater & Villarini, 2016a). However, in a majority of these studies, trends were not consistently increasing or decreasing across the country, and their spatial cohesion and extent were somewhat limited, meaning that patterns of changes in major floods did not closely follow regional divisions.

In fact, in spite of what a non-specialist might reasonably infer, changes in precipitation are often not directly reflected in streamflow response, and trends in heavy precipitation may be poor proxies for trends in flooding (Collins et al., 2022; Ivancic & Shaw, 2015; Sharma et al., 2018). This incongruence arises from the fact that streamflow, and subsequently flooding, does not result from precipitation alone and is instead subject to the pre-existing moisture conditions (e.g. antecedent moisture condition), as well as the physical characteristics of a catchment. Physical characteristics control, to a large degree, the storage and runoff rate of precipitation in a hydrological catchment, and there is evidence that changing landscape features such as urbanisation extent can have an effect on the proportional relationship between precipitation and streamflow.

The proportional change in streamflow for a 1% change in another variable is known as "streamflow elasticity". This is a simple, useful, metric for estimating the sensitivity or responsiveness of streamflow to environmental variables, typically precipitation, and is the primary hydrologic signature on which I focus in this thesis. Essentially a regression coefficient, elasticity is preferable to some other empirical metrics because it is a proportion, e.g. a ratio of percentages relative to average conditions, thus effectively capturing the effect of a change on the normal state of a system. In this way, elasticity allows for comparison across catchments with very different normal conditions.

Elasticity has been used for hydrologic projection as well as catchment classification. Spatial variation in elasticity has been widely studied in the U.S. The hydrologic signature is typically estimated for a single flow percentile, usually the central summary, and the possibility of temporal change has rarely been considered. These normative assumptions and approaches present a gap in our understanding of streamflow sensitivity to environmental changes. A reliance on statistical projection for engineering, infrastructure and the management of hydrological resources (François et al., 2019; Slater et al., 2020) means that non-stationarity in streamflow timeseries is an important topic for research in hydrology.

The application of stationary models assumes that the probability distributions of future events will resemble the historical record (Villarini et al., 2018). While substantial literature exists on the subject of how different landscape and climatological features influence streamflow responsiveness to precipitation and climate, there remains a good deal of uncertainty in the exact dynamics at play. For instance, the assumption of a temporally constant estimate of elasticity perhaps fails to adequately consider the effects of variations in surface and groundwater storage over time (ex. Bennett et al., 2018; Berghuijs & Slater, 2023), as is discussed in detail in the introduction of chapter 6.

Further, some empirical evidence (in addition to the works which comprise this thesis) implies that streamflow sensitivity is unlikely to be consistent across the entirety of the flow distribution (Harman et al., 2011). However, evidence to date is very limited since the metric is typically estimated for the central summary, or for specific percentiles in isolation, as is discussed and referenced extensively in chapter 5. Variation in response, or lack thereof, for low vs. high streamflow could carry implications for how elasticity, as a hydrologic signature, can be applied to relevant management problems.

In addition, techniques like pooled regression approaches, which combine information from many locations across time, present an opportunity for increased statistical robustness using observed data, despite fairly restrictive historical timeseries. They can help effectively control unobserved variable bias, and use other catchments as pseudo-counterfactuals. Examples of these approaches, such as panel regression, have become more widely-known in hydrology, so more robust estimates of hydrological sensitivity and change are now possible using relatively simple statistical models, a large-sample approach, and real-world data.

Large-sample hydrology, the field within which this thesis falls, relies on a comparative approach to learn from hydrological similarities and differences across large samples of

different hydrological catchments. Expanding sample sizes in this way helps to draw out underlying relationships and establish robust physical principles (Addor et al., 2020). The moniker "large-sample" in hydrology varies in actual size, as the number of catchments necessary to establish robust principles can vary widely, from tens of catchments to thousands, depending on the objective of the study (Addor et al., 2020). Statistical approaches applied to large data samples help ensure sufficient information is available to establish statistical significance and arrive at generalisable conclusions, which would be infeasible in research focussed on small datasets or individual catchments (Gupta et al., 2014). Further, the application of statistical approaches to large sample hydrological datasets allows for the classification of catchments according to their behaviour. This carries the potential to transfer information to ungauged locations (Hrachowitz et al., 2013), facilitates the identification of behavioural outliers which require additional attention (Andréassian et al., 2010), and allows for the estimation and reporting of uncertainty in most cases. Large-sample statistical approaches facilitate the examination of hydrological behaviour across many locations simultaneously, considering individual variation, and potentially leading to robust generalizable conclusions, which may or may not align with previously held hypotheses based on individual catchments.

For instance, it is generally regarded as true that a decrease in tree cover will result in an increase in streamflow and flooding (Douville et al., 2021; Maidment, 1992), and the opposite is expected for an increase in tree cover. However, recent large-sample empirical evidence demonstrates a much wider range of responses (Goeking & Tarboton, 2020). Similarly, while urbanisation is typically shown to increase flood risk, responses may be far more varied for streamflow occurring in drier periods. Uncertainty in all of these estimates poses a problem for their potential application to water and climate management issues, including nature-based solutions, natural flood management, flood risk management generally, and hydrological prediction in ungauged basins. Literature for these examples are discussed in depth in chapter 4.

Many of the questions addressed in this thesis could instead be addressed using hydrological or land surface models rather than an empirical statistical approach; indeed, similar questions to the ones I ask here have been addressed in this fashion many times. For instance, Buechel et al. (2022) examine the role of tree cover change on streamflow in the UK using a large land surface model, and others have used hydrological models to estimate streamflow elasticity to precipitation (ex. Vano & Lettenmaier, 2014). While modelling approaches are useful,

especially for simulating scenarios which are limited in the observed record but are plausible in the future, they can only tell us what we already know. These approaches require not only consideration of the data needed to test a given hypothesis, but also careful consideration of how to encode hypotheses as uniquely falsifiable predictions (Pfister & Kirchner, 2017). Further, detailed process-based modelling approaches are often limited at a large scale by the quantity of data required to suitably characterise any particular catchment (Harman et al., 2011). Empirical statistical approaches, on the other hand, carry their own challenges and susceptibility to data error, but allow for rigorous hypothesis testing within the confines of data availability (Peters-Lidard et al., 2017; Pfister & Kirchner, 2017) and allow for greater generalisation of catchment function by considering the ways in which sensitivity varies between catchments (Harman et al., 2011). Large sample empirical research can inform better hydrological and climatological modelling studies, hopefully resulting in more accurate simulations of hazards, associated risk, and better solutions to hydrological management problems.

By pooling observed data across many locations, we are able to interpret the drivers of hydrological change more robustly, allowing for a better understanding of the relationships between streamflow, climate, and the land surface, both historically and when projecting future conditions.

1.2 Aims and objectives

The principal objectives of this thesis are to advance understanding of the relationships between potentially changing environmental variables, namely land cover change and precipitation, and streamflow magnitude using a large sample of observed data and a statistically-derived approach to estimate elasticity.

My specific aims are: to assess the statistical robustness of a series of regression approaches for the quantification of the sensitivity of streamflow magnitude to land cover changes across the U.S; to determine the average influence of tree cover change and urbanisation on downstream flow magnitude in U.S. rivers; to develop an elasticity-based classification system for U.S. rivers which demonstrates the variable nature of elasticity across the annual and seasonal flow distribution at individual sites and between catchments; to assess the likely drivers of between-catchment variation in elasticity; and finally, to assess the suitability of an assumption of temporal stationarity in the streamflow-precipitation relationship.

The broader aims of this work are to generate information which is relevant to the management of hydrological resources and hazards. The papers which comprise my thesis share a number of common themes, including a consistent methodological approach, geographic range, and broader goals. Namely, I used regression-based models to infer the relationships of interest, focus on the catchments in the United States, and examined streamflow sensitivity to environmental variables. Further, each paper builds off of the work completed in the previous one. For clarity, any discussion of streamflow or flooding in this thesis refers to hydrological streamflow and fluvial flooding.

1.3   Questions and thesis structure

This thesis comprises seven chapters, three of which constitute original research. These chapters include this introduction (Ch. 1), a literature review (Ch. 2), a methods chapter (Ch. 3), the three core research chapters (Ch. 4-6), and a discussion and conclusions chapter (Ch. 7).

I have taken a publication-based approach and the core research chapters are comprised of one published article (Anderson et al., 2022), one article which is in the late stages of revision (Anderson et al., 2023), and one article which has been submitted. This meets the requirements of three submitted academic articles for a DPhil thesis. During my DPhil, I have co-authored an additional two published articles which complement this work, and which are cited but not included formally in this thesis (Lees et al., 2021; Slater et al., 2020). Signed statements from each co-author are included as appendices following each chapter. In the introduction, including the chapter summaries, methodology, and discussion and conclusions chapters I use the personal pronoun "I", while in the research chapters (4-6) I use the collective pronoun "we". This is because, while the research chapters are predominantly my own work, they were completed with the help of co-authors and were submitted to journals using the collective. This is retained in chapters 4-6 to maintain consistency with the published work andto acknowledge that contribution.

The core research chapters are summarised as follows:

Chapter 4: Statistical attribution of the influence of urban and tree cover change on streamflow: a comparison of large sample statistical approaches

In this paper, I explore the effect of tree cover change and urbanisation on streamflow and examine the performance of different modelling approaches. I do not frame the estimates as

elasticity, but they are calculated using the same methodology and could be characterised as such. I address the following specific questions:

1.   How are urbanisation and tree cover change associated with or affecting streamflow across the conterminous United States?

2.   How do the results of single-catchment and multi-catchment (panel) regression methods differ?

Chapter 5: Elasticity curves: a novel lens for interpreting the variable nature of the streamflow-precipitation relationship

In this paper, I explore streamflow elasticity to precipitation across the complete flow distribution and use the resultant elasticity curves as a tool for classifying hydrological behaviour and storage. While elasticity has been reported on widely, it is typically studied relative to individual components of the flow distribution. Here, I establish the concept of an elasticity curve, in which streamflow elasticity to precipitation is estimated for every 5th percentile of flow, and show that this combined information can be informative. I address the following specific research questions:

1.   Does elasticity curve shape vary systematically and predictably across catchments?

2.   What catchment attributes best explain between-catchment variation in elasticity curve shape?

Chapter 6: Streamflow sensitivity to precipitation shows large inter-annual and spatial variability

In this chapter, I explore how elasticity changes over time and the co-evolution of climatological and hydrological systems. It is typical to assume, implicitly or explicitly, that a single estimate of elasticity over the period of record captures the relationship between streamflow and precipitation. This, however, is unlikely to be physically realistic. In this paper, I build on the hypothesis that streamflow elasticity, particularly at lower flow quantiles, is largely controlled by the storage available within a catchment area. As available storage decreases (ex. soil approaches saturation), elasticity is likely to increase. Thus, in catchments where storage is changing, a change in elasticity may be expected. In this sense, climate and hydrological behaviour may evolve simultaneously – with climate change leading to a change not only in streamflow, but in hydrological response to precipitation in certain catchments. In

this chapter, I outline the theoretical argument for temporal and interannual variability in elasticity and assess these changes at a regional scale. I address the following specific research questions:

1.   Is interannual variation in elasticity well represented by the long-term average at the regional scale?

2.   Are statistically significant long-term trends in elasticity present?

2   Literature Review

I am primarily concerned with the sensitivity of streamflow to changes in land cover (Ch. 4) and changes in precipitation (Ch. 5 and 6). As mentioned previously, I take a statistical-empirical approach to address these questions and build my experiments and hypotheses using the hydrologic signature "streamflow elasticity". Streamflow elasticity is a (generally) empirically estimated metric representing the proportional change in streamflow associated with a proportional change in a variable of interest, and is usually represented as a percentage. Throughout this literature section, I explain the processes and concepts which relate to the questions addressed in this thesis, and discuss elasticity and other sensitivity metrics, in greater detail in section 3, but I re-introduce streamflow elasticity here first because it provides the context for the rest of the literature.

2.1   The hydrologic cycle

The hydrologic cycle (Figure 2.1) is arguably the most fundamental principle of hydrology. It describes the pathways by which water evaporates from the oceans and land surface, travels through the atmosphere, falls as precipitation, and is then filtered through the landscape. On the land surface, water may be held in storage through interception by vegetation, infiltration into soils and groundwater, or on the surface, for instance, as snow or in lakes and reservoirs. Alternatively, water may run off into rivers and streams, generating streamflow, the runoff which is held within the channel of a river or stream (Maidment, 1992).
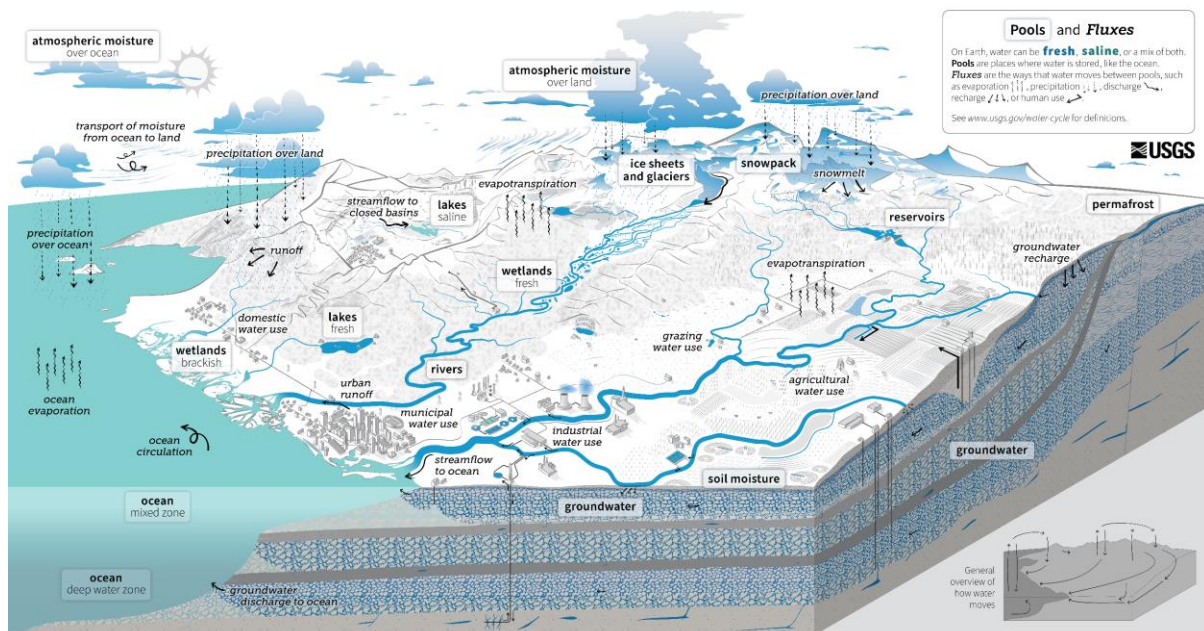
Figure 2.1 Graphical depiction of the hydrological cycle.
In the public domain (Corson-Dosch et al., 2022)*. a larger depiction of this diagram is included in appendix 1.*

2.1.1 Climate inputs

The primary climatological components of the water cycle are precipitation and evaporation. Precipitation (P) falls in various forms, or phases, but principally as snow and rain. Dominant precipitation phase in a catchment is typically the most influential control on runoff rate and timing in relatively natural catchments, and shifts between snow-dominated and rain-dominated hydrologic regimes may influence streamflow timing and magnitude. For instance, increasing average winter temperature can change the within-year distribution of streamflow (Berghuijs et al., 2014a; Jenicek et al., 2016; Stewart et al., 2005). In the U.S., catchments in which the dominant discharge to streamflow comes from snow meltwater tend to experience the highest flows in spring and summer due to melt processes, while low flow season is winter due to snow accumulation (Brunner et al., 2020). Rainfall dominated catchments may also exhibit different seasonal patterns which relate to precipitation regime patterns. For example, many catchments in the Pacific Northwest experience high flows in winter, the rainy season, and low flows in the autumn (Brunner et al., 2020). Meanwhile more arid catchments, for example throughout the Great Plains of the U.S. (Figure 2.4), may experience weak flow seasonality, as the majority of flow peaks are driven by individual precipitation events (Brunner et al., 2020; Farquharson et al., 1992). Types of precipitation events, such as convective storms,

and drivers of precipitation events, such as different climate indices, may also have an effect on the streamflow response, but because this work focusses on seasonal and annual scale streamflow, we do not consider these factors.

Potential evaporation (Ep) is the maximum amount of water vapor that could evaporate into the atmosphere under ideal conditions (Maidment, 1992). It is a conceptual entity, which indicates evaporative demand assuming no limit on water availability. Actual evaporation (E), on the other hand, is the amount water which is evaporated, taking into account limitations on the system, such as availability of water (Maidment, 1992). Actual evaporation can be difficult to measure, but is predominantly controlled by potential evaporation and available moisture (Blöschl, 2013; Budyko, 1974). Potential evaporation rate is controlled predominantly by temperature and solar radiation, although humidity, wind speed, and surface cover (including vegetation), may also be important. An abundance of equations for estimating Ep exist, which incorporate these variables to differing extents (Vörösmarty et al., 1998). Changes in land cover can result in shifts in potential evaporation.

2.1.2   Storage

Strong relationships exist between surface and subsurface storage and streamflow (McNamara et al., 2011), so the processes by which water is retained in a catchment for any period of time are essential in determining hydrologic sensitivity. Many studies consider the storage component of flow to be the contribution to streamflow from groundwater sources, although it also represents contributions from other stored sources (Hall, 1968; Smakhtin, 2001; Stoelzle et al., 2020). Depending on the physiography and climatology of a catchment, each of these sources could contribute to streamflow at different timescales (X. Zhang et al., 2020; Y. Zhang et al., 2022). Some draw a distinction between "active" and total water storage. Active storage refers to that which fills and releases water on timescales relevant to annual hydrological fluxes, although the boundaries between the two types can be fuzzy (McNamara et al., 2011; Pfister et al., 2017). A catchment's ability to store water determines, to a large extent, its capacity to buffer meteorological extremes and interannual variability in streamflow (Staudinger et al., 2017).

In the contiguous United States, for instance, snow and ice represent a source of water storage which, although seasonally delayed, is primarily formed and discharged within the same water year, contributing little to interannual variability in storage when water years are used. Other sources of hydrologic storage include water in lakes, wetlands, and reservoirs, soil water,

vegetation, and deep groundwater in aquifers. A shallow, rapidly recharging aquifer, for example, may contribute water to streamflow in under a year, while water entering a deeper, slowly recharging aquifer may take years or decades to re-enter streamflow naturally. Additionally, anthropogenic activity such as irrigation from groundwater sources may further influence the timescales of these interactions (Arrigoni et al., 2010; Kustu et al., 2011).

2.2   Potential drivers of hydrologic change

Natural climate variability, anthropogenic climate change, and land and water management decisions such as land cover changes, regulation and groundwater abstraction may all cause the distribution of streamflow magnitude to change over time. The specific causes of change in different locations can be wide-ranging and are often heavily debated. Some of these drivers are discussed in brief in the following section.

2.2.1   *Climate as a driver of non-stationarity*

Many studies have concluded that climate is the primary driver of changes in streamflow and hydrologic events like flooding (Berghuijs et al., 2017; Booij et al., 2019; Buechel et al., 2023). For example, evidence has shown that shifts in precipitation are strong predictors of flood frequency (Neri et al., 2019) as well as the distribution of flow magnitude (Slater & Villarini, 2017) in the Midwestern U.S. Changes in temperature due to greenhouse gas emissions may be responsible for altered streamflow timing and flood magnitude, with evidence in the mountainous western U.S. where snowmelt is a substantial contributor to flow (Davenport et al., 2020; Hidalgo et al., 2009; Stewart et al., 2005). Shifts in evaporation and snowmelt due to warming have been shown to be responsible for up to 30-50% of streamflow decline in the Colorado River basin (Douville et al., 2021; McCabe et al., 2017; Milly & Dunne, 2020; Woodhouse et al., 2016). Similar effects are likely in other locations.

Precipitation, potential evaporation, and temperature, as drivers of hydrologic non-stationarity are primarily controlled by natural climate variations and anthropogenically induced climate change, both of which may influence streamflow in different ways. Internal climate variability is the natural fluctuation in the Earth's climate that occurs over timescales ranging from months to centuries. These processes include large-scale modes of variability, such as the North Atlantic Oscillation (NAO), El Niño–Southern Oscillation (ENSO), or the Atlantic Multi-decadal Variability (AMV), among others (Seneviratne et al., 2021). These modes can affect the primary components of the water balance through changes in environmental conditions which result in differing rates of precipitation, evaporation, or snowmelt in line with a number

of dependent physical processes. Anthropogenic climate change, driven by greenhouse gas emissions, may further result in changes in the severity, frequency, or compound nature of climatic events which influence the water cycle in a similar fashion as internal variability (Seneviratne et al., 2021). It can be difficult to separate these processes from one another (Deser et al., 2012; H. Zhang & Delworth, 2018), and both may appear to represent non-stationarity in the observed hydrological record (Bayazit, 2015).

There is substantial evidence at the global scale that rare large precipitation events will increase in frequency as the amount of moisture that the atmosphere can hold increases in a warming climate (Seneviratne et al., 2021). Many have suggested that as climate change progresses, so too will shifts in the water cycle (Allan et al., 2020), potentially leading to greater floods, and shifts in the streamflow distribution overall. However, due to the intrinsic rarity of extreme events, observations of these have been somewhat limited and changes in precipitation extremes are often not reflected in flows (Sharma et al., 2018).

While the potential role of climate in driving streamflow change is clear, shifts in precipitation and temperature rarely fully explain variability and temporal trends in flow. Other anthropogenic drivers play an important additional role, and in some places may be outpacing the influence of climate change (Arrigoni et al., 2010; Douville et al., 2021; Vicente-Serrano et al., 2019). Beyond this, it is worth noting that climate and the land surface interact with one another and do not represent entirely independent variables. For example, snowpack represents both the primary contributor to streamflow and the primary source of groundwater recharge in large parts of the western U.S. (Rodell et al., 2018), thus changes in snowfall may effect streamflow directly and indirectly via pressure on groundwater-surface water interactions. Alternatively, land cover changes like urbanisation and water use like cropland irrigation may lead to changes in precipitation extremes (W. Zhang et al., 2018). The direct and indirect effects of land cover changes and water management on streamflow are touched on in the following sections.

2.2.2   Land cover change as a driver of non-stationarity

Land cover change represents a substantial source of potential human influence on the water cycle and streamflow specifically (Figure 2.1). The effects of land cover change on streamflow have been widely studied (Bassiouni et al., 2016; Blum et al., 2020; Buechel et al., 2022; Dudley et al., 2019; Neri et al., 2019; Prosdocimi et al., 2015; Villarini et al., 2009), but the quantification and projection of such impacts is challenging (Oudin et al., 2018a). Generalization of results from small sample hydrological studies remains an important aim

both for decision support, interpretation of physical processes, and for expanding information to less well understood systems (Blöschl et al., 2019; Hrachowitz et al., 2013), even considering the uniqueness of hydrological catchments (Beven, 2000b).

Land cover change can have a moderating or amplifying effect on streamflow. Altering the land surface can potentially lead to changes in the rate of runoff, evaporation, subsurface connectivity and drainage channels, the permeability and moisture storage capacity of soils, as well as interacting with and affecting climate at local scales (Rogger et al., 2017). Thus, land cover change can alter hydrological response to precipitation events by influencing the degree and the rate at which water is intercepted and evaporated, stored within a catchment, or allowed to run off into a river channel (Filoso et al., 2017; Jacobson, 2011; Shuster et al., 2005). The influence of land use and land cover changes on streamflow magnitude, are typically small (Buechel et al., 2023), but may be significant and particularly impactful depending on a number of factors.

Many land cover datasets are derived from satellite imagery and include classifications such as tree cover, cropland, grassland, urban area, and other broad categories of vegetation and surface cover. Each of these types of land surface are highly generalized, meaning that any large-scale, large sample, assessment of their influence on streamflow will be simplified. For instance, the category of "urban area" may look very different in a southern U.S. city vs. a European one, where norms around green space and building standards may differ widely. For this reason, assessing the influence of land cover changes on streamflow at a large scale is challenging. However, since questions regarding the influence of tree cover change (via deforestation and afforestation) and urbanisation on hydrology are important for management, generalized conclusions regarding their effects are still valuable. Changes in other land cover types certainly have an influence on flow, and a shift from one land cover type always results in reciprocal shifts in others. For instance, a great deal of research has demonstrated the influence of agricultural land and croplands on streamflow and flood risk (Neri et al., 2019; Poff et al., 2006; Schilling et al., 2014; D. Wang & Hejazi, 2011; Y.-K. Zhang & Schilling, 2006). However, I focus on urbanisation and tree cover change because of their prevalence in the literature, because of the evidence that urbanisation has a relatively large effect (Blum et al., 2020), and because tree cover has been proposed as potential nature-based solutions and as a flood management approach (Bastin et al., 2019; Buechel et al., 2022; Cohen-Shacham et al., 2016; Dadson et al., 2017).

2.2.2.1 Urbanisation

The effects of urbanisation on streamflow have been a key area of focus for many decades (e.g. Espey et al., 1965; Hollis, 1975) and studies have typically indicate that an increase in urban area results in increased streamflow and runoff, particularly for high and mean flows (Chen et al., 2017; Hodgkins et al., 2019; Oudin et al., 2018a; Shuster et al., 2005). For example, in a 2020 study which forms the methodological basis of much of the work carried out in this thesis, Blum et al. detected relatively large changes in annual floods (between 1.9 and 5.6% depending on the data sample used) in response to a 1% increase in urban area. Effect sizes, however, vary throughout the literature, and some exceptions exist (Brandes et al., 2005; Hopkins et al., 2015; McPhillips et al., 2019; Poff et al., 2006).

The primary mechanism by which urbanisation may affect streamflow is increased impervious surface area. Impervious surfaces, those which allow little infiltration, may increase "flashiness" resulting in faster and large rises in the flood hydrograph and thus higher peak and mean flows (Baker et al., 2004; Jacobson, 2011; Miller et al., 2014). However, the degree to which flashiness increases may depend on the fragmentation of urban area and the topography of a catchment (Jacobson, 2011). Increased impervious surface area has been shown to decrease low flows, as storm water runs off more quickly, rather than being held in a catchment (Jacobson, 2011).

Further, the effects of urbanisation may vary depending on the climate and previous conditions of a catchment. For instance, in dry urban areas, actual evaporation may be higher than in surrounding areas as irrigating non-native vegetation such as trees and lawns can contribute additional water to the catchment (Litvak et al., 2017; C. Wang et al., 2016). In wetter areas, urbanisation may result in reduced or increased low flows, depending on a number of factors. For instance, stormwater detention, imported waters, leaky infrastructure, and altered subsurface pathways may contribute to increased low flows (Cuo, 2016; Luthy et al., 2015), while impervious surface coverage and shallow groundwater withdrawal may result in decreased low flows (Meyer, 2002, 2005; Price, 2011). Thus, low flow response to urbanisation may be inconsistent across catchments (Price, 2011).

2.2.2.2 Tree cover

The influence of tree cover change on streamflow is less clear cut than urbanisation (Table 2.1), although some general rules of thumb do exist. For instance, afforestation is typically expected to result in decreased flows and deforestation to result in increased streamflow (Ahn

14

& Merwade, 2017; Booij et al., 2019; Brown et al., 2005; Farley et al., 2005, 2005; Goeking & Tarboton, 2020; Hibbert, 1965; M. Zhang et al., 2017). However, a number of studies (Table 2.1) have reported conflicting or non-existent streamflow response to afforestation (Bart et al., 2016; Biederman et al., 2014, 2015; Goeking & Tarboton, 2020, 2020; M. Zhang et al., 2017). Evidence suggests that the influence of deforestation on streamflow may also depend on the cause of forest loss (Goeking & Tarboton, 2020). For instance, stand replacing (e.g. severe fire or harvest) vs. non-stand replacing drivers (ex. drought, insects, low severity fire) may have very different effects on the vegetation structure, climate, and topography all play a role in determining response (Goeking & Tarboton, 2020). Other factors which potentially influence the effect of forest cover change on streamflow include drainage area, the total extent of tree cover area (Buechel et al., 2022), stand age, tree type, or post-disturbance vegetation type (Bart et al., 2016), among others. However, hydrologic behaviour is not perfectly explained by any of these drivers. In other words, while the general rules of thumb exist regarding streamflow response to forest cover change, actual response is complex and difficult to predict.

The mechanisms by which tree cover change influences streamflow are similar to those of urbanisation and that tree cover change alters the storage, run off, and evaporative components of the water balance. This can occur by for instance, changing surface roughness increasing evaporation following stand loss, shifts in transpiration which may include increases or decreases depending on post-disturbance vegetative growth (Goeking & Tarboton, 2020; Page et al., 2020).

Table 2.1 Hydrologic response to tree cover change in 78 papers in a review

Number of papers which found increases, decreases, and no change in hydrologic signatures plus response to particular types of disturbances in Goeking and Tarboton (2020) review.

Response	Total number of studies	Increase	No change	Decrease
Streamflow (annual water yield	31	26	16	9
Peak flow magnitude	22	19	10	7
Peak flow timing	18	14	7	4
Low flow magnitude	25	14	9	9
Maximum snow water equivalent	42	34	10	10

Type of disturbance	Total number of studies	Increase	No change	Decrease
Stand-replacing	17	15	7	3
Nonstand-replacing*	19	15	10	9

*Note: totals do not always equal the sum of papers across each row because many studies found variable responses and assessed multiple metrics. *Papers focused on nonstand-replacing disturbances included three papers based on conceptual models, which predicted an increase (three papers), no change (three papers), or decreases (one paper) in streamflow. Table and caption adapted from Goeking and Tarboton (2020).*

2.2.3   Water resource management practices

Beyond land cover and climate, direct human interactions with streamflow, in particular through irrigation, dam regulation, and groundwater abstraction may have substantial influence on streamflow. Groundwater is a substantial contributor to streamflow (Barlow & Leake, 2012; Berghuijs et al., 2022; Bloomfield et al., 2021; Rodell et al., 2018). For instance, experimental work indicates that storm runoff often contains large amounts of water which does not originate with from current rainfall (Kirchner, 2003; Neal & Rosier, 1990; Sklash & Farvolden, 1979), and a substantial portion of streamflow is older than three months, having originated from stored sources (Jasechko et al., 2016). Recent research has shown that groundwater represents a substantial contribution even to flood flows in U.S. river (Berghuijs & Slater, 2023). Natural geological structures may help control groundwater-surface water interactions and in some catchments in the western U.S., connectivity of groundwater stores have been shown to result

in different streamflow sensitivity to climate change between catchments (Safeeq et al., 2013; Tague et al., 2008).

Water extraction from groundwater wells can have an influence on streamflow predominantly by lowering the water table and reducing the groundwater contribution towards baseflow, or in some cases, by drawing streamflow into the underlying groundwater system (Barlow & Leake, 2012). Once extracted, groundwater recharge time can range from days to millennia (Barlow & Leake, 2012; Winter, 2000), so these changes may have long lasting effects on streamflow. Further, climatological events like prolonged drought can deplete groundwater stores, and in the absence of replenishing wet years, lower runoff ratios and reduced streamflow elasticity may occur (Fowler et al., 2022; Hughes et al., 2012; Saft et al., 2015).

Crop irrigation represents a large consumer of water extracted from groundwater sources and around 70% of all global water extractions (Grafton et al., 2018). Irrigation can result in a variety of streamflow responses, depending on the management practices undertaken. For instance, flood irrigation diversions from surface water may reduce streamflow, but help to recharge aquifers, thus sustaining low flows throughout the year (Kendy & Bredehoeft, 2006). Crop irrigation might further increase actual evaporation, resulting in increased rainfall during water-limited periods (Kustu et al., 2011; Pei et al., 2016; Wang-Erlandsson et al., 2018; Wey et al., 2015), although heavily irrigated watersheds are more likely to have higher evaporation and lower annual streamflow overall (D. Wang & Hejazi, 2011). Substantial effort has been expended in recent years to increase irrigation efficiency through technology, reducing the water lost to evaporation, with the aim to increase available water for crops. However, increasing irrigation efficiency has often resulted in reductions in runoff and subsurface recharge, which may exacerbate hydrologic drought (Allan et al., 2020; Grafton et al., 2018).

2.3 The water-balance

Water balance equations express the relationships described by the hydrologic cycle quantitatively at different spatial and temporal scales and in some cases have been used to analytically derive streamflow sensitivity. A simplistic, but effective representation of the water balance describes streamflow as a function of precipitation, storage, and evaporation (Tang et al., 2020; D. Wang, 2012; Wu et al., 2018).

$$Q = P - (\Delta S + E) \qquad \qquad \textit{2.1}$$

Where Q is annual streamflow at the catchment level P is precipitation, ΔS is storage change, and E is evaporation.

2.3.1 Ponce and Shetty

The Ponce-Shetty water balance model is an example of a functional water balance model which focusses on how water is partitioned between the components of the hydrologic cycle, and then stored or released (Ponce & Shetty, 1995; Sivapalan et al., 2011). The revised Ponce-Shetty model (Sivapalan et al., 2011) frames a catchment's annual water balance is two-stage process in which precipitation is partitioned into "fast flow" and "wetting" components. Fast flow refers to water which runs off rapidly into a river channel, while wetting refers to stored water. The wetting component is then further subdivided into "vaporization", or water which is evaporated, and "baseflow", the component of streamflow derived from delayed sources (Gnann et al., 2019). The equations for the revised Ponce-Shetty model are as follows (Sivapalan et al., 2011):

For the two partitioning stages:

$$P = Q_f + W \qquad\qquad 2.2$$

$$W = Q_b + V \qquad\qquad 2.3$$

And for the whole catchment:

$$P = V + Q \qquad\qquad 2.4$$

$$Q = Q_f + Q_b \qquad\qquad 2.5$$

Where W is wetting, V is vaporisation, P is precipitation, fast flow is Q_f, and baseflow is Q_b.

In the original presentation, Ponce and Shetty (1995) provide a series of general equations for the estimation of the model parameters. The water balance equations all have the same structure, so that a quantity of X can be expressed at the sum of the components Y and Z. This means that Y can be expressed as a function of X and Z, the general solution for which is as follows:

$$Y = \frac{(X - \lambda Z_p)^2}{X + (1 - 2\lambda)Z_p} \qquad\qquad 2.6$$

In this context, Y could be substituted for quickflow (Q_f), X for precipitation, and Z for wetting potential, the maximum hypothetical wetting capacity of a catchment. Alternatively, Y could

be substituted for baseflow, Z for vaporisation potential, and X for wetting, where wetting is equal to:

$$W = P - Q_f \qquad\qquad 2.7$$

The wetting and vaporisation potentials (Z_p) can be calculated using the following equations:

For the case when $\lambda = 0$

$$Z_p = \frac{X(X-Y)}{Y} \qquad\qquad 2.8$$

and when $\lambda > 0$

$$Z_p = \frac{X + 1\left(\frac{1}{(2\lambda)}\right)\cdot\left((1-2\lambda)Y - \left((1-2\lambda)^2 Y^2 + 4\lambda(1-\lambda)XY\right)^{\frac{1}{2}}\right)}{\lambda} \qquad\qquad 2.9$$

Two initial abstraction coefficients, one for the base flow equation and one for the quick flow equation, are estimated through a calibration procedure, described in the original paper (Ponce & Shetty, 1995). Once identified, the product of the initial abstraction coefficients (λ) and the wetting and vaporization potentials results in a wetting and a precipitation threshold value, indicating the quantity of precipitation necessary to generate both fast and slow flow in any given catchment. These values have been shown to relate closely to streamflow elasticity to precipitation, particularly in dry climates, where the difference between total precipitation and the thresholds for flow generation is generally small (Harman et al., 2011). Harman et al. (2011) hypothesize that climate and the long-term water balance exert control on these thresholds, and speculate that shifts in climate, and landscape characteristics, using the example of urbanisation, may alter these thresholds, and consequently, streamflow-precipitation elasticity. However, the authors did not account for "carry-over", or interannual storage components, which may influence elasticity on an interannual basis.

The Ponce Shetty water balance approach for estimating elasticity is considered an analytical method of estimation because it is a closed-form mathematical expression which can be used to describe hydrological processes in a straightforward way according to known relationships. This differs from statistical-empirical analyses, like those which are relied upon in this thesis, in which relationships are derived directly from empirical data, rather than analytically. While potentially highly informative, analytical methods involve simplifying assumptions which may

limit their accuracy in complex real-world situations, and thus can be complemented by the application of large sample statistical approaches.

2.3.2 *The Budyko hypothesis*

The partitioning of precipitation into streamflow and evaporation has been further described as a function of the aridity index, the ratio of potential evaporation to precipitation, via the Budyko framework (Budyko, 1951, 1974). Fu's equation (B. P. Fu, 1981) provides a parametric form through which the mean annual water balance is a function of aridity and other factors.

$$F(\phi, \omega) = \frac{E}{P} = 1 - \frac{Q}{P} \qquad 2.10$$

where *F* is an analytical equation describing the evaporative fraction (E/P). E is often unknown so Q may be used instead (Andréassian & Perrin, 2012), thus *F* may be the runoff ratio (Q/P), ϕ is the aridity index, and ω is a parameter that accounts for other factors that influence the mean-annual partitioning of precipitation (e.g., climate seasonality, soils, vegetation, and topography) (Berghuijs et al., 2017). This parameter ω lacks clear physical meaning (E. Daly et al., 2019).

Averaged over many years, a majority of catchments tend to fall along the Budyko curve which shows the evaporative fraction (or runoff ratio) vs. the aridity index. The Budyko curve is defined by Equation 2.11 and depicted in Figure 2.2.

$$\frac{E}{P} = \left[\frac{Ep}{P} tanh \left(\frac{P}{Ep} \right) \left(1 - exp \left(-\frac{Ep}{P} \right) \right) \right]^{\frac{1}{2}} \qquad 2.11$$

Where E is actual evaporation, and Ep and P are potential evaporation and precipitation as previously.

While there are some exceptions and the framework does not explain all spatial variation , the relationship between the aridity index and evaporative fraction is strong and generally consistent across catchments, including in the U.S. Catchments theoretically fall below the supply limit (E/P =1) and the demand limit (E/Ep =1), but tend to approach these limits under very arid or very wet conditions. The aridity index is the dominant factor in determining spatial variations in how precipitation is partitioned between mean annual run-off and evaporation (Blöschl, 2013; Budyko, 1974).
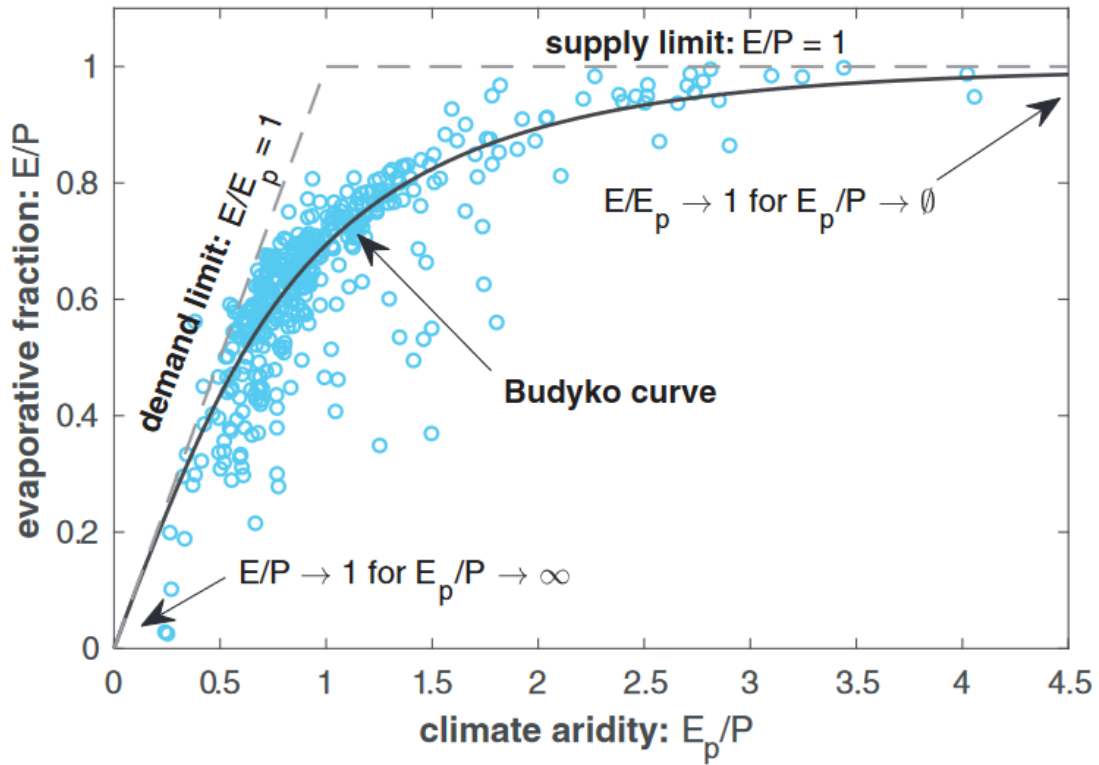
Figure 2.2 The Budyko curve

The x-axis contains the aridity index (Ep/P) and the y-axis the evaporative fraction (E/P), which often is approximated by one minus the runoff ratio (E/P = 1-Q/P) because storage changes are assumed to be negligible at multi-year timescales. 410 US MOPEX catchments are indicated by blue markers. Figure and caption taken from (Berghuijs et al., 2020).

Budyko-type equations have been used to analytically derive streamflow sensitivity to changes in aridity and other factors (Gudmundsson et al., 2016, 2017; Zhou et al., 2015). Often, studies quantify relative sensitivity – how responsive water availability is to each of the components of the Budyko framework as described in Equation 2.11, by assuming that E and Q follow the Budyko curve when aridity changes (Berghuijs et al., 2017; Berghuijs & Woods, 2016). Berghuijs et al. (2017) point out that this approach prohibits consideration of the effects of precipitation and provide technical improvements which allows for estimation of elasticity (proportional changes in total runoff) to separate changes in P, Ep, and ω as opposed to relative partitioning. Using this approach, they are able to analytically derive runoff sensitivity for individual grid cells at a global scale, determining the sensitivity of streamflow to changes in P, Ep, and other factors.

The Turc-Mezentsev or Turc-Pike (Equation 2.12) water balance formulas (Mezentsev, 1955; Pike, 1964; Turc, 1954) share many similarities with the Budkyo framework (Andréassian, Mander, et al., 2016; E. Daly et al., 2019). These have similarly been used to analytically estimate a theoretical streamflow elasticity (Andréassian, Coron, et al., 2016; Dooge, 1992; Sankarasubramanian et al., 2001).

$$E = \frac{P}{\sqrt{1+\left(\frac{P}{Ep}\right)^2}} \qquad\qquad 2.12$$

2.4 Hydrologic non-stationarity

2.4.1 Definitions and implications

As described, human activity and climate change, as well as natural variation in climatic processes can influence streamflow via changes in any other above-mentioned parameters of the water balance. Substantial changes in the streamflow time series can cause statistical non-stationarities, defined at the simplest level as a process by which the statistical properties, e.g. location, scale, or shape, of the streamflow distribution change over time (Slater et al., 2020). This is in contrast to a "stationary" streamflow time series, in which there is no significant change in any of these parameters (Figure 2.3). Approaches used for projection of hydrological hazards and for estimation of engineering-relevant metrics such as flood frequency, have generally assumed stationarity in streamflow timeseries, a paradigm which has been challenged in recent decades (Milly et al., 2008).

We are interested in changes in flooding, drought and water availability for somewhat obvious management reasons. Without a clear understanding of the probability of different magnitudes of streamflow, it is extremely difficult to plan for the associated risks. However, non-stationarity specifically, rather than just change, is particularly concerning for engineering practice. For instance, in the planning and design of conventional infrastructure that is subject to flood risk, such as dams, levees, or bridges, a "design flood" or flood of a certain magnitude, is used (François et al., 2019). Design floods are generally based on an assumption of stationarity, so any significant deviation on the statistical properties of hydrological distribution going into the future may have costly implications for design (François et al., 2019).
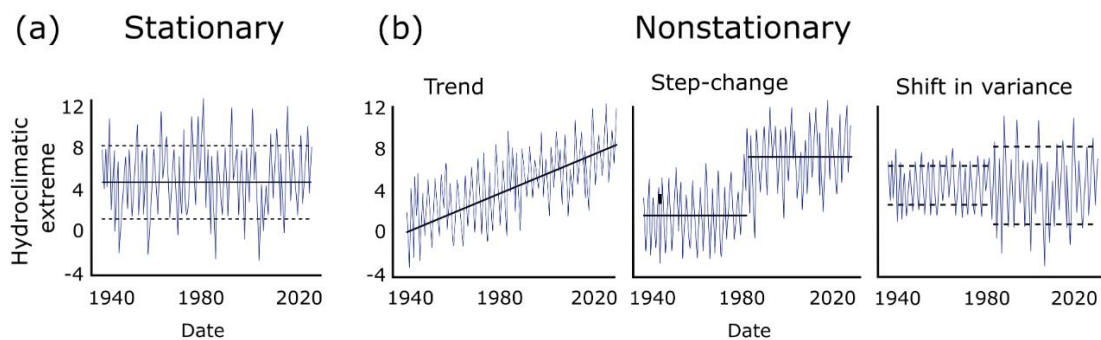
Figure 2.3 Examples of stationary and non-stationary time series
(a) a stationary time series with constant mean and variance; and (b) three non-stationary time series in the form of shift in mean (trend, step-change) and shift in variance. Solid and dashed black lines represent the mean and the variance of the time series, respectively. Figure from Slater et al. (2020).

Detection and attribution of the drivers of non-stationarity in hydrologic timeseries presents a number of challenges, but is essential in order to remove their influence from design problems (Bayazit, 2015). First, non-stationarity is usually detected using either a Mann Kendall trend test (Kendall, 1948; Mann, 1945) or the Pettit test for change point detection (Pettitt, 1979). These, and other tests, sometimes fail to detect known trends which are present in time series (Chappell & Tych, 2012; Mallakpour & Villarini, 2016). The tests struggle to adequately deal with autocorrelation, and require a priori assumptions regarding underlying physical processes (Serinaldi & Kilsby, 2015). Further, the timescale used to asses non-stationarity may strongly influence whether or not its presence is detected, because natural climate variability has its own structure which can be mistaken for trends at short timescales (Barros et al., 2014; Blöschl & Montanari, 2010; François et al., 2019; Koutsoyiannis, 2006). In other words, a stationary

streamflow time series will always exhibit natural variation in streamflow, but is expected to fall within an envelope of typical behaviour, that is behaviour which exists within the historical record. This includes anomalously dry or wet periods, and the question of how long the record should be to assess stationarity, hasn't been adequately addressed (François et al., 2019). Others have even argued that it is impossible to determine whether a trend is truly a trend, or the product of long term persistence in the hydrological record (Matalas, 2012).

In light of these concerns, it is possible that in the case of hydrologic projection, when non-stationarity exists but the underlying model structure or drivers of change are unknown, it may be preferable to use stationary models for the design and management of extremes. Stationary models may help avoid increasing uncertainty (Serinaldi & Kilsby, 2015; Slater et al., 2020). Further, stationary models may be preferable to non-stationary ones when the change in time cannot be reliably predicted (Bayazit, 2015). Thus, understanding non-stationarity when the drivers are physically plausible, explainable, or known (Clarke, 2013) can help with interpretation of hydrologic change, improve projection in cases where future flow distributions deviate from historical ones, and reduce losses due to hydrological risk. There is some evidence that aggregation of data to larger spatial scales can reduce the noise of interannual climate variability, and improve detection of trends (Fischer & Knutti, 2014; François et al., 2019). While substantial debate exists regarding the appropriate methodologies for detecting and dealing with non-stationarity, it seems clear that it remains a useful concept.

2.4.2   Existing trends

In the past two decades countless studies have attempted to assess the presence of trends in hydrologic timeseries over periods ranging from the mid-20th century, often to differing results. Trends in streamflow have often been assessed at the regional scale. On the whole, the literature suggests that while fluvial floods, and especially large floods, are not necessarily becoming more extreme or more frequent at a national scale, they may be in some regions and catchments (Douville et al., 2021). Further, significant trends in low and average streamflow are present in many locations in the U.S. at the annual timescale.

For example, Archfield et al. (2016) found some statistically significant regional scale trends in the frequency, magnitude, duration, and volume of floods across the United States. However, these trends lacked substantial geographic cohesion, e.g. the direction of the trends was not consistent in space, and many regions showed no significant trends. These findings suggested that trends in flooding may be better studied at the catchment scale. Again, at the regional scale,

Collins et al. (2023) found evidence of increasing frequency of the occurrence of large floods in only two regions in the U.S. and concluded that on average, large floods were not becoming more common in the 1966 to 2015 study period. Slater and Villarini, (2016), on the other hand, found clear regionally consistent trends in flood risk as defined according to impact-based categories. They found increasing trends in Northern Midwest and in the Great Lakes region, and decreasing trends in the South and Pacific Southwest, although using a shorter time series (1985-2015). Douglas et al. (2000) found statistically significant regional-scale trends in low flows in the Midwest, but no statistically significant trends in flood flows across the country. Rice et al. (2015) found that average streamflow across the United States was generally increasing but that high flows were becoming less extreme. Patterns at the regional scale varied substantially, with significant increasing trends in the annual mean streamflow across the Northeastern portion of the country and significant decreasing trends spread more across the Western and Central U.S. Sadri et al. (2016) in an analysis of low flow magnitude found a general pattern of increasing low flows in the Northeast and decreasing low flows in the Southeast of the U.S. Approximate regions of the United States are outlined in Figure 2.4.
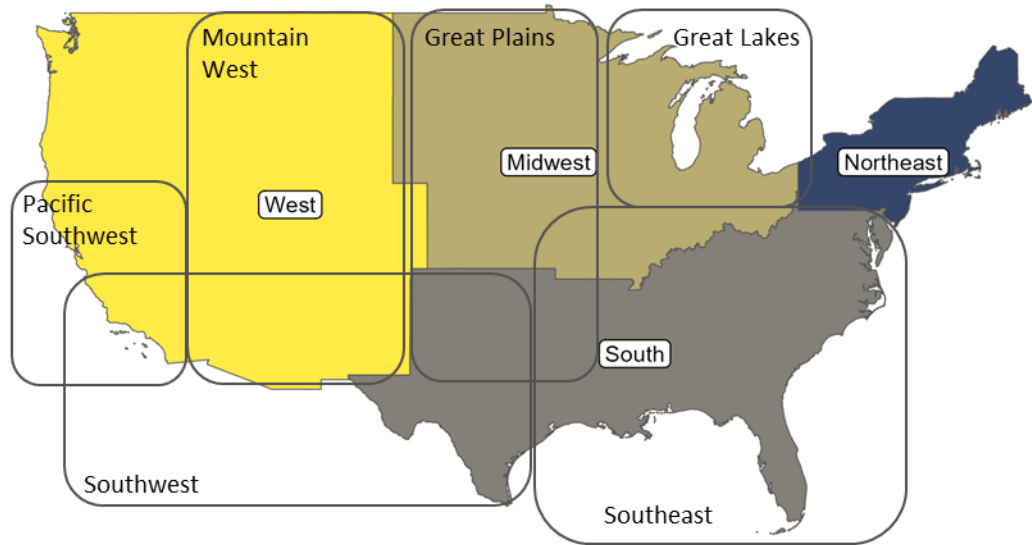
Figure 2.4 Generic map of the regions of the United States

Approximate locations of the regions of the United States mentioned in the literature review.
White labels indicate U.S. Census regions. Black boxes correspond to black labels and
indicate rough approximations of the outlines of further subregions mentioned throughout
this literature review.

3   Methodology

3.1   Hydrologic signatures

3.1.1   Streamflow

Streamflow timeseries can be represented as a statistical distribution, and are often visualized via a hydrograph, a diagram which shows discharge rate over time. This can be quantified and examined using a range of hydrologic signatures (Figure 3.1), or quantitative metrics that describe the statistical or dynamical properties of streamflow (McMillan, 2020). Of particular interest have been: streamflow magnitude, the actual volume of water passing by a specific point in a river channel, typically measured in cubic metres per second; timing, the specific time of year in which certain flow events, such as annual maximum, occur; frequency, the regularity of occurrence of four events of particular magnitudes; and duration, the length of time for which a specific flow condition occurs (McMillan, 2021; Richter et al., 1996). Other useful metrics exist which describe, for instance, the shape of different segments of the flow

duration curve (FDC), the cumulative distribution function of flow that shows the percent of time that flow values are exceeded (Searcy, 1959), or the aridity index, defined as the ratio of potential evaporation to precipitation.

These signatures can be linked to hydrologic processes or used as proxies to estimate them (McMillan, 2020). For instance, the slope of the low end of the FDC has been shown to be a useful representation of flow from deep groundwater (Farmer et al., 2003), the slope of the midsection of the curve can be used to capture low behaviour e.g. during the snowmelt season (Kelleher et al., 2015), and the slope of the high end of the FDC can be used to represent "flashiness", where a high slope may indicate a lack of storage (McMillan, 2020; Safeeq & Hunsaker, 2016). Elasticity can also be considered a streamflow signature, which links to a range of processes relating to the sensitivity of a catchment to hydrologic response.
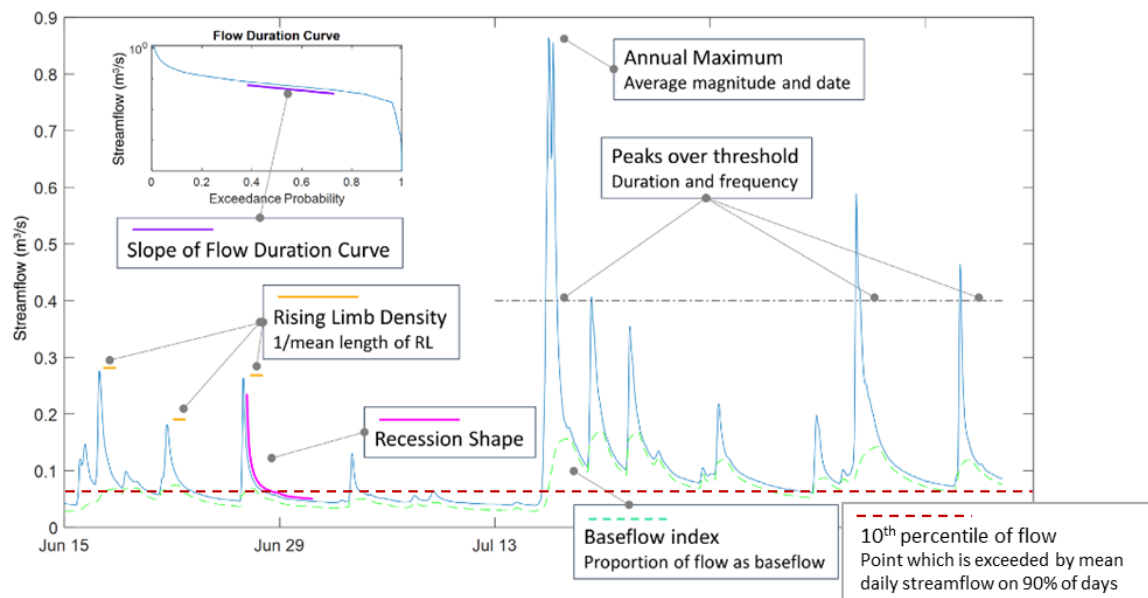
Figure 3.1 Example hydrograph demonstrating different hydrologic signatures.
Figure adapted from Mcmillan, 2021 to include the 10th percentile of flow as an example of streamflow percentiles.

In this thesis, I focus primarily on the sensitivity of streamflow magnitude across the different segments of the flow distribution to variation in environmental inputs. Streamflow magnitude varies naturally over different time periods. For instance, a snow dominated river may experience diurnal peaks and troughs during snow melt season, in line with temperature (Lundquist & Cayan, 2002) as well as annual high and low flow seasons in line with snow melt occurrence and rainfall.

Different segments of the flow distribution provide different information about the physical, ecological, and management properties of a catchment. Low flows are important indicators for ecological health (Poff & Zimmerman, 2010) and commerce. They generally occur in dry periods and are seasonally recurrent (Smakhtin, 2001). While they are often correlated with drought events, they are not synonymous with hydrological drought, as low flows are regularly occurring and droughts are anomalous events normally associated with implications for resource availability (Smakhtin, 2001). Low flow is somewhat analogous to baseflow, although extreme low flows below the baseflow separation threshold are not uncommon. This means that baseflow does not necessarily correspond to minima, and instead represents an average flow during drier periods. Low flows are predominantly fed by discharge from surface water and groundwater sources (Brutsaert, 2008; Dudley et al., 2019; Smakhtin, 2001). A number of specific hydrologic signatures exist for quantifying low flow magnitude, including the widespread low flow index "7Q10", which represents the 7-day average low flow with a 10-year return period (Fouad et al., 2016; Smakhtin, 2001), as well as exceedance probabilities, the discharge amount exceeded a given percentage of the time (Pyrce, 2004).

High streamflow magnitudes occur during wet periods and are somewhat analogous with flooding. This portion of the annual flow distribution is of interest due to its association with flood hazards and their associated financial risks. High flows are typically driven by precipitation events in combination with antecedent moisture conditions (Bennett et al., 2018; Berghuijs et al., 2016; Farquharson et al., 1992; Guo et al., 2014; Neri et al., 2019; Slater & Villarini, 2017). Some recent evidence indicates that inputs from storage sources represent an important contribution to high flows, especially when predicting long-term trends in these values (Berghuijs & Slater, 2023). Similar to low flow metrics, high flows can be quantified using a number of different approaches including peaks over threshold techniques, flood recurrence intervals, mean annual floods, etc.

While streamflow magnitude can be quantified using a variety of signatures, one simple and common approach which allows consideration of multiple segments of the annual hydrograph using the same methodology is to calculate percentiles or quantiles of flow in a given time period (Figure 3.1). Streamflow percentiles are very similar to exceedance probabilities, a straightforward representation of the percentage of time in which the flow of a given magnitude is achieved in a river. Since percentiles offer a consistent methodology across the distribution, this is the approach used throughout this thesis. Depending on the country of origin of the research there is some debate regarding the syntax of flow percentiles, where Q90 (90th

percentile) may refer to flow exceeded 90% of the time, or 10% of the time. Throughout this work, I use the second approach e.g. the 90th percentile of streamflow is flow exceeded only 10% of the time, corresponding to high flows.

The magnitude of streamflow itself is determined by a number of factors. Climatic variables such as precipitation and antecedent wetness are typically identified as the primary drivers of hydrologic variability (Berghuijs et al., 2016; Neri et al., 2019; Slater & Villarini, 2017). The literature generally shows that flood magnitude is dependent on a combination of precipitation events and the antecedent wetness of soil and groundwater (Bennett et al., 2018; Berghuijs & Slater, 2023; Ivancic & Shaw, 2015; Slater & Villarini, 2017), and that low and median flows are better predicted by antecedent moisture availability (Slater & Villarini, 2017). Physiographic characteristics, those relating to the underlying geology and geomorphology of a catchment, may also play a role in determining flow magnitude. For instance, channel width, catchment slope, as well as catchment roughness and vegetation, may all help determine streamflow magnitude. The roles of these different components of the water balance are discussed in the next section.

3.1.2 Storage

There are several approaches for quantifying catchment water storage which include direct measurements, hydrologic tracer studies, and modelling exercises (Kalbus et al., 2006; McNamara et al., 2011). The storage-discharge relationship could also be calculated using an indirect water balance approach with average baseflow (McMillan, 2020; McNamara et al., 2011), or by examining the streamflow recession curve following rain events (Krakauer & Temimi, 2011). Point estimates of hydrologic storage are common, but spatially distributed estimates which more completely capture hydrologic storage are less widely available.

Satellite remote sensing, particularly NASA's Gravity Recovery and Climate Experiment (GRACE) mission, offers another pathway for quantifying changes in total water storage (TWS), the sum of groundwater, soil moisture, surface waters, snow and ice (Girotto & Rodell, 2019; Rodell et al., 2018; Rodell & Famiglietti, 2002) which may address this concern to some extent. GRACE measures anomalies in Earth's gravity field, allowing estimation of changes in storage, including groundwater, which is especially difficult to assess (Rodell et al., 2018). However, GRACE data has coarse spatial resolution, lacks the ability to partition component mass changes, and has a limited time series length (Rodell et al., 2018). For these reasons, it is not considered in depth in this thesis, but does provide important context on total water storage

changes throughout the literature. It is worth noting that a machine learning-based reconstruction of a GRACE-based TWS timeseries back to 1940 was published in the final weeks of the writing of this thesis (Yin et al., 2023). This and similar work may prove useful in future analyses. Figure 3.2 shows trends in total water storage between 2002 and 2016 derived from the GRACE satellite mission (Rodell et al., 2018).
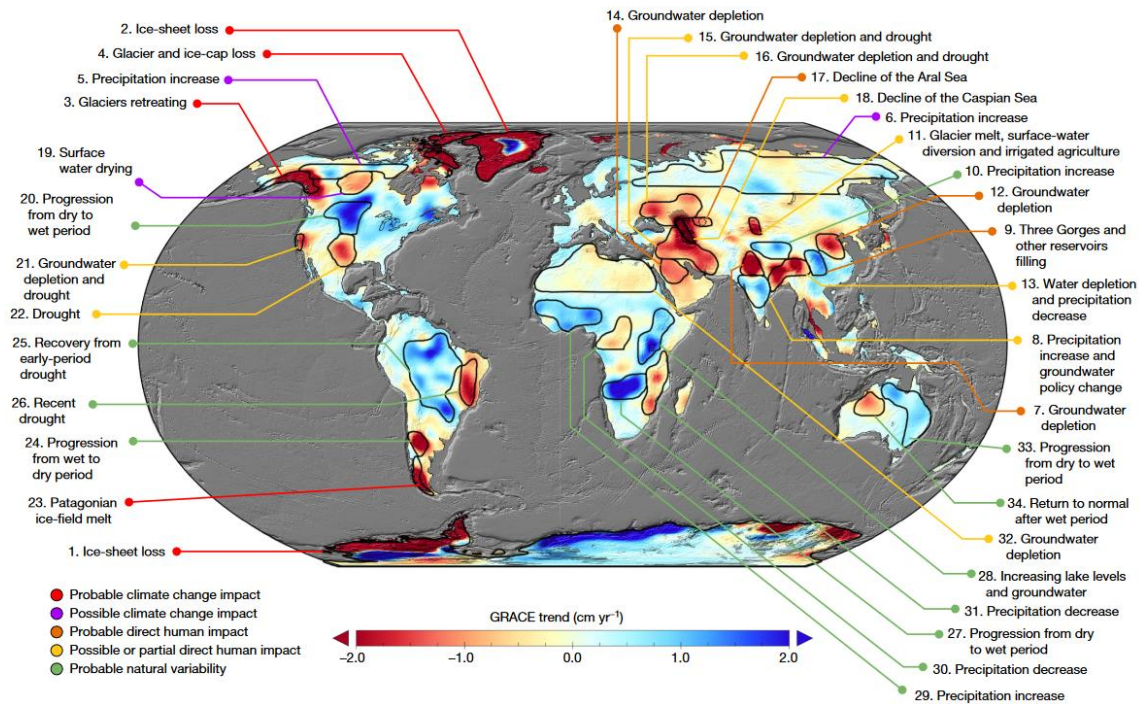
Figure 3.2 Annotated map of TWS trends

Trends in TWS (in centimetres per year) obtained on the basis of GRACE observations from April 2002 to March 2016. The cause of the trend in each outlined study region is briefly explained and colour-coded by category. The trend map was smoothed with a 150-km-radius Gaussian filter for the purpose of visualisation; however, all calculations were performed at the native 3° resolution of the data product. Figure and caption taken from Rodell et al., 2018.

Catchment water storage can be difficult to quantify and is thus often excluded from analyses of streamflow. However, much recent hydrological research has incorporated storage by using a proxy. These proxies include the baseflow index (BFI) (Berghuijs & Slater, 2023; Y. Zhang et al., 2023), the residuals of the water balance closure (Sayama et al., 2011; D. Wang & Alimohammadi, 2012), or a similar analytically derived signature such as the baseflow fraction (Gnann et al., 2019). While it is important to note that these do not directly measure storage or the storage-discharge relationship, these do indirectly provide information about storage processes and structure (McMillan, 2020; Pfister et al., 2017) and may be used as an

indicator of the importance of groundwater influence (Kalbus et al., 2006; Wrede et al., 2015).

As described previously, streamflow is broadly comprised of fast flow and baseflow components. Fast flow, including direct surface runoff and rapidly draining subsurface flow, is the portion of streamflow which results directly from rainfall events. Baseflow is the flow derived from groundwater and other sources of storage, resulting in a delayed input into rivers in the absence of rainfall (Beck et al., 2013; Gnann et al., 2019; Smakhtin, 2001). Stored sources of water, like groundwater often also contribute to streamflow even during rainy periods (Wittenberg, 1999). Thus, baseflow is the portion of streamflow which is sustained during dry parts of the year and a provides an important and influential contribution to flow.

In the absence of data derived by hydrometric or tracer approaches (Smakhtin, 2001; Stoelzle et al., 2020), several data-based approaches exist for estimating the baseflow index, the most common being two-component hydrograph separation methods. An example of this, the approach proposed by Lyne and Hollick (Lyne & Hollick, 1979) uses a recursive digital filter for baseflow separation. The similarly popular UK Institute of Hydrology approach performs separations based on progressively identified streamflow minima (Gustard et al., 1992).

The hydrograph separation class of approaches has been criticized due to the ambiguity of the results produced, especially insofar as they fail to provide differentiable information regarding the sources of water contributions (Klaus & McDonnell, 2013; Stoelzle et al., 2020). In other words, the typical BFI mixes different delayed flow sources into one parameter (Hellwig & Stahl, 2018; Parry et al., 2016; Partington et al., 2012; Stoelzle et al., 2020). Further, it is broadly acknowledged that these numerical approaches are not closely related to the underlying physical processes, although they can provide useful information for certain applications (Ladson et al., 2013).

Recent research has indicated that the estimation of a "delayed flow index" (DFI) in addition to a traditional BFI can provide more robust information regarding the sources of discharge from storage (Gnann et al., 2021; Stoelzle et al., 2020). The DFI has been shown to more closely relate to geology (Gnann et al., 2021) and may provide a more reliable estimate of the ability of a catchment to sustain streamflow during dry periods than does the traditional BFI estimation (Stoelzle et al., 2020). The DFI may be considered to represent the more stable multiannual baseflow from seasonal variations, while BFI separates event driven flow from interevent flow, although they different components do not necessarily relate directly to any

particular source (Gnann et al., 2021). The DFI can be estimated by generating multiple baseflow components over different time periods, rather than the standard approach of five days (Stoelzle et al., 2020).

3.2 Elasticity

Elasticity is typically calculated for climatological variables such as precipitation or temperature, although the land cover questions can be addressed through similar approaches. Streamflow elasticity to precipitation is usually estimated from the central summary (e.g. median or mean) of streamflow at the annual timescale (Berghuijs et al., 2017; Chiew, 2006; Chiew et al., 2014; Milly et al., 2018b; Sankarasubramanian et al., 2001; Tang et al., 2020; Tsai, 2017) and can be calculated using a number of methods, including multivariate regression.

Precipitation elasticity of streamflow was originally defined by Schaake (1990) as:

$$\varepsilon_p(P, Q) = \frac{dQ/Q}{dP/P} = \frac{dQ}{dP} \cdot \frac{P}{Q} \qquad 3.1$$

Where precipitation elasticity is ε_p is a function of the normalised ratio of change in streamflow *(dQ)* to change in precipitation *(dP)*. Estimation of elasticity using this approach can be difficult, in part because the ratio of $\frac{dQ}{dP}$ has typically been estimated from a hydrologic model (Sankarasubramanian et al., 2001). Sankarasubramanian et al. (2001) developed a more robust estimator presenting an alternative approach in which elasticity was defined at the mean of the climate variable(s), such that it represents the effect on streamflow of a 1% deviation from the long-term mean of precipitation or (evaporation). Variations of this approach, defined by Equations 3.2 and 3.3 below, have become the dominant expression in subsequent years.

$$\varepsilon_p(\mu P, \mu Q) = \frac{dQ}{dP}\bigg|_{P=\mu P} \frac{\mu P}{\mu Q} \qquad 3.2$$

Where *μP* and *μQ* represent mean precipitation and mean streamflow respectively. The median of these values is then used as a nonparametric elasticity estimator:

$$e_P^1 = median\left(\frac{Q_t - \bar{Q}}{P_t - \bar{P}}\right)\frac{\bar{Q}}{\bar{P}} \qquad 3.3$$

Where $\bar{Q}$ and $\bar{P}$ are the long term means of streamflow and precipitation respectively. An analogous estimator exists for potential evaporation.

Sankarasubramanian et al. (2001) also propose a bivariate linear estimator which incorporates potential evaporation and Potter et al. (2011) suggested computing elasticity as a multiple linear regression between annual transformed streamflow values to annual precipitation and temperature anomalies. Many recent studies have employed similar approaches. For instance, Andréassian et al. (2016) compared five different regression approaches to a reference estimator and found that a bivariate generalized least square regression approach which incorporated both precipitation and potential evaporation was able to account for covariation in the climate variables, minimizing bias and RMSE. Others have created more complex forms, estimating elasticity using multivariate regression models which include storage parameters in addition to climatological ones, in an attempt to more fully integrate the water balance (Konapala & Mishra, 2016; Tang et al., 2020; Y. Zhang et al., 2023).

3.2.1 Streamflow sensitivity

Elasticity is, broadly speaking, a method for estimating the sensitivity of streamflow to changes in a related variable. A variety of other frameworks and approaches have been used to study these relationships including hydrological models, conceptual models, paired catchment studies, and statistical and machine learning approaches. A very brief introduction to these themes is provided here.

Andréassian et al. (2016) argue that the majority of studies which estimate elasticity are theoretical in nature. They use the Turc-Mezentsev water balance equations (Mezentsev, 1955; Turc, 1954) as a reference estimation approach to analytically derive elasticity, as have many previous studies (Arora, 2002; Donohue et al., 2011; Sankarasubramanian et al., 2001). Other approaches, such as the previously described Budyko-based and the Ponce and Shetty water balance approaches perform a similar function – quantifying the sensitivity or responsiveness of streamflow (or precipitation or temperature) to changes in some other variable using analytical means.

Attribution studies are similar to studies which estimate sensitivity and aim to discern the most likely causes of a change after it is detected in the observed record (De Niel & Willems, 2019). They relate trends to various drivers which may have caused those shifts in hydrological behaviour. However, the term is often used more broadly. Attribution studies may be carried out using experimental or quasi-experimental approaches when assessing changes in a limited number of catchments, but the primary aim is to discern the principle causes of change.

Before After Control Impact (BACI) approaches involve sampling both before and after a known impact occurs, when no spatial control group is available and then comparing changes across similar individuals. One frequently used approach are paired catchment studies, in which two or more catchments that are very similar in their characteristics—i.e. size, soil type, climatology, vegetation, are compared after half of each pair has experienced some change which the other has not. This approach has been used extensively (Bart et al., 2016; Boggs et al., 2016; Brown et al., 2005; Loon et al., 2019; Prosdocimi et al., 2015) and possibly represents the most robust method for detecting and attributing changes at a small scale (Salavati et al., 2016). Reference catchments, such as the Plynlimon catchments in the UK have been exceptionally useful in these types of experiments (Archer, 2007). However, this approach is highly limited by the existence of two similar, neighbouring catchments which differ only in their recent interventions.

Modelling approaches have been widely used to examine the possible influences of land cover and climatological changes on streamflow, both for the purposes of attribution and estimating sensitivity. In contrast to relying on the observed record, this allows for the generation of scenarios which may not exist in the observed record, and which can test effects based on known relationships. For instance, Buechel et al. (2022; 2023) simulated large scale afforestation in the United Kingdom using a land surface model, demonstrating a range of hydrologic responses. Others, for example, Salavati et al. (2016), used the pre-urbanisation period in the same catchment to simulate and compare unaltered flow to post-urbanisation streamflow using a physically based model. Others have incorporated hydrological models such as the abcd model (Sankarasubramanian et al., 2001), the SIMHYD and AWBM models (Chiew, 2006), as well as the Soil Water Assessment Tool (SWAT) (e.g. Schilling et al., 2014), the Variable Infiltration Capacity (VIC) model (Frans et al., 2013; Hidalgo et al., 2009; Matheussen et al., 2000; Nijssen et al., 2001) to simulate unaffected conditions in riverine systems and then compare them with observed conditions which can then be compared with observed streamflow. Simulations may also be carried out using statistical tools and observed data. Simulation-based approaches such as these are limited in that their results are often the product of model assumptions (Blöschl et al., 2007; Salavati et al., 2016)

3.2.2 Elasticity estimation approaches

In this thesis, I use a range of regression model types which consist of generalized linear models formulated within the Generalized Additive Models for Location, Scale and Shape (GAMLSS) framework (D. Stasinopoulos et al., 2017; D. M. Stasinopoulos & Rigby, 2007), and fixed-

effects panel, or longitudinal, regression models. Each of the general model types that I use are described below and are discussed in the context of their applications.

3.2.2.1   Single-site approach

In two chapters of my research I apply what I refer to as a single site regression approach to estimate streamflow sensitivity to land cover changes and then climate. This approach involves regression models independently fit to each catchment, where the regression slope (β coefficients) represents the relationships between the variables of interest. This is an exceptionally common approach to estimating elasticity, as a large number of papers have similarly applied single site regression models in this fashion.

I use the GAMLSS package in R, which represents a framework for distributional regression modelling where the parameters of the assumed distribution can be flexibly modelled as functions of the explanatory variables (Rigby & Stasinopoulos, 2017; D. M. Stasinopoulos & Rigby, 2007), to create models in my first research chapter (Ch. 4). The GAMLSS framework allows any parametric distribution for the response variable. Models applied within the GAMLSS framework have been used in hydrology to model non-stationary processes. Some applications have included flood frequency analysis (López & Francés, 2013; Villarini, Serinaldi, et al., 2009; Villarini, Smith, et al., 2009), drought indices (Shao et al., 2022; Sun et al., 2020), and to assess the impacts of urbanisation on streamflow (Han et al., 2022).

The growing popularity of this approach in hydrology precipitated its application in the early stages of my research, however following testing of more complex parameterization, a simpler formulation seemed best suited to the problems I address. Estimation within the GAMLSS framework is based on the maximum likelihood principle, and if only the location parameter is modelled, regression models may resemble more common generalized linear models (dependent on distributional assumptions), as they do in my work. In the second research chapter (Ch. 5) I instead fit simple ordinary least square regression log-log linear regression models to the data, as this saves on computation time and the results were nearly identical. Discussion of these choices is included in the relevant research chapters.

3.2.2.2   Multi-site approach

The primary concerns with single site regression models are that because hydrological data is noisy, it is difficult to separate out the effects of different drivers from one another, and thus results are not robust to spatial variations. Panel regression models, which are borrowed from

econometrics and are relatively new in hydrology, offer some mechanisms to address these issues. These are regression models which are applied to a two-dimensional dataset, typically one in which data is recorded across space and time simultaneously. These models can distinguish between within-variability (temporal) and between-variability (spatial) in the same model, allowing for the control of confounding relationships (Steinschneider et al., 2013). Because panel regression models merge characteristics of both cross sectional (across space) and single-site time series models, they are better able to leverage heterogeneity between catchments, and isolate specific relationships.

In this thesis I use fixed effects panel regression models. In these models we assume that there are omitted variables from our regression which are correlated with regressors that are included in the model (Steinschneider et al., 2013). Assigning fixed effects at the catchment level controls for these omitted variables. An alternative assumption would be that omitted variable are not correlated with the predictors or residuals in the model, and are therefore considered through random variables. Fixed effects panel regression models address the majority of omitted variable bias, by requiring that confounding variables either be directly measured or be invariant along at least one dimension of the data, for instance, time (Nichols, 2007).

Panel regression allows for improved statistical robustness in the estimation of relationships between environmental changes and streamflow. These methods are starting to develop traction in hydrology, largely having been applied for the estimation of land cover changes on flow (Blum et al., 2020; Steinschneider et al., 2013; W. Yang et al., 2021). Bassiouni et al (2016) used random effects, fixed effects, and pool panel regression models to estimate low flow elasticity to precipitation in Hawaii. They showed that as compared to traditional regression in which space was substituted for time (Singh et al., 2011), panel regression models were able to reproduce behaviour while also reducing standard errors, even when unobserved heterogeneity is significant and substantial multicollinearity exists. Blum et al. (2020) used fixed effects panel regression models to assess the influence of impervious surface area on floods. Figure 3.3 shows a causal diagram (concept explained in section 3.2.2) and the panel regression model used in their study. This diagram outlines how each of the included variables accounts for potential bias in the system.
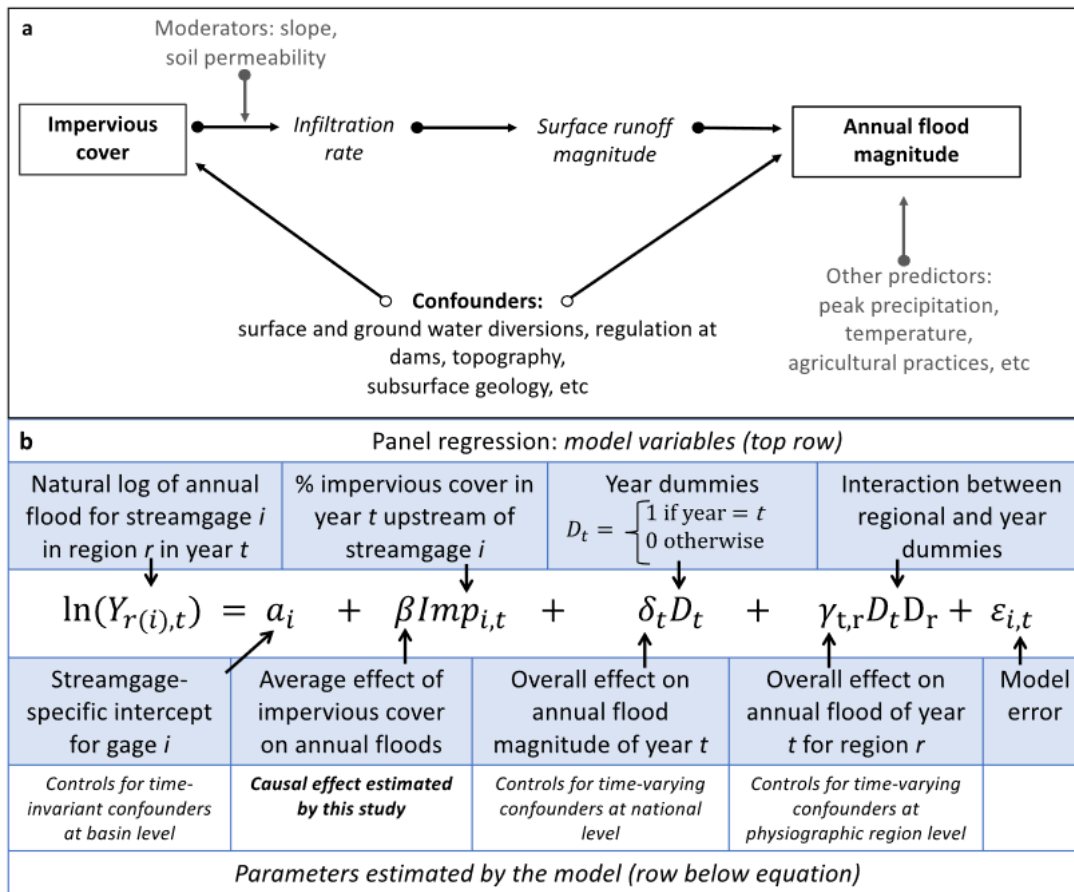
Figure 3.3 Causal diagram and example panel regression model.

(a) Causal graph illustrating the relationship between impervious cover and changes in annual flood magnitude. Arrows illustrate causal relationships, not physical pathways. (b) Panel data regression estimator used to estimate the average causal effect of impervious cover on annual flood magnitude. Figure and caption taken from Blum et al. (2020). See original source for details of the regions described.

The reader is directed to Steinschneider et al (2013) for a detailed introduction to panel models in hydrology. It is worth noting also that mixed effects models, which are not explored in detail here, are statistically similar to panel regression and provide many of the same advantages. These models are better suited for hierarchical data and combine random and fixed effects in the same models giving them the capability to control for within catchment variation and allowing for statistical significance testing between groups.

3.3   Statistical and causal inference in hydrology

In this thesis, I am interested in discerning causal relationships using empirical data and simple statistical tools. Perhaps the most common refrain in science is that *correlation does not equal*

causation. This idea refers to the fact that a statistically significant relationship does not necessarily indicate that one process has *caused* another. The broader field of causal inference attempts to address this disparity. The specific methodological approaches used in this thesis are described in the individual research chapters, however the concepts are introduced briefly here.

The way in which most of us use statistical tools typically falls under the umbrella of *statistical inference*. Within this framework, the parameters of a distribution can be assessed by drawing samples from that distribution. This involves the use of probabilistic tools like confidence intervals, p-values and hypothesis testing to make inferences from the data (Otsuka, 2023; Pearl, 2009). When making statistical inferences, we generally avoid claiming causality and instead refer to "associations" between variables. *Causal inference* approaches may employ the same or very similar tools, but with the objective of inferring relationships which persist across changing conditions and which cannot be clearly defined by distribution functions alone. In other words, causal inference aims to identify relationships which clearly represent our underlying assumptions about the causal pathways within a system (Otsuka, 2023; Pearl, 2009). Correlation does not equal causation nor is correlation directly necessary to infer causation (Sugihara et al., 2012). Causal inference is fundamentally concerned with whether or not particular mechanisms are explicitly linked in an ordered fashion, rather than just consistently associated with one another (Huntington-Klein, 2021). For instance, we may want to know whether carbon emissions *cause* warming temperatures, or if they just happen to be occurring at the same time.

Ideally, causality is inferred through experimental research because these approaches allow for carefully designed counterfactuals and control groups. A number of experimental catchments have been over the years in attempts to quantify the impacts of changes on hydrology (Bosch & Hewlett, 1982; Moore & Wondzell, 2005). For example, the famous Plynlimon catchment in Wales, which has been used as an experimental catchment to observe the influence of forest cover change for over 40 years, about which hundreds of publications have been written (Robinson et al., 2013).

However, experimental research is often infeasible or impractical because of ethics, plausibility, cost, or accuracy, and small catchment scale experiments limit capacity to develop robust, generalizable hypotheses. Thus, a number of different methods exist for inferring causal relationships from observed data. Applications in hydrological science, as outlined by Ombadi

et al. (2020) include frameworks such as Granger causality, which states that variable X has a causal effect on variable Y if variable X provides statistically significant information about future values of variable Y (Granger, 1969), transfer entropy (Schreiber, 2000), convergent cross mapping (Sugihara et al., 2012), and graph based causal methods such as causal diagrams, which define casual relationships (Pearl, 1995; Runge et al., 2019). Quasi-experimental research designs, which include regression techniques (Nichols, 2007; Shadish et al., 2002), have also been used in combination with causal diagrams to discern causality from observed data (Blum et al., 2020; Ferraro et al., 2019; Williamson et al., 2014).

Some of the fundamental concerns in causal inference include how to deal with biases which arise from *confounding* and *omitted variables,* and how to express causal relationships, since these cannot be adequately represented through associations as in frequentist statistics. Statistical tools, like those used in this thesis, can be used to help address some of these biases, while the causal inference methods mentioned above may help address others.

3.3.1   Causal diagrams

Finally, causal diagrams, or directed acyclic graphs (DAGs), can be useful for defining difficult or untestable assumptions which underlie statistical models by outlining causal pathways, as well as confounding, moderating, and mitigating variables present within a natural system. They are popular in econometric and epidemiological research and have recently been employed in hydrology and earth sciences (e.g. Blum et al., 2020). These graphical models can then be used to guide variable selection by integrating subject-matter information with statistical models in order to facilitate causal inference from observed data They also make the underlying assumptions explicit, so that the interpretation of results is clearer to the reader.

Confounding variables are those which may influence both independent and the dependent variables simultaneously, leading to confusion regarding which processes are the actual causal mechanism. When using causal diagrams to help design regression models for causal interpretation, confounding variables are typically "controlled for" in order to limit their influence on the outcome and allow for the isolation of the direct effect. Moderating variables, on the other hand, are variables which indicate under what conditions the relationship between the variables of interest is stronger or weaker (Huntington-Klein, 2021).

To illustrate this point, we can use a simple example, for instance to model the relationship between playing football and getting frostbite, the D.A.G. outlining causal assumptions might look something like Figure 3.4.
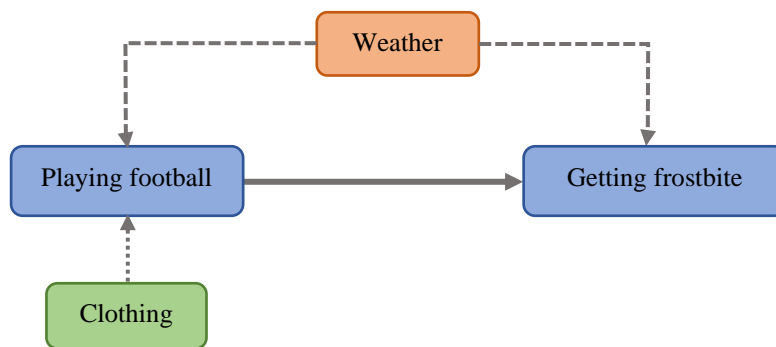
Figure 3.4 Example directed acyclic graph or causal diagram.

Demonstrates the relationship between playing football and getting frostbite. Blue is the direct causal relationship being measured, green indicates a moderating variable, and orange indicates a confounding one.

Where playing football is the independent variable, getting frostbite is the dependent variable, weather is a confounding variable because the weather will affect both playing football and getting frostbite, and finally, clothing is a moderating variable because the clothing worn will influence the extent to which a person will experience negative effects of the cold. Obviously, the model excludes some important factors, for instance, amount of time exposed to cold, but the assumptions of the model are explicit. The same approach can be used when creating empirical models with the aim of inferring causality within hydrological research, and such causal diagrams can be used to inform and make explicit the assumptions behind specific modelling choices.

3.4   Comment on data availability

The stream gages used in the analyses presented in the following chapters were selected from the same dataset and similar selection procedures were applied in each case. For instance, I required a minimum of 95% complete daily streamflow data in each catchment and at least 20 (in chapter 4) or 30 (in chapters 5 and 6) years of consecutive data since at least the 1980s. The specific requirements for each analysis are detailed in those chapters respectively, and slightly different sample sizes were included in each case depending on the data requirements.

Relatively short time series were used in part because of the limited temporal availability observed land cover data from satellite imagery, which is generally only available since the

1980s, a limitation which is relevant to the work carried out in Chapter 4. The length of time series included in this work is most likely to affect the single-site regression models, as the panel regression approach is more robust against problem, as spatial variation accommodates a wider distribution of hydrological behaviour and increases statistical robustness. Regardless, longer timeseries lengths would likely be beneficial to the analyses, and may result in conclusions which are more representative of longer-term hydrological behaviour. In chapters 5 and 6, timeseries lengths were limited because older climatological data (prior to 1981) from the PRISM dataset (Hart & Bell, 2015b) were based on less extensive historical observations, and thus were thought to be less reliable.

Overall, the methodologies employed in this thesis require data from relatively large numbers of catchments, and thus may not be applicable in parts of the world where data is less available. Further, it is worth noting that while the dataset employed in this study covers a fairly wide range of climatological and geophysical zones, stream gages which meet the selection criteria are substantially less abundant throughout much of the Great Plains and Southwest regions (Figure 2.4). This factor represents a limitation in terms of how representative the results can be of hydrologic behaviour over e.g. drier regions, but is not more restrictive than previous work on the subject. Each of these points is noted with greater specificity in throughout the research chapters.

4 Statistical attribution of the influence of urban and tree cover change on streamflow: a comparison of large sample statistical approaches

4.1 Abstract

The strengths and weaknesses of different statistical methodologies for attributing changes in streamflow to land cover are still poorly understood. We examine the relationships between high (Q99), mean (Qmean), and low (Q01) streamflow and urbanisation or tree cover change in 729 catchments in the United States between 1992 and 2018. We apply two statistical modelling approaches and compare their performance. Panel regression models estimate the average effect of land cover changes on streamflow across all sites, and show that on average, a 1%-point increase in catchment urban area results in a small (0.6-0.7%), but highly significant increase in mean and high flows. Meanwhile, a 1%-point increase in tree cover does not correspond to strongly significant changes in flow. We also fit a generalized linear model to each individual site, which results in highly varied model coefficients. The medians of the single-site coefficients show no significant relationships between either urbanisation or tree cover change and any streamflow quantile (although at individual sites, the coefficients may be statistically significant and positive or negative). On the other hand, the GLM coefficients may provide greater nuance in catchments with specific attributes. This variation is not well

represented through the panel model estimates of average effect, unless moderators are carefully considered. We highlight the value of statistical approaches for large-sample attribution of hydrological change, while cautioning that considerable variability exists.

4.2 Introduction

4.2.1 Land cover effects on streamflow

Significant trends have been detected in historical streamflow records across the United States (Archfield et al., 2016; Douglas et al., 2000; Lins & Slack, 1999, 2005; Rice et al., 2015; Sadri et al., 2016; Slater & Villarini, 2016b; Tamaddun et al., 2016; Y.-K. Zhang & Schilling, 2006). Shifts in climate characteristics, such as precipitation totals, phase, and timing are widely considered to be the dominant drivers of hydrologic change, but land cover changes, consisting of changes in the biological or physical features present in a landscape (e.g., forested or urban area), also have the potential to drive changes in streamflow, potentially even offsetting the influence of climate (Slater et al., 2021a). The effects of land cover changes on hydrological extremes, such as worsening flood risk (Bradshaw et al., 2007; Dijk et al., 2009), or as potential mechanisms by which hydrological and climatic risks may be managed or offset (Dadson et al., 2017; Dixon et al., 2016) remain poorly understood. It is clear that land cover changes can alter hydrological response to precipitation events by influencing the degree and rate at which water is intercepted and evaporated, stored, or allowed to run off into a river channel (Filoso et al., 2017; Jacobson, 2011; Shuster et al., 2005), however, the extent to which they do so lacks clear definition.

There are a number of ways in which land cover change might be expected to influence the magnitude of high, mean, and low daily streamflow. Widely discussed in the literature, urbanisation is typically expected to increase high flows and flood risk (Blum et al., 2020; Hodgkins et al., 2019; Hollis, 1975; Prosdocimi et al., 2015; Salavati et al., 2016; W. Yang et al., 2021) and to a lesser extent, water balance or mean annual flows (Oudin et al., 2018b; Salavati et al., 2016). Urbanisation has been associated with a wide range of changes in low or base flows in general, including significant decreases and increases in flow, likely dependent on the specific activities associated with urbanisation in a given catchment (Dudley et al., 2020; Jacobson, 2011; O'Driscoll et al., 2010).

The relationship between streamflow and tree cover change is less well defined in the literature. Typically, one might expect tree loss or deforestation to be associated with increases in streamflow, and afforestation to be associated with decreases in streamflow. This expectation

remains consistent in the literature for low, mean and high flows generally (Ahn & Merwade, 2017; Bladon et al., 2019), although the relationship appears particularly well defined for mean flows (Brown et al., 2005; Hibbert, 1965; Swank et al., 2001) and low flows (Smakhtin, 2001) with evidence suggesting pronounced decreases in flow following afforestation (Farley et al., 2005) and vice versa. While there is generally agreement regarding the ways in which changes in each of these land cover types affect streamflow, there exists a range of research which reports contrasting or even non-existent associations between these variables (Bart et al., 2016; Biederman et al., 2014, 2015; Goeking & Tarboton, 2020; Guardiola-Claramonte et al., 2011; Slinski et al., 2016).

Understanding the role that both tree cover change and urbanisation play in the hydrologic cycle is increasingly important in light of growing discussion around natural flood management, which focusses on the use of land management strategies for mitigating flood risk (Dadson et al., 2017), and nature-based solutions to climate change (Cohen-Shacham et al., 2016). This discussion centres the potential for afforestation to sequester carbon (Bastin et al., 2019), and in a hydrological context, begs the questions: what effects would large-scale afforestation have on water availability and risk? Conversely, to what extent is urbanisation altering our susceptibility to water related risks?

While many have attempted to understand the influence of land cover changes on streamflow, the breadth of knowledge that we have about those relationships is deep, but incomplete. Most of these studies have used small sample sizes and employed methods such as paired catchment analysis (Brown et al., 2005; Prosdocimi et al., 2015; Seibert & McDonnell, 2010; L. Zhang et al., 2012), or simulation and modelling-based approaches (Hejazi & Markus, 2009; Hidalgo et al., 2009; Schilling et al., 2014). These approaches can be very powerful when good counterfactuals exist, e.g. when a suitable paired catchment is present and small sample analyses are useful for understanding physical relationships within a single catchment or a limited number of catchments. They are not, however, particularly well-suited to extrapolating findings across larger regions and making generalized statements about hydrological behaviour.

Much large sample research has relied on regression techniques to develop our understanding of the potential effects of land cover changes on streamflow. There are some definitive benefits to these approaches. For example, statistical approaches may allow for the quantification of relationships between flow characteristics and catchment descriptors for which data are

available (e.g. potential for differing effect sizes based on soil type or air temperature), as well as the ability to state a level of confidence in results (Gupta et al., 2014). There are two dominant statistical approaches to attribute the drivers of large-sample hydrological change in the literature. First, a single-catchment approach involves fitting distinct regression models to often 'lumped' time-series data for individual catchments, then assessing the fit of these models and signs of their coefficients (Neri et al., 2019; Prosdocimi et al., 2015; Slater et al., 2021; Villarini, Smith, et al., 2009). Alternatively, a combined multi-catchment approach involves fitting panel regression models to estimate average causal effects across many sites (Bassiouni et al., 2016; Blum et al., 2020; Brady et al., 2019; De Niel & Willems, 2019; Lombard & Holtschlag, 2018; Steinschneider et al., 2013; W. Yang et al., 2021). The deceptively simple nature of regression approaches means that they have been widely applied, however, while both single-catchment and multi-catchment approaches have their unique benefits, they are best suited to slightly different questions.

4.2.2 Study scope

The aims of this work are two-fold. We first seek to improve understanding of the relationships between land cover changes, specifically tree cover and urban area, and low, mean, and high annual streamflow. Then, we compare the results from two different statistical techniques, a multi-catchment panel regression approach, and a single-catchment regression approach applied to the same sites. Specifically, we address the following research questions and hypotheses:

1: How are urbanisation and tree cover change associated with or affecting streamflow across the conterminous United States?

In accordance with prior research, we hypothesize that urbanisation may result in increased mean and high flows, and that low flow relationships will be more varied (Blum et al., 2020; Prosdocimi et al., 2015; Villarini, Smith, et al., 2009), while afforestation (deforestation) may decrease (increase) streamflow for all parts of the hydrograph.

2: How do the results of single-catchment and multi-catchment (panel) regression methods differ?

We expect that our panel regression model coefficients will roughly correspond with the central summary of the distribution of the combined single-catchment regression coefficients; however, the panel model estimates will exhibit less variability, demonstrating that they are a

more reliable metric for estimating the typical effect of different drivers on flow across a wide scale.

4.3 Data

We use the Geospatial Attributes of Gages for Evaluating Streamflow version II (GAGES II) dataset (Falcone, 2011) as a basis for selecting catchments to include in the analysis. GAGES II contains geospatial characteristics and catchment descriptors for 9,322 gauged river basins in the United States which had a long flow record at the time of its creation in 2011 (Falcone, 2011). We downloaded the daily streamflow data between 1992 to 2018 for all catchments in the GAGES II dataset from the U.S. Geological Survey using the R package 'dataRetrieval' (DeCicco et al., 2024) and calculated the annual 0.99 (Q99), mean (Qmean), and 0.01 (Q01) quantiles of the daily streamflow to represent high, mean and low flows respectively for each calendar year. We then used the catchment boundaries associated with each of these gauge sites from the National Hydrography Dataset version 1 (NHDv1) Watershed Boundary Dataset (WBD) (United States Geological Survey and United States Department of Agriculture., 2020) to quantify the annual average percentage of tree cover and urban area in each catchment from the European Space Agency (ESA) Climate Change Initiative (CCI) Global Land Cover dataset (300 m resolution; 1992-2018). The dataset is described in more detail later in this section (ESA CCI, 2017).

We then assessed the daily streamflow data from all catchments to ensure that they had:

- at least 20 years of 95% complete daily streamflow records (more than 347 days per year) in the years for which land cover data are available;
- no 0 flow values in the annual flow quantiles;
- not experienced more than one day of upstream dam storage, calculated by dividing the total upstream dam storage by the estimated catchment annual runoff, both taken from GAGES II (Blum et al., 2020; Falcone, 2017; Hodgkins et al., 2019);
- and a minimum of four distinct values in the land cover timeseries being assessed (i.e. experienced land cover changes over time).

We further require: the presence of some land cover change to have occurred in each catchment because this is necessary to be able to fit the single site models, no 0 flow values because the low flow behaviour of these catchments may be too complex to model well with a regression-based approach, and removed catchments with greater than one day of dam storage because we

expect that dam storage may be used to counter high runoff relevant to flood events. Other factors, like possible diversions from watersheds, were not considered here.

The final dataset included the high, mean, and low annual quantiles of the daily streamflow data for 729 catchments. On average, sites used in the analysis had 26.5 years of complete daily streamflow data. We organized the discharge data according to calendar years (Jan-Dec) rather than water years (Oct-Sept) to maintain consistency between the climate and land cover datasets. Calendar years are also better suited to studying low flows, which occur in the autumn in many US catchments (Sadri et al., 2016).

While the ESA-CCI land cover dataset is based in large part on the medium resolution imaging spectrometer surface reflectance (MERIS SR) timeseries, the urban area class is supplemented by two external sources (ESA CCI, 2017): The Global Human Settlement dataset, created using Landsat imagery and validated using population, amongst other sources (Pesaresi et al., 2016) and the Global Urban Footprint dataset, derived using high-resolution information from the synthetic aperture radar satellites, TerraSAR-X and TanDEM-X (Esch et al., 2013). The ESA-CCI dataset has an estimated overall accuracy of around 71% (ESA CCI, 2017) and the user accuracy estimates for the land cover classes which we use in this work are generally higher, e.g. for four of the tree cover classes (ranging from 75% to 90% accuracy) and urban area (75%). It is worth noting that these accuracy estimates are global averages based on the 2015 land cover data. Actual accuracy is likely to be much higher over the study area because the number of valid observations is high for the USA (ESA CCI, 2017). It is also possible that accuracy varies from year to year.

We aggregated the original land cover classification categories into seven broad groups (Supplementary Materials Table 8.1) based on the recommendations of the United Nations Convention to Combat Desertification (UNCCD) good practice guidance for Sustainable Development Goals (SDG) Indicator 15.3.1 (Sims et al., 2017), prior to calculating the catchment percentages of land cover area. We then retain the data for urbanisation and tree cover change for analysis (Figure 4.1). In referring to tree cover change, we use the term "afforestation" as the equivalent of a net increase in tree cover, and "deforestation" to refer to a net decrease in tree cover, but we do not consider the mechanisms by which tree cover change has occurred (e.g. reforestation). Evidence suggests that effects are similar regardless of the mechanism by which change occurred (Filoso et al., 2017). Urban area did not decrease in any catchment. On the other hand, tree cover change was not unidirectional; a given catchment may

have experienced relative gains and losses in tree cover in different years, over the period of record.

We compute catchment-averaged annual precipitation and mean annual temperature from 4 km x 4 km resolution annual Parameter-elevation Regressions on Independent Slopes Model (PRISM) data (C. Daly et al., 2008; Di Luzio et al., 2008) accessed using the R package 'prism' (Hart & Bell, 2015a). The PRISM dataset is the most widely used spatial climate dataset in the United States, and is the official climate dataset for the US Department of Agriculture (C. Daly & Bryant, 2013). Finally, we use the U.S Geological Survey Physiographic divisions of the United States (Figure 4.1) to represent geomorphic and geologic characteristics in our multi-catchment models (Fenneman & Johnson, 1946).
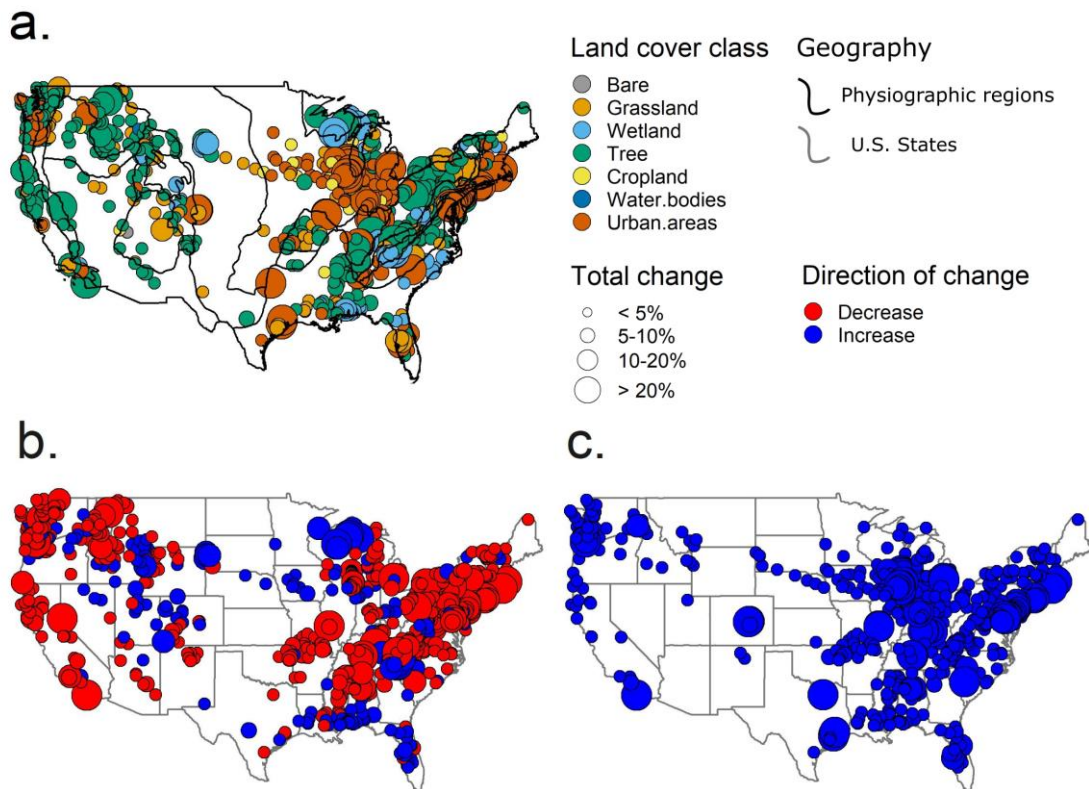
Figure 4.1 Dominant land cover types and total absolute change in land cover in the United States.

a. Total absolute change (1992-2018) in the land cover class which in 2018 occupied the largest percentage of the catchment area in the 729 catchments used in this analysis. Land cover classes represent aggregated groups (Supplementary Materials Table 8.1) based on the ESA CCI Global Land Cover dataset. Panels b and c show total changes in tree cover (b) and urban areas (c) respectively in the 729 catchments used in this analysis.

4.4   Methods

4.4.1   Causal diagrams

We construct causal diagrams to outline the potential relationships within the hydrological system (Blum et al., 2020), and inform the design of our regression models. The mechanisms by which tree cover change and urbanisation might influence streamflow are outlined in the causal diagrams in Figure 4.2. The arrows in these diagrams denote causal relationships and not physical pathways. Confounders are variables which could potentially influence both the land cover variable in question as well as streamflow. Moderators are catchment characteristics which are likely to influence the degree to which different land cover changes influence

streamflow, but not whether or not there is a relationship between the two. The diagrams are used to help construct the models outlined in the remainder of this section.
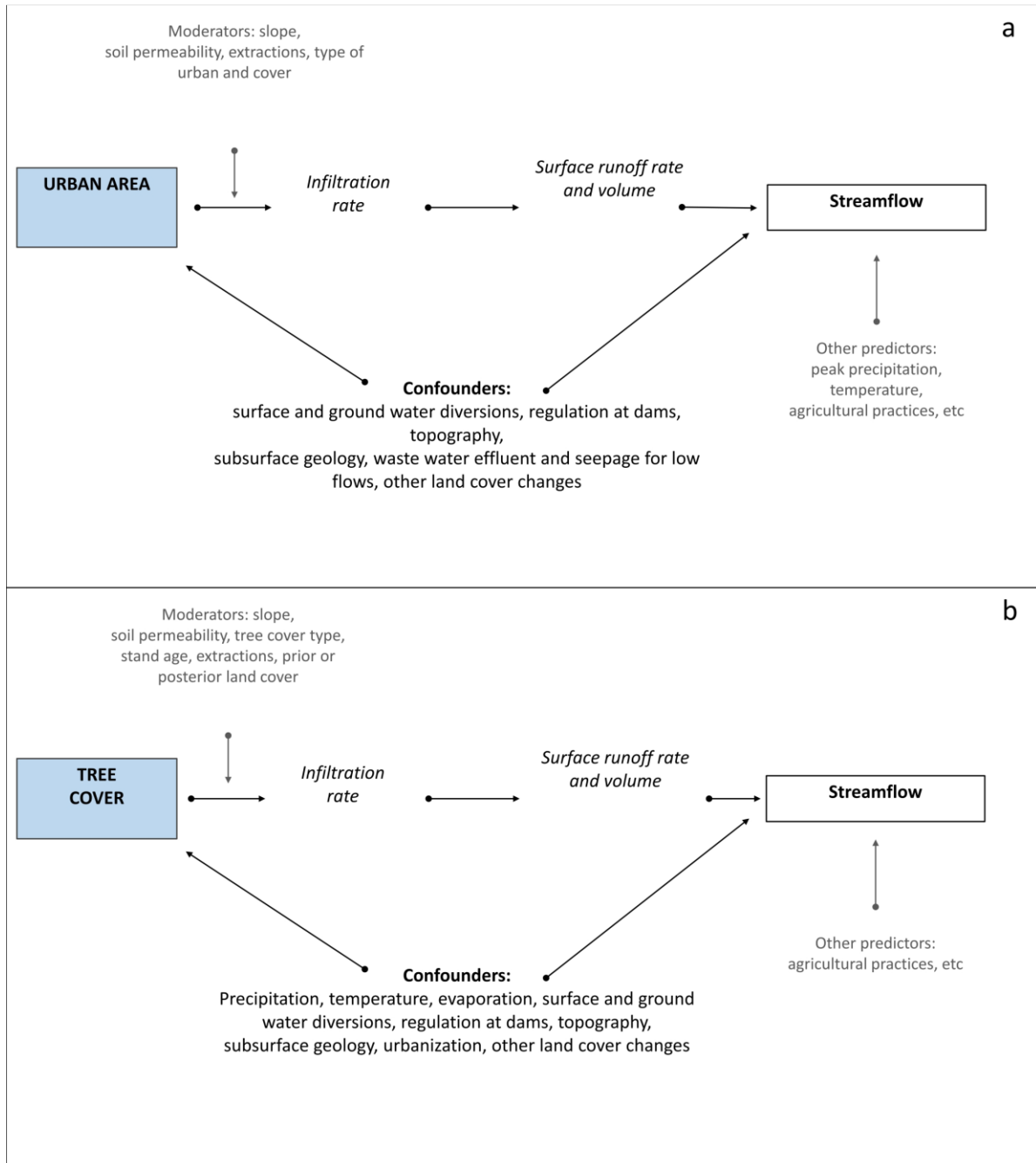
Figure 4.2 Causal diagrams depicting the relationships between streamflow and urban area or tree cover.

Urban area is in panel a and tree cover in panel b. Figures are adapted Blum et al. (2020).

Arrows denote causal relationships, not physical pathways.

4.4.2 Single-catchment models

Next, we fit generalized linear models (GLMs) using the R package gamlss (Rigby & Stasinopoulos, 2005) to each individual streamflow quantile time series (Table 4.1). The GAMLSSframework is highly flexible, however, the use of more complex models was attempted but did not result in increased goodness of fit. For more detailed discussion on the GAMLSS framework the reader is directed to Rigby and Stasinopoulos (2005) and Stasinopoulos & Rigby, (2007)

We fit GLMs using the lognormal distribution (LOGNO in the gamlss package) for the response variable Y (Table 4.1; Table 4.2.b.), so that the coefficients would be directly comparable across both modelling approaches. The lognormal distribution is parametrised by μ and σ, which, in the GAMLSS framework, can both be modelled as a function of explanatory variables. The μ parameter models the location of the distribution while the σ parameter (scale) relates to the dispersion of the distribution. Here we are concerned with the typical behaviour of the flow variables, and so focus on μ parameter modelled as a function of explanatory variables (Table 4.2.b.). We hold the σ parameter constant in our GLMs (Table 4.2.b) meaning that its estimation does not vary with time or any other variables. In effect, a GLM based on a log-normal distribution with constant σ is equivalent to the traditional multiple linear regression based on the ordinary least-squares estimator, in which the original response variable is log-transformed: the σ parameter is estimated using the residual sum of squares of the OLS fit.

Table 4.1 Equations for GLMs and Panel models.

VIF is the variance inflation factor used to select catchments where tree cover and urbanisation are strongly correlated. The variables used and terms estimated by these equations are described in Table 4.2.

Type	#	Model	Equation	Description
GLM	4.1	Both	$$Y_{i,t} \sim \ln(\mu_{i,t}, \sigma_i^2)$$ $$\mu^{i,t} = \alpha^i + \beta_1{}^i urban_{i,t} + \beta_2{}^i tree_{i,t} + \varepsilon_{i,t}$$ $$\sigma_i = k;$$	Model fit exclusively to estimate VIF
GLM	4.2	Tree Cover	$$y_{i,t} \sim \ln(\mu_{i,t}, \sigma_i^2)$$ $$\mu^{i,t} = \alpha^i + \beta_2{}^i tree_{i,t} + \theta_1{}^i \ln(P_{i,t}) + \theta_2{}^i Tmean_{i,t} + \varepsilon_{i,t}$$ $$\sigma_i = k;$$	Single catchment model for tree cover only
GLM	4.3	Urbanisation	$$y_{i,t} \sim \ln(\mu_{i,t}, \sigma_i^2)$$ $$\mu^{i,t} = \alpha^i + \beta_1{}^i urban_{i,t} + \theta_1{}^i \ln(P_{i,t}) + \theta_2{}^i Tmean_{i,t} + \varepsilon_{i,t}$$ $$\sigma_i = k;$$	Single catchment model for urban area only
Panel	4.4	Tree Cover	$$\ln(Y_{r(i),t}) = \alpha_i + \beta_2 tree_{i,t} + \theta_1 \ln(P_{i,t}) + \theta_2 Tmean_{i,t} + \delta_t D_t + \gamma_{t,r} D_t D_r + \varepsilon_{i,t}$$	Panel model for tree cover only
Panel	4.5	Urbanisation	$$\ln(Y_{r(i),t}) = \alpha_i + \beta_1 urban_{i,t} + \theta_1 \ln(P_{i,t}) + \theta_2 Tmean_{i,t} + \delta_t D_t + \gamma_{t,r} D_t D_r + \varepsilon_{i,t}$$	Panel model for urban area only

In some instances, tree cover and urban area changes will be correlated with one another. Urbanisation is a confounder for tree cover because changes in tree cover are potentially caused by urbanisation and, urbanisation is also likely to have an effect on streamflow. We therefore introduce a variable selection procedure to select the data incorporated into the models for each catchment. We prioritize reducing collinearity between land cover variables to improve our coefficient estimations because when two or more variables in the model are highly correlated they are more likely to provide redundant information about the response, and reduce our ability to interpret the results in a meaningful way (James et al., 2013).

In catchments in which only tree cover or urban area change was present in the study period, we are not concerned about collinearity between the land cover variables. In those catchments, we use the land cover change variable which was present and do not apply the variable selection procedure. In catchments where both tree cover change and urbanisation occurred, we examine the collinearity between the land cover variables by fitting a log-normal GLM (Table 4.1) for which the $\mu^{i,t}$ parameter includes both tree cover and urbanisation variables, as defined in Equation 4.1 (Table 4.1; Table 4.2.a; Table 4.2.b). We then estimate the variance inflation factor (VIF) for Equation 4.1 (with Qmean as the predictand) using the R package car (Fox & Weisberg, 2011) to determine the impact of collinearity on the precision of the model parameter estimation. VIF has a minimum possible score of 1 (no collinearity), and as a rule of thumb, a VIF of greater than either 5 or 10 can be considered to have a potentially dangerous level of collinearity (James et al., 2013). Since our intention is to interpret the regression coefficients as a form of attribution, we adopt a conservative VIF threshold of 2.5. Then if VIF is > 2.5 we retain only urban area in the model, and if VIF < 2.5 for a catchment, we retain the land cover variable which experienced the largest absolute change between 1992 and 2018 and exclude the other, using Equations 4.2 and 4.3 as our final single site regression models for the analysis (Table 4.1; Table 4.2.a; Table 4.2.b). Our approach prioritizes urbanisation when the two variables are collinear because it makes intuitive sense that urbanisation is a likely driver of changes in tree cover, rather than the reverse.

Table 4.2 Description of variables used and terms estimated by the GLMs and Panel models which are described by the equations in Table 4.1.

Name	Models	Description	Purpose
a.	**All independent variables**		
$urban_{i,t}$	1-5	Annual catchment averaged urbanisation in %	
$tree_{i,t}$	1-5	Annual catchment averaged tree cover in %	
$\ln(P_{i,t})$	1-5	Natural logarithm of catchment averaged total annual precipitation in mm	
$Tmean_{i,t}$	1-5	Mean annual catchment averaged temperature in degrees C	
D_t	4-5	Dummy variable for year	
$D_t D_r$	4-5	Interaction between annual dummy variables and physiographic regions	
b.	**GLM terms**		
$Y_{i,t}$	1-3	One of three annual streamflow quantiles (0.1, mean, and 0.99) representing low, mean, and high flows, estimated from daily streamflow data, for each streamgauge (i) and each annual timestep (t) in m³/s	Response variable
$\mu^{i,t}$	1-3	Median of the GLM for each site *(i)* at time step *(t)* using the maximum likelihood estimator	Represents the predicted value of Y for each site and year
σ_i	1-3	Sigma parameter for the GLM model estimation	
k	1-3	A constant defined as 1	
α_i	1-5	Streamgauge specific intercept of the fitted model	
$\beta_1{}^i$	1,2	Estimated influence of a 1%-point increase in urban area on streamflow for a single site	Association estimated by this model
$\beta_2{}^i$	1,3	Estimated influence of a 1%-point increase in tree cover on streamflow for a single site	Association estimated by this model
$\theta_1{}^i$	1-3	Estimated influence of a 1% change in total annual precipitation on streamflow for a single site	Controls for association between precipitation and streamflow
$\theta_2{}^i$	1-3	Estimated influence of a 1 degree change in mean annual temperature on streamflow for a single site	Controls for association between temperature and streamflow
c.	**Panel terms**		
$Y_{r(i),t}$	4-5	One of three annual streamflow quantiles (0.1, mean, and 0.99) representing low, mean, and high flows, estimated from daily streamflow data, for each streamgauge (*i*), in region (*r*), and annual timestep (*t*)	Response variable
α_i	4-5	Streamgauge specific intercept of the fitted model	Controls for time invariant confounders at the basin level

β_1	4	Average effect of a 1%-point increase in urban area on streamflow	Causal effect estimated by this model
β_2	5	Average effect of a 1%-point increase in tree cover on streamflow	Causal effect estimated by this model
θ_1	4-5	Estimated average influence of a 1% change in total annual precipitation on streamflow; not considered causal	Controls for confounding effect of precipitation on tree cover; allows estimation of precipitation elasticity
θ_2	4-5	Estimated average influence of a 1 degree change mean annual temperature on streamflow; not considered causal	Controls for confounding effect of temperature on tree cover; allows estimation of temperature elasticity
δ_t	4-5	Overall effect of annual flow magnitude on year (t)	Controls for time-varying confounders at the national level
$\gamma_{t,r}$	4-5	Overall effect of annual flow magnitude on year (t) in region (r)	Controls for time varying confounders at the physiographic region level

While temperature and precipitation are not confounding variables for urbanisation, they are important predictors of flow and are therefore included in our models. We use the natural log of the precipitation variable in order to make the coefficients interpretable as equivalent to associated percentage change in the original flow scale. We apply an exponential transformation to the land cover and temperature coefficients so that they are interpretable relative to the non-log transformed streamflow quantiles as described in Supplementary Materials Text 8.1.

4.4.3   Multi-catchment (panel) regression design

Panel regression models have recently been applied in hydrological research (Blum et al., 2020; Davenport et al., 2020; De Niel & Willems, 2019; Ferreira & Ghimire, 2012; Levy et al., 2018; Müller & Levy, 2019; W. Yang et al., 2021). The approach allows consideration of the data across both time and space; here we quantify the average effects of individual drivers (changes in tree cover and urban area) across all sites while controlling for the influence of a wide range of confounding variables (Figure 4.2). Through careful consideration of the data and the aid of the causal diagrams, we attempt to isolate a causal effect (Ferraro et al., 2019; Pearl, 2009). Therefore, when discussing the results of the panel models, we refer to an "effect" size and not merely an association between variables. We formulate our panel regression models based on the design proposed in Blum et al. (2020) with some modifications, and fit them using the R package `plm` (Croissant & Millo, 2008).

The panel models in Equations 4.4 and 4.5 test the effects of change in tree cover and urban area on streamflow respectively (Table 4.1; Table 4.2.a; Table 4.2.c). These models replicate

the Blum et al. (2020) model with some changes. While precipitation and temperature are unlikely to be significant confounders for the effect of urban area on streamflow, they are confounding variables in the tree cover model because, in addition to streamflow, they each might influence tree growth directly. As an example, an event such as prolonged drought might affect both tree cover and streamflow, and if we failed to include precipitation in the tree cover model, these effects might not be captured in the annual dummy variable. For this reason, we equally tested the urban area model in Equation 4.5 (Table 4.1; Table 4.2.a; Table 4.2.c) without the climatological variables (Supplementary Materials Text 8.1.3). Results showed that the inclusion of total annual precipitation and mean annual daily temperature had a small influence on the urbanisation coefficient (Supplementary Materials Table 8.2) resulting in slightly higher and more significant coefficients.

While there was substantial overlap in the confidence intervals of these two models, the difference in significance suggests that climatological confounders might not be fully controlled for in the model design if these variables are not explicitly defined. Therefore, we include climatological variables in both models so that all of the models are comparable, and so that climatological coefficients may be considered across models. We apply the same transformation to the land cover and temperature coefficients as we do to the GLM coefficients (Supplementary Materials Text 8.1).

The "fixed effects" (α_i) are intercepts specific to each streamgauge (Table 4.2.c). These account for possible confounding variables which are constant over time. Similar to Blum et al. (2020), we use an annual dummy variable to control for time-varying national scale confounders, an interaction term for the U.S Geological Survey Physiographic divisions of the United States (Fenneman and Johnson, 1946), with the annual dummy variables to control for regionally time-varying confounders (Figure 4.2; Table 4.2.c). The physiographic regions represent broadscale geomorphic regions based on similar terrain texture, rock type, geologic structure, and history (Fenneman and Johnson, 1946). These models provide a minimally-biased estimate of effect size assuming that there are no sub-regional time-varying factors impacting both streamflow and urban area.

Autocorrelation in fixed effects panel models can lead to the underestimation of standard errors. We address this concern by clustering standard errors at the streamgauge level (Arellano, 1987, 1987; Bertrand et al., 2004; Blum et al., 2020). The process is described explicitly in Supplementary Materials Text 8.2.

*Table 4.3 Threshold of minimum land cover change used in sensitivity analysis.
"Total" is the number of sites included in the model for each of the sub-datasets tested in the sensitivity and robustness analysis.*

Threshold	Model	Total
0%	Tree Cover	388
0%	Urbanisation	341
1%	Tree Cover	333
1%	Urbanisation	168
5%	Tree Cover	122
5%	Urbanisation	109

4.4.4   Sensitivity and robustness testing

In the first instance, we fit the models described in Equations 4.2–4.5 (Table 4.1; Table 4.2) to data from catchments in which change in tree cover or urban area was greater than 0% over the 1992 to 2018 time period. We then conducted a sensitivity analysis to challenge the robustness of the estimated land cover and climatic coefficients by incrementally increasing minimum thresholds of land cover change, then refitting each of the panel models and resampling the GLM results to the adjusted datasets. The number of catchments included in each subsample is described in Table 4.3 and depicted in Figure 4.3.
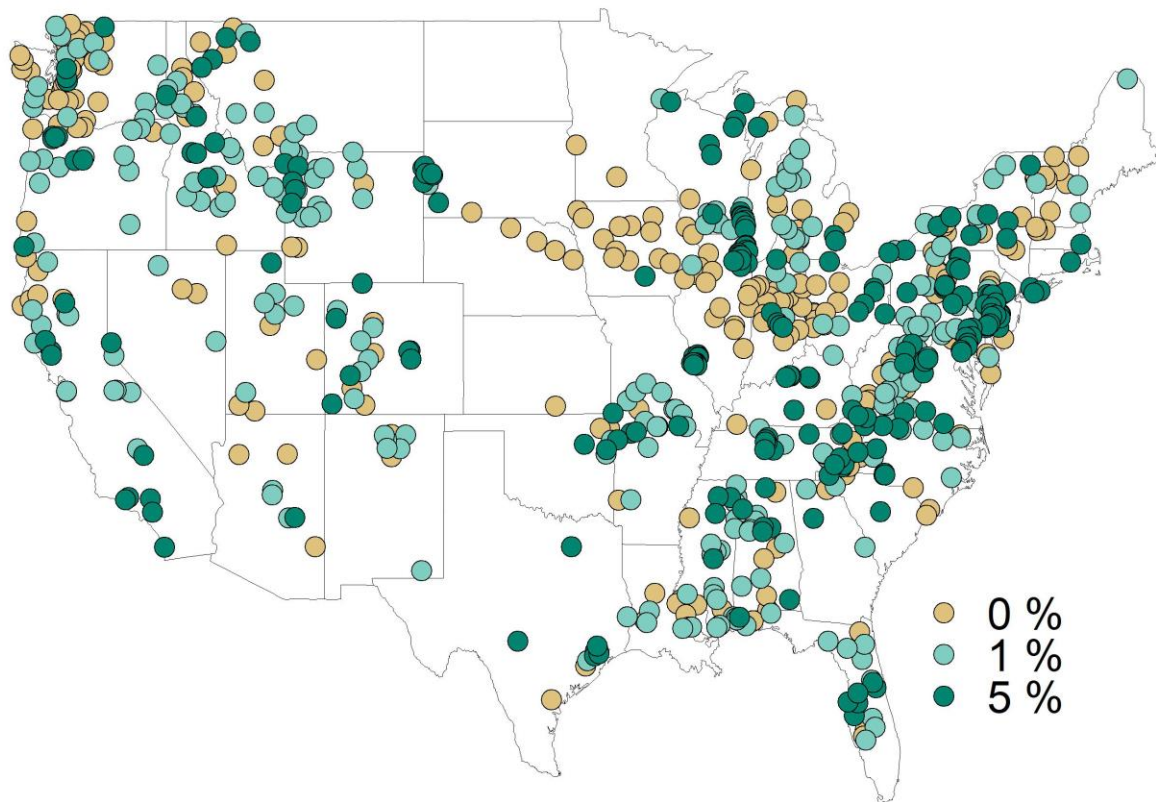
Figure 4.3 Location of catchments which met each land cover change threshold in the sensitivity analysis.

Thresholds are cumulative, so that sites in each category are also included in samples with lower thresholds for land cover change.

4.5   Results and discussion

4.5.1   Land cover coefficients

Our first question addresses the expectation that urbanisation would, on average, increase streamflow and that afforestation and deforestation would decrease or increase streamflow respectively. In sections 4.1.1. and 4.1.2. we focus on the models which were fit to all catchments which experienced any total absolute change in land cover ($> 0\%$) between 1992 and 2018.

4.5.1.1   Multi-catchment approach: 0% land cover change threshold

The effect of urbanisation on mean and high flows is positive ($\sim 0.6\%$) and highly statistically significant according to the panel model results (Table 4.4). The estimated effect sizes are small compared to those of Blum et al. (2020) who looked at annual floods, defined as maximum

annual daily discharge, using a similar methodology and estimated an average 3.3% effect of a 1%-point increase in impervious surface area on annual floods, from a sample of 280 catchments for sites which did not experience substantial dam storage. It is possible that the difference in our results when compared to Blum et al. (2020) relate to the urbanisation data, which was created from a different origin. However, Yang et al. (2021), also use the ESA CCI land cover data for a similar application in China, and estimate an approximately 3.9% effect of a 1%-point increase in urban area on annual floods in a sample of 757 catchments. It is more likely then, that the difference in results relates to the segment of the streamflow hydrograph which is being examined. Results for low flows in our urbanisation model are not statistically significant.

Table 4.4 Model results for land cover coefficients: 0% land cover change threshold.
The p values less than 0.1 are presented in bold font. Lower and upper bounds indicate the middle 90% of the distribution of GLM coefficients and the 90% confidence intervals of the panel models. Coefficients represent the % change in streamflow expected for a 1 %-point change in the land cover variable.

Q	Model	Variable	Estimate	p value	Lower bound	Upper bound
Q01	Urban GLM	Urban	-1.686		-98.045	80.679
Qmean	Urban GLM	Urban	-2.099		-85.091	17.349
Q99	Urban GLM	Urban	-2.704		-89.332	27.933
Q01	Urban Panel	Urban	0.34	0.4276	-0.363	1.048
Qmean	Urban Panel	Urban	0.623	< 0.0001	0.397	0.85
Q99	Urban Panel	Urban	0.675	0.0002	0.374	0.977
Q01	Tree GLM	Tree	-1.053		-20.271	21.946
Qmean	Tree GLM	Tree	0.241		-18.252	20.188
Q99	Tree GLM	Tree	1.223		-20.274	32.641
Q01	Tree Panel	Tree	0.275	0.209	-0.085	0.636
Qmean	Tree Panel	Tree	-0.261	0.1505	-0.559	0.038
Q99	Tree Panel	Tree	-0.384	0.1057	-0.773	0.006

The estimated panel effects of tree cover change are not statistically significant (Table 4.4). The weaker statistical significance may be indicative of unidentified confounding which dampens the estimated effect, may reflect the complex relationship between tree cover change and streamflow, or the challenges and limitations of the data used. However, strictly in terms of p-values, the model results indicate that tree cover change does not have a statistically significant average effect on any examined streamflow quantile. The low flow model also differs from the other tree cover panel models in that the resulting coefficient is positive (the inverse of those for mean and high flows), although not statistically significant.

While the insignificance of the effect of urbanisation on low flows and of tree cover overall indicates that these land cover changes have no effect in an average sense, it does not necessarily mean that these land cover changes have no effect in particular circumstances. Rather, it may speak to the wider range of possible effects. For low flows, smaller overall differences in runoff and the intricacies of land cover change may have a proportionally larger influence on streamflow than they do for higher flows. For instance, the location of waste water treatment plants in urban areas (Oudin et al., 2018b), landscape irrigation and water withdrawals and returns, as well as forest life stage (age) or the mechanisms by which deforestation occurs (Biederman et al., 2014, 2015; Slinski et al., 2016), or the aridity of a region (Goeking & Tarboton, 2020) may play an important role in determining the directionality of the effects of land cover change on low flows, depending on the particular characteristics of what comprises "urbanisation" or "tree cover change" (Smakhtin, 2001). In other words, if many outcomes are possible, developing a meaningful average across sites is far more difficult.

4.5.1.2   *Single-catchment approach: 0% land cover change threshold*

We address the same question using the single-catchment regression approach and expect that the land cover coefficients might vary more widely for streamflow in relation to tree cover change and for low flows than for higher flows in relation to urbanisation, based on the previous literature.

The expectation that urbanisation would, on average, increase streamflow and that afforestation and deforestation would decrease or increase streamflow, is not clearly supported by the single-catchment regression models (Figure 4.4; Table 4.4). In fact, the median coefficients for both land cover variables are indistinguishable from zero relative to the width of the middle 90% of

the distribution. In other words, in an average sense neither has an effect on streamflow. These results are contrary to expectations.

The distribution of coefficients is not dramatically different than the results of similar studies, however. For instance, Salavati et al. (2016) which, when using a paired catchment and a simulation driven residual analysis approach for 24 paired U.S. catchments all with more than 9% urbanisation, found a range of different associated changes in streamflow. The estimated changes in their study often widely overlapped 0 and the medians were small (median change adjusted for a 1%-point change in urban area for paired catchment models: Q95 = 0.5%, Qmean = 0.36%, Q05 = 0.36%; median change adjusted for a 1%-point change in urban area for regression-based models: Q0.95 = -0.03%, Qmean = -0.06%, Q0.05 = -0.06%). Oudin et al. (2018) used the same residual analysis approach to examine the effects of urbanisation on flow change in 142 catchments in the United States. They find that the median of these single catchment models for mean and high flows were negative and the middle 90% of the distribution overlapped 0 until a minimum urbanisation change threshold of 10% was used. Even then, the distribution was wide, and error bars crossed zero.
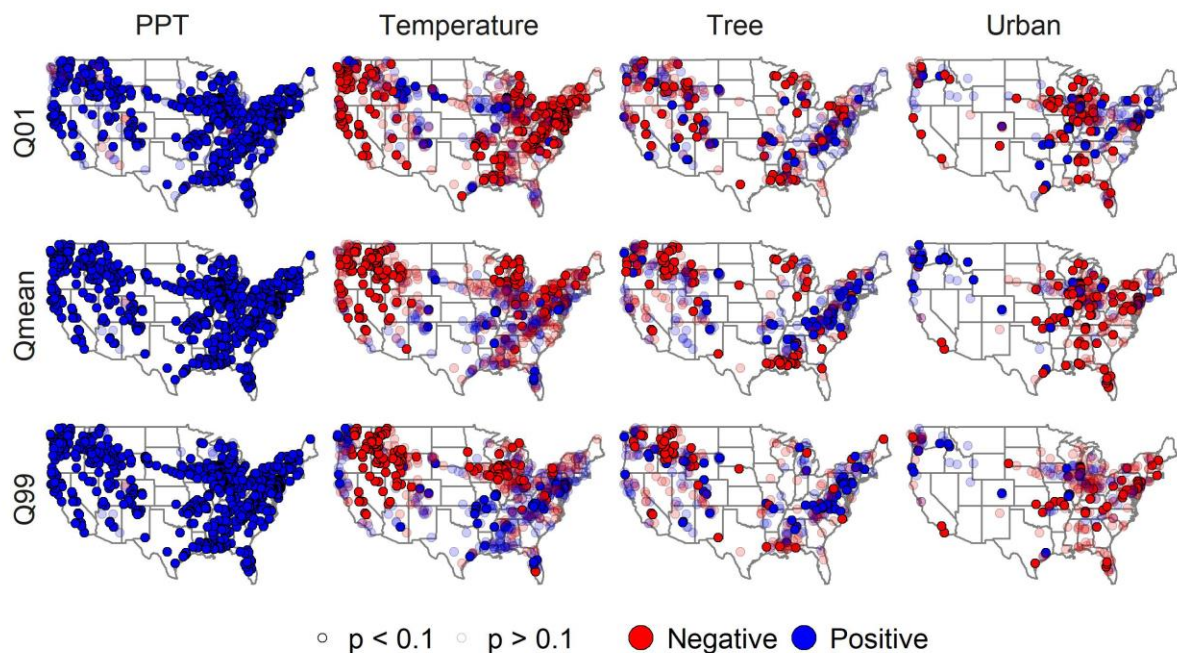
Figure 4.4 Sign of the coefficients for the land cover and climatological variables in the single-catchment models at the 0% land cover change threshold.

Red points represent a negative coefficient while blue represent a positive coefficient. Coefficients which are not significant at the p < 0.1 level are presented in a transparent shade.

We do notice some geographical patterns in the direction of the land cover coefficient resulting from the single site models (Figure 4.4). Most noticeably, there is a predominantly positive association between tree cover and streamflow in the eastern United States, and a predominantly negative or insignificant association in the western portion of the country. The insignificance of the tree cover coefficients in the southwestern United States is in line with the conclusions of Goeking and Tarboton (2020) who suggest that deforestation may often fail to increase water yield in semi-arid western watersheds. There is no immediately apparent explanation for the largely negative urbanisation coefficients which do not appear to follow a clear regional pattern, although, estimated coefficients may be more reliable in catchments with greater land cover change or better-fitting models.

4.5.1.3   A multi-catchment vs. single-catchment approach

In the previous subsections, results for two statistical modelling strategies have been presented. We next consider how the results of single-catchment and multi-catchment (panel) regression methods differ. In principle, the multi-catchment, panel regression approach is substantially more robust than the single-catchment models.

We conducted a sensitivity analysis to examine the robustness of each modelling approach to changes in the data sample, refitting the panel models or adjusting the GLM sample three times each. In the first instance, we fit each model to all sites with any percentage change in the land cover variable; then only to sites with more than 1%-point change; then, finally, to sites with more than a 5%-point change in land cover.

We first consider the land cover coefficients resulting from both the GLMs and the panel regression models fit for each of the data samples selected according to the sensitivity analysis (Table 4.3). The variation in resulting GLM coefficients is dramatically reduced as the threshold for land cover change is increased (Figure 4.5). Similarly, the standard errors of the coefficients decrease as the absolute values of catchment averaged land cover change increase (Figure 4.6). This improvement in the reliability of the estimated GLM coefficients is especially apparent for the urbanisation coefficients, and suggests that implementing a minimum threshold for land cover change may remove some catchments with spurious results.

Figure 4.5 Sensitivity analysis and comparison of the land cover variables in the GLMs and panel models.

Coloured bars represent the middle 90% of the distribution of GLM coefficients, and the 90% confidence intervals of the panel models. The medians of each group are represented by a black horizontal line. The y-axis is presented on a pseudo-log scale which maps numbers to a signed logarithmic scale with a smooth transition to linear scale around 0.

Despite this shift, the middle 90% of the distribution of GLM coefficients always overlap 0, even with increasing land cover change thresholds. This behaviour differs from that of temperature, for which the distributions of GLM coefficients continue to vary widely even with reduced sample sizes, and from the precipitation coefficients, for which the distributions are narrow, but consistent over time (Figure 4.7, Section 4.2). Meanwhile, the effects of land cover change on streamflow may occur inconsistently or be too small to be reliably detected at the catchment scale, unless substantial land cover change has taken place (Figure 4.6). On the other hand, the estimated effects of urbanisation and tree cover change in the panel models are relatively consistent, even in instances where the effects are not statistically significant, speaking to the robustness of the modelling approach – and its potential to isolate relatively marginal effects.
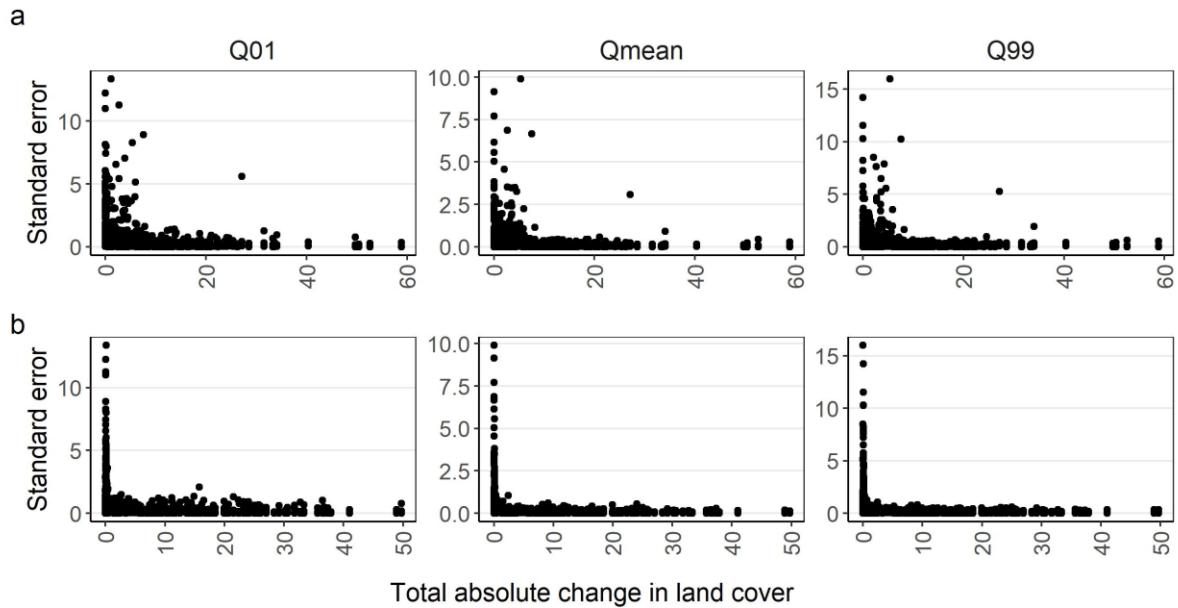
Figure 4.6 Scatterplot of the standard errors of the land cover coefficients from the GLM models (y) vs. the absolute change in a. tree cover and b. urban area between 1992 and 2018.

4.5.2 Climatological coefficients

While the focus of this study is not to quantify the relative importance of land cover change when compared to climatological variables, or to quantify the climatological elasticity of streamflow, we consider the coefficients for mean annual temperature (% change in streamflow per 1-degree Celsius change in temperature) and total annual precipitation (% change in streamflow per 1% change in precipitation) and compare them between models as one mechanism by which we validate model performance.

When comparing the coefficients from the GLM and panel models across all thresholds, the relationship between precipitation and high, mean, and low flows is clear. The median precipitation coefficients from the GLMs closely align with the coefficients from the panel models (Figure 4.7; Table 4.5), for the most part. The coefficients are positive, strongly significant (panel models), and the 90% confidence intervals of the panel models consistently overlap with the middle 90% of the distribution of GLM coefficients (Figure 4.7). The precipitation coefficients in the panel models are smallest for low flows (Table 4.5). Similarly, the range of resulting GLM coefficients is widest for low flows and narrowest for mean flows, indicating wider ranging possible relationships between more extreme flows and precipitation change depending on catchment specific characteristics (Figure 4.7; Table 4.5).

Table 4.5 Model results for climatological coefficients: 0% land cover change threshold. The p values less than 0.1 are presented in bold font, as are GLM coefficients for which the confidence intervals do not cross zero. . Lower and upper bounds indicate the middle 90% of the distribution of GLM coefficients and the 90% confidence intervals of the panel models. Coefficients represent the % change in streamflow expected for a 1% increase in total annual precipitation or a 1°C increase in mean annual temperature.

Q	Model	Variable	Estimate	p value	Lower bound	Upper bound
Q01	Urban GLM	Total annual PPT (%)	1.365		0.235	3.334
Qmean	Urban GLM	Total annual PPT (%)	1.683		1.161	2.535
Q99	Urban GLM	Total annual PPT (%)	1.796		1.151	2.746
Q01	Urban Panel	Total annual PPT (%)	0.908	0.004	0.39	1.427
Qmean	Urban Panel	Total annual PPT (%)	1.12	0.0027	0.506	1.733
Q99	Urban Panel	Total annual PPT (%)	1.256	0.0026	0.57	1.943
Q01	Tree GLM	Total annual PPT (%)	1.001		0.178	2.656
Qmean	Tree GLM	Total annual PPT (%)	1.487		0.793	2.184
Q99	Tree GLM	Total annual PPT (%)	1.687		0.844	2.696
Q01	Tree Panel	Total annual PPT (%)	1.058	< 0.0001	0.94	1.175
Qmean	Tree Panel	Total annual PPT (%)	1.452	< 0.0001	1.344	1.56
Q99	Tree Panel	Total annual PPT (%)	1.853	< 0.0001	1.719	1.986
Q01	Urban GLM	Mean annual temperature (C)	-6.802		-24.212	12.305
Qmean	Urban GLM	Mean annual temperature (C)	-1.057		-7.006	5.992
Q99	Urban GLM	Mean annual temperature (C)	0.197		-12.267	14.224
Q01	Urban Panel	Mean annual temperature (C)	-2.512	0.1876	-5.559	0.634
Qmean	Urban Panel	Mean annual temperature (C)	-3.196	0.0496	-5.794	-0.526
Q99	Urban Panel	Mean annual temperature (C)	-2.857	0.1293	-5.863	0.245
Q01	Tree GLM	Mean annual temperature (C)	-8.419		-27.1	6.276
Qmean	Tree GLM	Mean annual temperature (C)	-3.929		-15.009	4.644
Q99	Tree GLM	Mean annual temperature (C)	-0.338		-20.779	15.957
Q01	Tree Panel	Mean annual temperature (C)	-4.366	0.0108	-7.082	-1.571
Qmean	Tree Panel	Mean annual temperature (C)	-1.321	0.1585	-2.839	0.222
Q99	Tree Panel	Mean annual temperature (C)	-0.793	0.5881	-3.165	1.636

The precipitation coefficients are larger in the tree cover panel models relative to urbanisation models, while the reverse is generally true for the GLMs. However, the confidence intervals overlap, indicating that this difference is not substantial. Finally, as expected, the middle 90% of the distribution of the GLM precipitation coefficients generally exhibits more variability than the corresponding confidence intervals of the panel regression models. However, the 90% confidence interval of precipitation in the mean and high flow urbanisation panel models are wider (lower statistical significance), rivalling those of the GLMs.

Across all land cover change thresholds tested, the precipitation coefficients from the two models remain fairly consistent, always with overlapping distributions, speaking to a generally robust relationship between precipitation change and flow change (Figure 4.7; Table 4.5). The most notable change is a narrowing of the panel model confidence intervals and an increase in the estimated effects for all flow quantiles when the threshold for urbanisation was increased from 1% to 5%. It is possible that this hints at a nonlinear relationship between precipitation and streamflow, i.e. more urbanized catchments experience higher precipitation elasticity of streamflow than less urbanized ones. However, because we do not explicitly test interaction terms between precipitation and urbanisation, we cannot say for certain. It is also possible that the change in the estimation is due to the shift in sample size. Further exploration of the potentially varying effects of precipitation on streamflow under different land cover scenarios is needed in the future, particularly as the spatial distribution of catchments across the study area remains relatively consistent even with the increased thresholds (Figure 4.3).
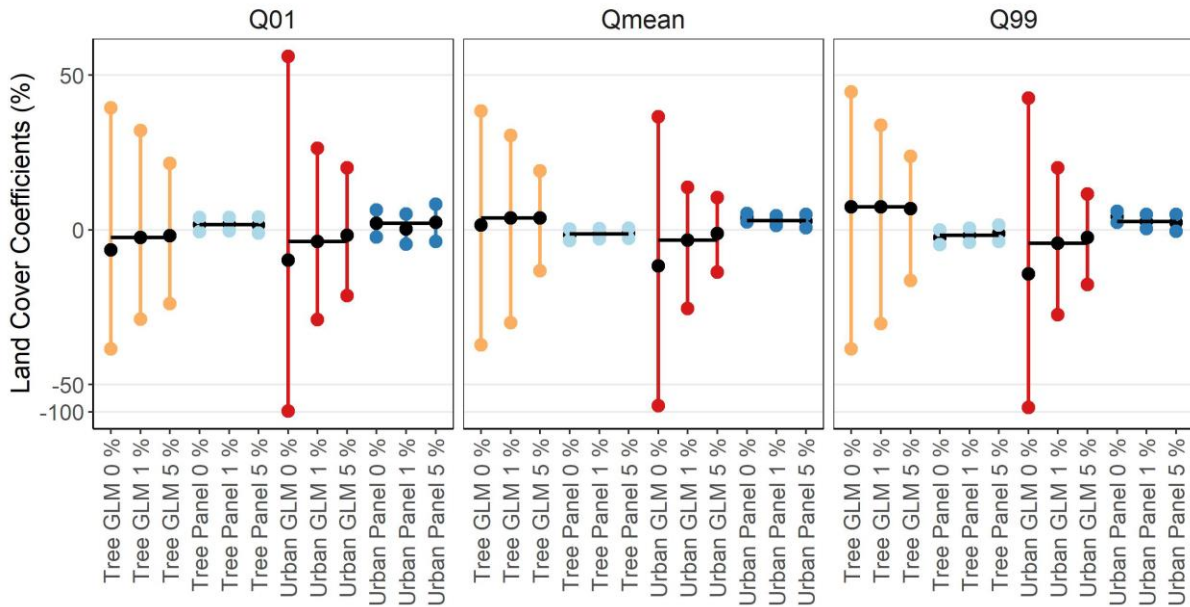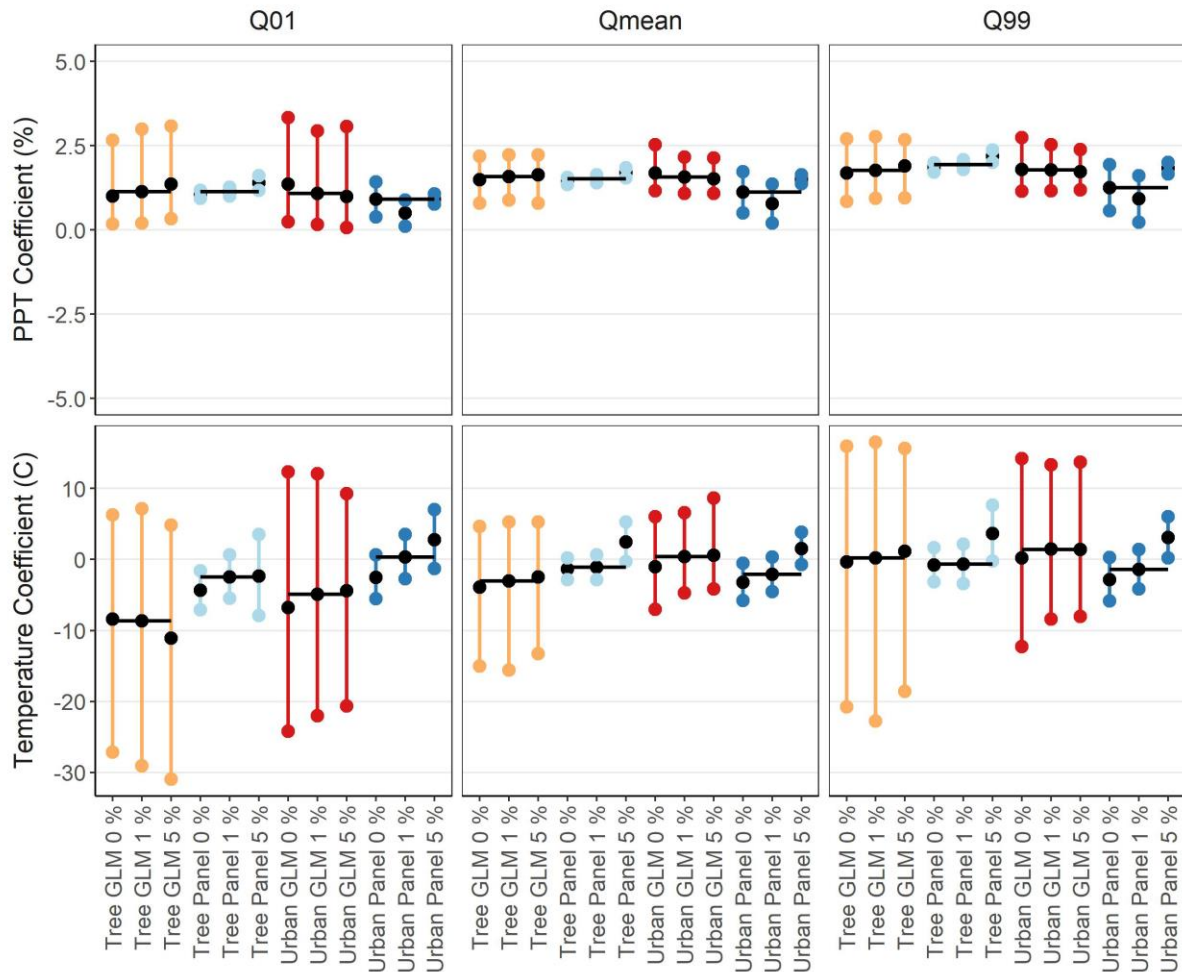
Figure 4.7 Sensitivity analysis and comparison of the climatological variables in the GLMs and panel models.

Coloured bars represent the middle 90% of the distribution of GLM coefficients, and the 90% confidence intervals of the panel models. The medians of each group are represented by a black horizontal line. Coefficients for temperature represent the expected % change in streamflow for each 1℃ change in mean annual temperature, and coefficients for precipitation represent the expected % change in streamflow for a 1% change in annual total precipitation.

The temperature coefficients vary more widely for different land cover sensitivity thresholds (Figure 4.7; Table 4.5). While the coefficients resulting from the panel models with 0% change threshold are statistically significant for mean flows in the urbanisation model (p <0.10) and for low flows in the tree cover model (p <0.05), the statistical significance of these coefficients is inconsistent as the land cover change threshold increases. Further, the estimated coefficients from the panel models shift from negative to positive. While the median temperature coefficients of the GLMs are similar to the panel models in that they are all negative, the middle

90% of the distribution is wider and overlapping zero. GLM temperature coefficients follow a similar pattern as the precipitation coefficients in that the range is narrower for mean flows, and more variable for the extremes.

In many ways the variation in temperature coefficients and statistical significance, as compared to the precipitation coefficients, is important for contextualising the landcover coefficients. We would expect the influence of temperature on streamflow to be more varied than precipitation and therefore less likely to be significant across all catchments because it is reasonable to expect temperature to have differing effects on streamflow depending on the other characteristics of the watersheds in question. These differences can be seen, to some extent, in the regional variation in GLM temperature coefficients (Figure 4.4).

For the most part, the climatological coefficients are not particularly surprising. We expect the changes in precipitation to be the most important driver of changes in streamflow magnitude in our models (Slater et al., 2021; Slater & Villarini, 2017). Contrary to our results, there is some evidence in the literature that lower flows may be more sensitive to changes in annual precipitation than are mean and high flows respectively (Allaire et al., 2015; Lins & Cohn, 2002), however, studies which explore this relationship at the annual timescale are limited. We might also expect changes in temperature, which affects streamflow by modifying evaporation, and may be serving as a proxy for snowfall, to be important for low flows which already occur in drier periods, but not for higher flows, when precipitation could be expected to be the primary influencer. Milly et al., (2018a) published a global dataset of regression based and theoretical precipitation and temperature sensitivities of the annual water balance between 1901-2013, the central summaries of which closely resemble our average single catchment and the urbanisation panel model coefficients for temperature (median regression and theoretical coefficients: ~ -0.02) and the coefficients of all models for precipitation (median regression and theoretical coefficients: ~ 1.6) (Milly et al., 2018b). The authors acknowledge that the models are subject to errors and that ignored processes may carry regional importance.

4.5.3   *Model context and assumptions*

We build on the question of comparability between methodological approaches posed in Salavati et al. (2016) by comparing single-catchment attribution approaches with the multi-catchment panel methods. These two general modelling approaches where selected due to their relative prevalence in the literature.

The GLMs used in our analysis are applied to timeseries data from individual sites. GLMs are models in which the response variable is expected to follow an exponential family distribution, in our case, the log normal distribution, and in which there is a linear relationship between the transformed response in terms of the link function and the explanatory variables.

The single site models are susceptible to issues caused by time series length, influential cases -- points in the time series which would significantly alter the regression coefficients if removed, and an inability to control for omitted variable bias. The specific formulation of our GLMs was selected so that the models would be approximately comparable to the panel regressions.

In a broad sense, panel regression models combine the data from the time series observations of multiple individuals, resulting in an increase in the degrees of freedom and model variability, and therefore improve the inference accuracy of model parameters (Hsiao, 2007). A key problem in single site regression on observed data is that of omitted variable bias which gives rise to endogeneity, meaning that a regressor is correlated with the error term. While endogeneity can arise from a number of sources, most important are omitted variable bias, error in the explanatory variable, and simultaneity, instances in which there is an explanatory variable that is jointly determined with the response variable (Croissant & Millo, 2018). Fixed effects panel regression models address the majority of omitted variable bias, enabling a causal interpretation, by requiring that confounding variables either be directly measured or be invariant along at least one dimension of the data, for instance, time (Nichols, 2007). This bias remains unaddressed in single site regression models. Additional sources of endogeneity, in particular, time-variant error in the explanatory variables, and omitted variable bias due to time-varying sub-regional variables may not be fully controlled for by the panel regression method either.

While we know that each catchment has unique and unmeasured characteristics which may affect streamflow, a panel regression approach allows us to explicitly define time variant attributes which affect multiple catchments, while minimizing the signal of catchment attributes which are constant in time, to isolate an average effect of a specific driver of change, in this instance, land cover change (Croissant and Millo, 2018). In essence, the panel regression approach does not consider every point in every time series as uniquely and specifically important, instead relying on inter-individual differences. It is therefore better equipped to formulate a robust average estimation by controlling for cross-sectional heterogeneity. This

robustness is exhibited by the relative consistency in the estimated effects of land cover changes across the sensitivity test thresholds, even when considering only sites where minimal land cover change occurred. That being said, pooling sites in this way may not be the best approach if we anticipate that land cover changes will affect streamflow in opposite ways (increasing or decreasing flow) depending on the specific attributes of the sites examined.

In contrast, the single-catchment regression models are more susceptible to bias arising from to the limited length of the time series, quality of the data used in fitting, and our inability to control for all confounding variables at individual sites. As such, in an average sense, the results can be misleading, especially if expected effects are marginal. On the other hand, the methods perform similarly when the effects are consistent, as is true for precipitation (Figure 4.7), and panel regression models require data from a large number of sites in order to perform well.

With longer time series, we might discover more land cover change and more certain associated streamflow responses, which may improve the performance of the single site models, however, it is also possible that because the effects are so small, the differences in catchment moderators may continue to render the land cover effects indistinguishable. For instance, if one is interested in a particular catchment for which a long, reliable, time series is available, the national or even regional average deliverable by a multi-catchment regression may not provide the most accurate description of the land cover-streamflow relationship. However, due to the dearth of long, consistent, timeseries, particularly for land cover, averaging across space and time in a multi-catchment regression may provide the best available estimates of the effects of land cover on streamflow.

4.6   Limitations

It is possible that there are other, unidentified, time-varying factors which have been omitted from our models, and which may bias the coefficients of the single site models. For instance, it is imaginable that water management practices, especially the presence of wastewater treatment facilities as they relate to urbanisation (Oudin et al., 2018), flood alleviation schemes, or other land cover changes may affect the influence of land cover change on streamflow to a high degree, particularly for low flows. Both the type and age of trees (Brown et al., 2005), and the land cover types which tree cover or urban area replace are likely to alter how these land cover types influence streamflow. The location of the land cover change within a catchment is also likely to influence how streamflow responds to land cover change, and fragmentation of urban area can have a key influence on flow response, particularly for low flows (Oudin et al.,

2018). We also do not consider the potential effects of allowing nested catchments to exist within the dataset, a factor which might positively bias the panel model significance if the same relationship exists as catchment size increases.

While the regional dummy variable in the urbanisation panel regression models controls for national and regional scale changes, it is possible that sub-regional trends that vary over time may be overlooked by this approach. Some examples of possibly influential omitted variables include antecedent moisture, and water management practices which may be specific to a city of subregion. Similarly, it is possible that average annual precipitation may not be the ideal metric for predicting low or high streamflow, so the inclusion of different precipitation quantiles could potentially affect the research outcomes. Removal of sites which are ephemeral at the annual timescale may have unintentionally excluded sites with strong signals from the analysis, particularly as there is some evidence that large-scale land cover changes can lead to 0 flows in formerly perennial rivers (Brown et al., 2013). A land cover dataset with a higher spatial resolution, or improved accuracy and longer time period might also result in different land cover classifications and results. Similarly, error in the land cover time series may vary year to year, which could potentially cause coefficient attenuation. While these data improvements could hypothetically increase confidence in the results, it is unlikely that the influence of small changes in land cover will be easily detected in the single site models

Future work should include more detailed consideration of confounding variables, as well as the numerous potential moderators which determine the effect of land cover changes on flow at different sites. Improved observational data, and consideration of other land cover types or classifications would likely also prove useful.

4.7   Conclusions

In order to effectively manage water resources and their associated risks it is important to understand where and to what extent streamflow distributions are being modified by anthropogenic drivers such as land cover change. Regression modelling approaches are popular for the purpose of attempting to attribute such changes to different drivers. Here, we use two statistical approaches to explore land cover changes in a large-sample dataset and compare the outcomes of the models. Using both single-catchment and multi-catchment (panel) regression methods, we address the following questions: 1. How do urbanisation and tree cover change associate with or affect streamflow across the conterminous United States; 2. How do the results of single-catchment and multi-catchment regression methods differ?

The panel regression models generally conform with our expectations regarding the direction of expected changes in streamflow in response to land cover change. According to our panel models, a one percentage point increase in catchment urban area leads to an average increase of approximately $0.6 – 0.7\%$ in both mean and high flows ($p < 0.0001$). Meanwhile, the effects of tree cover changes are generally not significant. The panel models indicate that there is not a statistically significant causal relationship between either land cover class and low flows.

Interestingly, the panel models also indicate that streamflow response to changes in annual precipitation is less certain (wider confidence intervals) and relatively smaller in catchments where urbanisation is also considered. We also demonstrate that at the annual timescale, low flows appear to be less sensitive to changes in precipitation than mean and high flows respectively.

The single-catchment GLM approach reveals an impressively wide range of coefficients, the median of which is largely close to zero in any case. However, due to the limitations of the data, and of the single catchment models for explaining bias, much of the variability is also related to confounding variables and omitted variable bias which are overlooked by the single-catchment models. The panel regression approach provides an immense increase in statistical robustness, and is capable of detecting the essentially marginal effects on flow consistently.

The panel regression approach may be inappropriate in instances where an average effect is not a helpful metric and at-site estimates are needed, or where we are unable to control for the influence of catchment specific attributes in a meaningful way. This is further characterized by the climatological coefficients, which are consistent and significant for precipitation in both approaches, but which vary more widely, and are frequently statistically insignificant for temperature. The systematic failure of the GLMs to detect a meaningful "average" association between either tree cover change or urbanisation, and streamflow, despite increasing the minimum land cover change threshold required for analysis, suggests that the relationship is, in an average sense, so small at the individual site level, that the effects of these land cover changes on flow are cancelled out by other intra-catchment processes. When using a longitudinal approach, as we do with the panel repression models, we are better able to tease out the small changes which are consistent over many individual catchments, over time.

This analysis therefore serves as a word of caution against the over-interpretation of single-catchment approaches as an optimal strategy for hydrologic attribution. Conversely, while panel regression approaches provide a more robust estimation of average land cover effects on

streamflow, the complicated nature of these relationships means that an average estimated effect may be not be a useful metric to capture the complexity of different climates and physical environments.

5 Elasticity curves describe streamflow sensitivity to precipitation across the entire flow distribution

5.1 Abstract

Streamflow elasticity is the ratio of the expected percentage change in streamflow for a 1% change in precipitation; a simple approximation of how responsive a river is to precipitation. Typically estimated for the annual average streamflow, we propose a new concept in which streamflow elasticity is estimated for multiple percentiles across the full range of the streamflow distribution in a large-sample context. This "elasticity curve" can then be used to develop a more complete depiction of how streamflow responds to climate. Representing elasticity as a curve which reflects the range of responses across the distribution of streamflow within a given time period, instead of as a single point estimate, provides a novel lens through which we can interpret hydrological behaviour. As an example, we calculate elasticity curves for 805 catchments in the United States and then cluster them according to their shape. This

results in three distinct elasticity curve types which characterize the streamflow-precipitation relationship at the annual and seasonal timescales. Through this, we demonstrate that elasticity estimated from the central summary of streamflow, e.g. the annual median, does not provide a complete picture of streamflow sensitivity. Further, we show that elasticity curve shape, i.e. the response of different flow percentiles relative to one another in one catchment, can be interpreted separately from between-catchment variation in the average magnitude of streamflow change associated with a one percent change in precipitation. Finally, we find that available water storage is likely the key control which determines curve shape.

5.2 Introduction

The relationship between streamflow and meteorological variables such as precipitation, temperature, and evaporation are often represented simplistically and may be poorly understood through modelling experiments. Analyses based on observations can provide better insight into assumed physical relationships. One data-based approach for quantifying the relationship between streamflow and precipitation, and for estimating future changes in streamflow, is the concept of "elasticity". Streamflow elasticity describes the sensitivity of streamflow to changes in any given climatic variable (relative to the long-term mean of the time series) and is defined most frequently as the percentage change expected in the annual water balance or mean annual streamflow which results from a one percent change in a variable of interest, typically precipitation (Schaake, 1990).

Streamflow elasticity to precipitation, as estimated for average flows, has been reported on extensively at the annual timescale (Berghuijs et al., 2017; Chiew, 2006; Chiew et al., 2006; Milly et al., 2018b; Sankarasubramanian et al., 2001; Tang et al., 2020; Tsai, 2017), and more recently, at aggregated multi-annual scales (Y. Zhang et al., 2022). At seasonal to annual timescales, streamflow magnitude represents the aggregated components of precipitation, evapotranspiration, and storage, including antecedent moisture conditions and water use. Thus, a one percent change in precipitation is unlikely to result in a one percent change in streamflow. Instead, changes in precipitation tend to be amplified in streamflow, and elasticity estimates are typically greater than one. Reported values range between 0.75 and 2 depending on the region and methodology (Allaire et al., 2015; Sankarasubramanian et al., 2001; Tsai, 2017) and may differ for increases vs decreases in precipitation. For instance, average streamflow in arid regions tends to be more sensitive to precipitation decreases than increases (Tang et al., 2019). Additionally, in some cases, elasticity has been quantified for low flows (Bassiouni et al., 2016;

Kormos et al., 2016; Tsai, 2017) and high flows individually (Brunner et al., 2021; Prudhomme et al., 2013; Slater & Villarini, 2016c).

Few studies, however, have quantified the elasticity of different segments of the flow distribution within the same catchment simultaneously. Harman et al. (2011) examined the elasticities of the slow and quick flow components of the annual hydrograph, approximately equivalent to low and high streamflow, and the total annual discharge in catchments in the United States using an analytical-functional water balance modelling approach. They found that quick flow frequently experienced much higher elasticities relative to total discharge or slow flow, respectively. Further, they showed that the elasticities of the slow flow component were highly variable between catchments, while the elasticity of quick flow was relatively consistent across sites, and the variability of total flow fell somewhere in between (Harman et al., 2011). Anderson et al. (2022) found a similar pattern using a different approach, also in the United States.

The dominant sources of streamflow are dependent on the segment of the hydrograph which is considered. For instance, low flows or base flows in natural rivers are typically the result of inflow from catchment storage sources, such as groundwater, lakes, or wetlands (Smakhtin, 2001). Meanwhile, high streamflow magnitudes are controlled, in large part, by precipitation events and antecedent soil moisture conditions (Ivancic & Shaw, 2015; Slater & Villarini, 2016c). Thus, it stands to reason that different percentiles of streamflow at both the annual and seasonal timescales will experience different elasticities to precipitation change. The variations in streamflow sensitivity to precipitation at different flow percentiles evident in Anderson et al. (2022) and Harman et al. (2011), when considered relative to one another in the same catchment and in aggregate, may provide new information, or a new lens for interpreting information about how rivers might react to climate changes. This is especially relevant for lower streamflow, as hydrologic behaviour has been shown to have a lower degree of regional similarity for low flows when compared to higher streamflow percentiles because local geographic conditions have greater influence over low flow regimes (Patil & Stieglitz, 2011).

Streamflow elasticity is the ratio of the expected percentage change in streamflow to a 1 % change in precipitation – a simple approximation of how responsive a river is to precipitation. Typically, streamflow elasticity is estimated for average annual streamflow; however, we propose a new concept in which streamflow elasticity is estimated for multiple percentiles across the full distribution of streamflow. This "elasticity curve" can then be used to develop a

more complete depiction of how streamflow responds to climate. Representing elasticity as a curve which reflects the range of responses across the distribution of streamflow within a given time period, instead of as a single-point estimate, provides a novel lens through which we can interpret hydrological behaviour. As an example, we calculate elasticity curves for 805 catchments in the United States and then cluster them according to their shape. This results in three distinct elasticity curve types which characterize the streamflow–precipitation relationship at annual and seasonal timescales. Through this, we demonstrate that elasticity estimated from the central summary of streamflow, e.g. the annual median, does not provide a complete picture of streamflow sensitivity. Further, we show that elasticity curve shape, i.e. the response of different flow percentiles relative to one another in one catchment, can be interpreted separately from between-catchment variation in the average magnitude of streamflow change associated with a 1 % change in precipitation. Finally, we find that available water storage is likely the key control which determines curve shape.

We propose the use of a new concept, the "elasticity curve", as a means to interpret hydrological responses to precipitation across many segments of the flow distribution simultaneously (Figure 5.1 A). This new approach allows for the visualization and comparison of the varied responses of streamflow across the flow distribution to precipitation changes at the annual and seasonal timescales. The main principle being that the response of streamflow to a shift in total precipitation across the period of interest will differ for higher streamflow percentiles, which result from more immediate responses, than for low flows, which are typically driven by storage in drier periods. We expect that hydrological catchments which have greater storage availability will be better able to sustain low flows, resulting in flatter elasticity curves, as opposed to those with lower storage capacity. Elasticity curves are generated by estimating elasticity for a series of discrete percentiles of streamflow. The combination of these discrete point estimates then forms a curve which represents the variation in streamflow sensitivity to climate across the annual and seasonal streamflow distributions (Figure 5.1 B).

We generate streamflow elasticity to precipitation curves ($\varepsilon_{c,P}$) for 805 rivers in the United States using statistical modelling and clustering approaches. We address the following questions:

Does $\varepsilon_{c,P}$ shape vary systematically and predictably across catchments?

What catchment attributes best explain between-catchment variation in $\varepsilon_{c,P}$ shape?
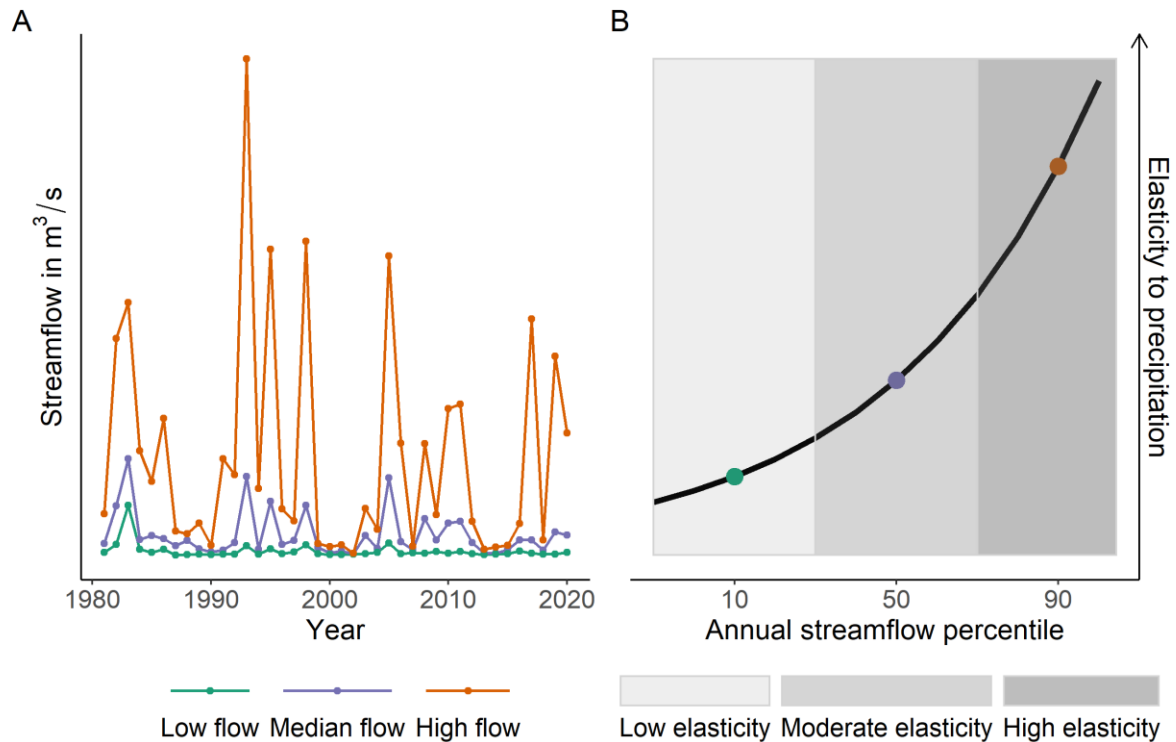
Figure 5.1 Conceptual diagram demonstrating how to read an elasticity curve.Where plot panel A. shows hypothetical high, low, and median annual streamflow (10th, 50th and 90th percentiles of the flow distribution in each year) and plot panel B. shows the hypothesised relative elasticity of each of these streamflow percentiles to changes in annual precipitation. For simplicity, this diagram shows only 3 points, but a typical curve in this study would normally include 21 points (one for every 5th percentile from 0-100 inclusive). Note: In practice, elasticity curve shape may vary from this simplified example, and a monotonically increasing line is not necessary.

5.3   Methods

5.3.1   Data

We estimate the elasticity of streamflow to changes in precipitation at every 5th percentile of annual and seasonal flow in 805 perennial U.S. rivers. This sample of catchments was selected from the Geospatial Attributes of Gages for Evaluating Streamflow version II (GAGES II) data set, having met the following criteria. All catchments were required to have less than 1 day of upstream dam storage (Anderson et al., 2022; Blum et al., 2020; Hodgkins et al., 2019), calculated by dividing the total upstream dam storage by the estimated catchment annual runoff (Falcone, 2017), as evidence that they were minimally influenced by dam storage. Additionally, all catchments had at least 30 years of 95% complete, consecutive daily

streamflow data between 1981 and 2022. Finally, we removed all ephemeral rivers and streams, defined as streamflow records having any 0 flow days. Gages II was used because the dataset provides geospatial data for a large number of catchments in the United States, facilitating analysis.

Catchment attributes, including total upstream dam storage, average annual runoff and watershed boundaries were taken from the same source (Falcone, 2017). The daily streamflow time series for the period 1981–2020 were taken from the USGS using the *R* package dataRetrieval (DeCicco et al., 2024). Gridded monthly precipitation and temperature (4 km resolution) were extracted from the Oregon State PRISM project using the R package prism (Edmund & Bell, 2015). We estimated average daily precipitation (mm/day) annually and seasonally within the upstream drainage area (watershed boundary) of each gaging station. We calculated average daily PET (mm/day) for each timescale in R using the Hamon equation (Hamon, 1963a; Lu et al., 2007a) with monthly temperature as previously described and estimated solar radiation from latitude and Julian date. While GAGES II (Falcone, 2017) includes PET estimates, also calculated using the Hamon equation, we recalculated these because the existing dataset did not cover our desired time period. The Hamon equation was used to retain consistency with the GAGES II data set and because this method has been shown to perform well relative to other approaches, despite its simplicity (Lu et al., 2007a). Annual values were calculated for water years (defined here as September to August), and seasonal values were estimated for winter (December, January, February), spring (March, April, May), summer (June, July, August) and fall (September, October, November) within each water year.

5.3.2 Single catchment models

Historically, streamflow elasticity has been estimated using a reference approach as proposed initially by Schaake (1990) and further developed into a nonparametric estimator by Sankarasubramanian et al. (2001), in which elasticity is expressed as the median of the ratio of the annual streamflow anomaly to precipitation anomaly, relative to the long term mean. Many recent studies have instead relied on the coefficients from multivariate regression models, such as generalised and ordinary least squares regression (Andréassian et al., 2016; Potter et al., 2011), or regionally-constructed panel regression models (Bassiouni et al., 2016), to estimate elasticity. These types of approaches are often functionally equivalent (Cooper et al., 2018) to the reference approaches. The benefits of regression-based approaches include simultaneous estimation of sensitivity to potential evaporation and precipitation, accounting for co-variation in these phenomena and providing a more robust estimate of elasticity (Andréassian et al.,

2016).   Probabilistic statistical tools also enable straightforward calculation of confidence intervals and panel regression models, like those included in the appendix of this paper (Appendix 3), are capable of controlling for a large portion of omitted variable bias, allowing for a more causal interpretation of regression results (Croissant & Millo, 2018; Hsiao, 2007; Nichols, 2007). These have been shown to produce more reliable elasticity estimates than single catchment models, when the expected effect is relatively uncertain (Anderson et al., 2022; Bassiouni et al., 2016), although their application for the explicit estimation of elasticity thus far, is limited.

In the first instance, we fit simple log linear models (lm) using the ordinary least squares estimator, to every 5th percentile of the annual and seasonal flow regimes from the minimum streamflow magnitude (Q_0) to the maximum (Q_{100}) for each historical streamflow record (Equation 1).

$$ln\left(Q_{i,t}^q\right) = \ \alpha_{i,t} + \varepsilon_P^q \ln\left(P_{i,t}\right) + \varepsilon_E^q \ln\left(E_{i,t}\right) + \eta_{i,t}^q \qquad\qquad 5.1$$

where $\ln\left(Q_{i,t}^q\right)$ is the natural logarithm of a streamflow percentile (q) calculated for time period (t) for catchment (i), $\alpha_{i,t}$, is the intercept, $\ln\left(P_{i,t}\right)$ is the natural logarithm of catchment averaged annual or seasonal mean of daily precipitation, and $\ln\left(E_{i,t}\right)$ is the natural logarithm of catchment averaged annual or seasonal mean of daily potential evaporation in that period. Note that mean seasonal or annual climate time series (*P* and *E*) are used, not percentiles equivalent to the streamflow percentile of interest (denoted with the superscript "q"). In other words, while $Q_{i,t}^q$ refers to a different percentile of annual or seasonal streamflow ranging from 0-100 in each iteration of the model, $P_{i,t}$ and $E_{i,t}$ refer to the annual or seasonal average in all iterations. The point estimate of precipitation elasticity is represented by the regression coefficient: ε_p^q and potential evaporation elasticity is represented by ε_E^q. The error term is $\eta_{i,t}^q$.

The elasticity curve $\varepsilon_{c,P}$ is simply the combination of the percentile specific point estimates of elasticity (ε_P^q). For visualization purposes, we linearly interpolate between the points. As presented in this study, the elasticity curve characterises the sensitivity of different percentiles of annual and seasonal streamflow to changes in the average annual or seasonal precipitation. For example, an elasticity of 0.5 for the 15th percentile of annual streamflow would indicate that a 1% change in the overall mean annual precipitation would correspond to a 0.5% change in the 15th percentile of annual flow.

Understanding the shape of the elasticity curve is important in order to assess the responsiveness of different streamflow percentiles to changes in precipitation within a given catchment area. We do not explicitly try to explain spatial variation in actual magnitude of elasticity in this work because this has been done extensively in other literature. We aim, instead, to identify catchments with a similar elasticity behaviour across streamflow quantiles, and therefore seek to cluster the curves based on their shape, rather than the magnitude of the elasticity estimates. To achieve this, we normalize the curves relative to the elasticity of the minimum streamflow at each timescale, by subtracting ε_P^0 from each of the ε_P^q estimates.

We then use Ward's minimum variance method (Ward, 1963) for agglomerative hierarchical clustering in R to group the complete elasticity curves for the individual catchments into clusters with similar shapes. Hierarchical clustering methods were chosen because the results are reproducible and not influenced by initialisation and local minima (Murtagh & Contreras, 2012). We used the Euclidean distance measure for clustering, and Ward's algorithm was selected because it had the highest agglomerative coefficient as compared to the complete linkage, single linkage, and UPGMA algorithms, indicating stronger clustering structure.

The number of clusters for each temporal scale was selected through visual inspection of the dendrograms, silhouette plots, and the gap statistics. We additionally performed a sensitivity analysis in which we fit 2, 3, 4, and 5 clusters to the data and examined the spatial distribution of the prospective clusters. This resulted in the selection of 3 clusters for the annual, winter, and summer timescales and 2 clusters each for the spring and fall timescales. We then determined cluster type based on the difference between the average elasticity of the minimum and maximum flow in a given period. The number of clusters were chosen so that the fewest clusters possible would be selected for each temporal scale while still capturing the general shapes of the $\varepsilon_{c,P}$s. In spring and fall additional clusters did not result in a more informative classification.

In addition to these models, a panel regression approach was applied to help validate the results. This model and its results are included in Appendices A and B.

5.3.3 Attribution of elasticity curve classification

Finally, we are concerned with the drivers behind variability in elasticity curve shape. Therefore, we consider explanatory variables which have previously been shown to be related to between-catchment variation in the magnitude of elasticity as well as additional hydrologic signatures related to streamflow sensitivity. These variables, presented in Table B1, include:

the slope of the flow duration curve calculated for low flows (lowest third - fdc_{bl}), average flows (middle third – fdc_b), and high flows (highest third fdc_{bu}), runoff coefficient (RC), average annual temperature, aridity index, mean elevation, slope, drainage area, snow fraction, and average permeability and latitude (Falcone, 2017). We additionally consider the baseflow index (BFI) calculated over a time window of five days, and a longer "delayed flow index" (DFI) calculated over a time window of 90 days as in Gnann et al., (2021). Our intention here is to capture baseflow from different sources – BFI aims to separate event from inter-event flow and DFI aims at separating seasonal variation from inter-annual baseflow (Gnann et al., 2021; Stoelzle et al., 2020). DFI has been previously shown to be much more clearly related to geology as compared to BFI. The full equations and specifications for the explanatory terms are included in Table B1. Finally, we consider six categorical seasonality variables: most important precipitation season (winter, spring, summer, fall), calculated as the season in which the largest precipitation amount falls, least important precipitation season, calculated as the season in which the least amount of precipitation falls, low flow season and high flow season. Further, we include combinations of most important precipitation season and low flow season, as well as least important precipitation season and low flow season (ex. winter_summer, in the instance that winter is the most important precipitation season and summer is the most important flow season). These final two seasonality metrics are intended to shed light on whether streamflow is in phase with precipitation.

To attribute the drivers of between-catchment variation in elasticity curve shape and determine the predictability of elasticity curve cluster membership, we use random forest classification models to estimate relative variable importance for the prediction of cluster membership at each temporal scale. The clusters are frequently imbalanced in terms of the number of sites in each group, so we train the model on a sub-sample of the data set which consists of 80% of the sites in the smallest cluster and equivalent quantities of each additional cluster randomly selected from the complete data set. We then test the model performance using a sample which consists of the remaining 20% of the smallest cluster, and quantitatively equivalent samples of each additional cluster. We repeat the random sampling and model fitting process 10 times per temporal scale and then calculate the average actual accuracy across 10 iterations.

5.3.4 *Example catchments*

Elasticity curves computed at individual sites typically have wide confidence intervals and should be applied cautiously, but we select three sites which may serve as an example of the elasticity curve concept, and put the limitations of the approach in context. The three

catchments provide a detailed example of the approach and mechanistic insights. These example catchments are: Turnback Creek above Greenfield (gauge id: 06918460), Current River at Van Buren (gauge id: 07067000), and Reddies River at North Wilkesboro (gauge id: 02111500), examples which coincide with Gnann et al. (2021) who proposed a framework for incorporating regional knowledge into large sample hydrology when studying baseflow processes and drivers. They include detailed examples of the processes controlling baseflow and delayed flow partitioning in catchments in different regions of the U.S., some of which happened to be included in our analysis. These example catchments were selected due to the availability of information for comparison, and because two of them are located near one another but have differing physiographic profiles, while a third is physically distant but has similar baseflow metrics. These relationships allow for comparison of the elasticity curves for each site.

5.4 Results

5.4.1 Normalized elasticity curves

Figure 5.2 shows the average normalized elasticity curves for each temporal scale (annual and seasonal). The normalized curves have been clustered so that catchments with similar curve shapes are in the same group. The curves were produced using linear regression models fit to each catchment individually (Equation 1), then the normalized values were averaged within each cluster and plotted with the interquartile range of the respective ε_p^q values. We use the interquartile range because the lms result in a distribution of ε_p^q values for each streamflow percentile (one per stream gauge) and the resultant curve is an average of all sites in a cluster.
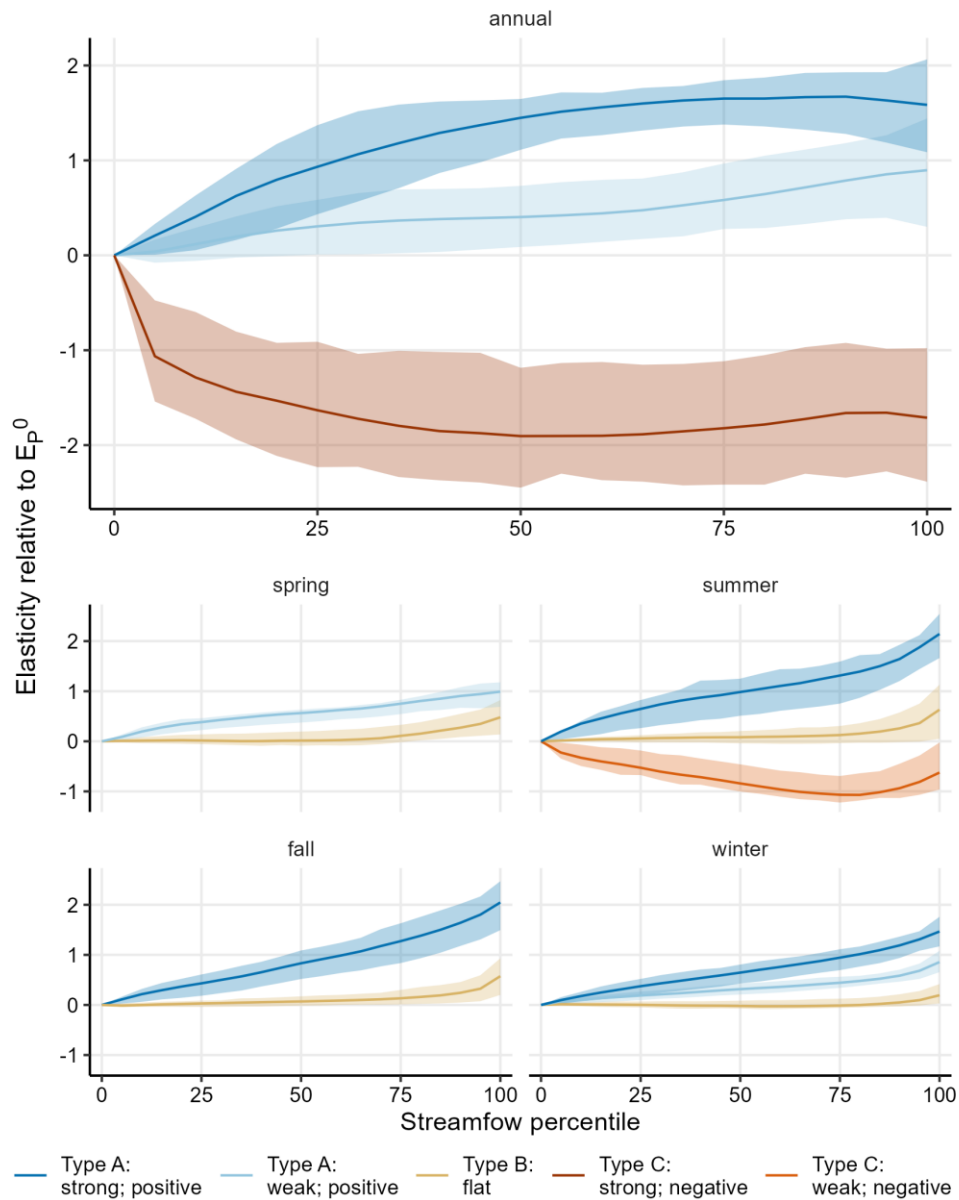
Figure 5.2 Normalized elasticity curves

shows the curves resulting from the single catchment linear models (lm) where the line is the mean of the distribution of elasticity point estimates (for a cluster of sites) and the bands are the inter-quartile range. Note that spring and fall have 2 clusters while winter, summer, and annual have 3 and that seasonal streamflow percentiles represent subsets of the annual flow.

We find three main curve types which we define as: curve type A - where the cluster average curve is positively sloping and the difference between ε_P^0 and the largest point estimate in the average curve is greater than 0.75 percentage points; curve type B - where the cluster average curve is relatively flat and the absolute difference between these points falls between -0.75 and 0.75 percentage points; and curve type C - where the cluster average curve is negatively sloping

and the difference between ε_P^0 and the largest point estimate of the average curve is less than -0.75 percentage points. We further define two sub-types of curve types A and C: "strong" with greater than a 1.25 percentage point difference between ε_P^0 and the largest point estimate and "weak" (0.75 - 1.25 percentage points). This division is merely a heuristic for separating the clusters. Some individual catchments within each type class have total absolute differences in elasticity estimates which do not comply with this division.

At the annual timescale, 91% of catchments exhibited type A curves, demonstrating that in an overwhelming majority of cases larger streamflow quantiles are proportionally more responsive to precipitation. Of these, 31% (251 catchments) were grouped into a single class for which the average $\varepsilon_{c,P}$ has a strongly positive slope (curve type A: strong), and 60% of catchments (495) were clustered into a weakly positive class (type A: weak). In catchments with curve type A, where $\varepsilon_{c,P}$ has a positive slope, higher streamflow percentiles are increasingly more responsive to a one percent change in precipitation than are low flows. Some catchments, predominantly in the eastern portion of the country, exhibit different behaviour. 7% of catchments (58 catchments) were clustered into a group with strongly negative $\varepsilon_{c,P}$ (curve type C: strong). A negatively sloping elasticity curve shape indicates that high flows are relatively less responsive to precipitation variation than are lower flows. In other words, variation in precipitation predominantly effects the hydrologic response of larger streamflow percentiles for catchments with a positively sloping $\varepsilon_{c,P}$, and lower streamflow percentiles in catchments with negatively sloping $\varepsilon_{c,P}$.

In winter, fall, and spring, none of the cluster-average elasticity curves are negatively sloping. 31% of catchments (246) in the fall, 26% (211) in winter, and 65% (524) in spring are grouped into a cluster for which $\varepsilon_{c,P}$ can be described as relatively flat (curve type B), defined here as having a range of normalized ε_P^q values between -0.75 and positive 0.75. In winter, catchments with curve type B are mostly concentrated at high latitudes and mountainous regions, while in the fall, these catchments are geographically more widespread (Figure 5.3 C), existing both in the north, the southwest, and to some extent, throughout the gulf coast. A flat elasticity curve denotes a catchment in which the responsiveness of streamflow to changes in precipitation is consistent across the distribution. The remaining clusters are positively sloping curves. Similarly, 78% (626) of catchments in the summer season are curve type B. Meanwhile 111 catchments (~14%) are curve type A (strongly positive), and a cluster with the remaining 8% (68) of catchments is generally negatively sloping (type C: weak). Finally, 281 catchments

(~35%) in spring are weakly positive (type A: weak). In spring, the absolute difference between the cluster-specific ε_P^0 and ε_P^{100} across all curves, is small, not exceeding one percentage point on average for any group.

Elasticity curve shape and the actual magnitude of expected streamflow change in response to a one percent change in precipitation do not necessarily correspond (Figure 5.3). For instance, in the summer, 78% of catchments exhibit a flat elasticity curve (Figure 5.3 summer: A; C). However, while skewed towards zero, the distribution of possible elasticity magnitude is widespread (Figure 5.3 summer: B), indicating that the streamflow response to a one percent change in precipitation in this group ranges from between about zero to two percent. Conversely, the distributions of magnitude for flat elasticity curves in winter is concentrated around zero, indicating that streamflow across the majority of catchments has a very low responsiveness to precipitation variation in this season. In other words, a flat elasticity curve indicates that low and high flows have approximately the same response to precipitation changes within a particular catchment, but that the response is not necessarily small or consistent across catchments with the same elasticity curve shape. The highest actual elasticity values are predominantly in the eastern U.S. in all seasons. High magnitude elasticity values also occur in the Pacific Northwest especially in the fall, winter, and summer seasons.

It is worth noting that the distribution of streamflow in each season represents a subset of the streamflow in a year. For example, the streamflow magnitude which corresponds to high flows in the winter season may be equivalent to average or lower streamflow at the annual timescale.
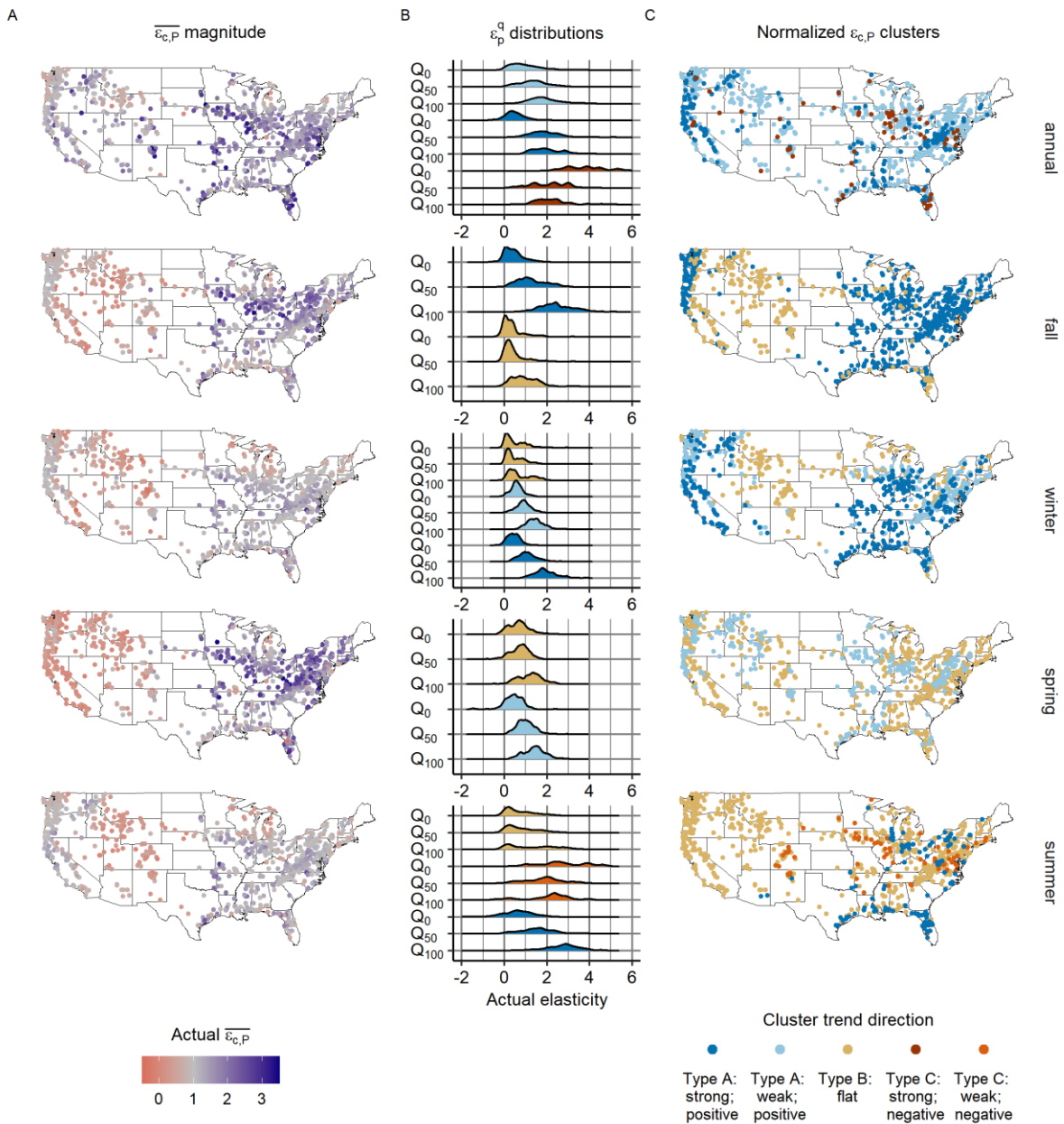
Figure 5.3 Actual elasticity compared to normalized elasticity curves.

Panel A shows the geographic distribution of the means of the non-normalized, site specific, elasticity curves. These values are referred to as actual elasticity in the text. Smaller mean elasticity values (less responsive) are highlighted in lighter shades and higher mean elasticity values in darker shades. Panel B shows the distributions for non-normalized point estimates of elasticity at the lowest, median, and highest streamflow (Q0, Q50, Q100) in each time period (annual, winter, spring, summer, fall). The distributions in Panel B are coloured according to the cluster membership of the normalized curves (Figure 2), the geographic distribution of which is shown in Panel C.

5.4.2 Attribution and predictability of between-catchment variation in streamflow elasticity

We conduct a multivariate variable importance analysis using random forest models to determine the extent to which catchment attributes are able to predict elasticity curve shape. The following catchment characteristics are included in this analysis: Aridity index, DFI, BFI, Slope of the flow duration curve (calculated at the 0-33rd, 33-66, and 67-100th percentiles), latitude, coefficient of variation for daily streamflow in each season, mean annual temperature, mean catchment elevation, drainage area, mean catchment slope, and snow fraction, as well as, precipitation and streamflow seasonality and timing metrics (Table B1). Averaged over 10 iterations each, the random forest model accurately predicted class membership in approximately 70% of cases at the annual timescale, 95% for fall, 79% for winter, 63% for spring, and 79% for summer, all rounded to the nearest integer.

For each temporal scale, different variables were selected as the best predictors of cluster membership using both the Gini coefficient and the mean decrease accuracy metric. For both the annual and summer periods, fdc_{bl} was the best predictor for every iteration of the random forest model. At the annual timescale, DFI, fdc_b and aridity are the second and third best predictors of cluster membership depending on the model run. The second and third best predictors for summer class membership vary between iterations. In winter, the best predictors for both metrics were either average annual temperature, or the time delay between the least important precipitation season and low streamflow season. In addition to these metrics, mean catchment elevation and other seasonality metrics were frequently selected as the second or third most important predictors for winter depending on the model run. For fall, the time delay between the least important precipitation season and low streamflow season, mean catchment elevation, and BFI were the top three predictors in the majority of iterations of the model for both metrics and typically had very similar mean decrease accuracy scores and Gini coefficients. No variable was clearly the best predictor of cluster membership in springtime, as over the course of 10 model runs, eight different variables had the highest Gini coefficient or mean decrease accuracy score.

5.5 Discussion

In this paper, we use multivariate statistical models to investigate whether streamflow elasticity to precipitation varies across the distribution of streamflow at the annual and seasonal timescales. We then use a clustering algorithm and random forest regression model to examine the extent to which that variation is systematic and predictable.

By creating elasticity curves, which represent the range of elasticity across the streamflow distribution (Figure 5.2), we show that at the annual and seasonal timescales, the highest streamflow percentiles are typically more responsive to long-term precipitation change relative to lower streamflow percentiles in the same catchment and time period. This is especially true for annual elasticity and in the spring, winter, and fall. The finding that low flows are less responsive to precipitation change than higher flows is in line with existing literature. Low flows are typically sustained by groundwater, saturated soils, and surface water storage which require precipitation for recharge, but for which the effects of changes in precipitation are inherently delayed and moderated (Gnann et al., 2021; Price, 2011; Smakhtin, 2001).

There are, however, catchments which do not have positively sloping elasticity curves at some timescales. Approximately 7% of catchments at the annual timescale and 8% in summer are clustered into groups with generally negative trends, indicating that low flows are relatively more responsive to precipitation than are higher streamflow percentiles. Further, the elasticity curves of roughly 31% of catchments in fall, 78% in summer, 65% in spring, and 26% in winter are nearly flat, having very low slopes for the majority of the curve, with ε_P^q estimates only increasing marginally for the highest streamflow percentiles.

The best predictors of elasticity curve shape are those related to the hydrologic storage capacity of the catchments. For instance, fdc_{bl}, the most important catchment attribute at the annual timescale and in summer, provides information about a catchment's ability to sustain flows of a certain magnitude during the dry season. The flow duration curve (fdc), here calculated using daily streamflow for the entire study period, is a cumulative frequency curve which shows the percentage of time that a certain magnitude of streamflow is equalled or exceeded (Searcy, 1959). When the slope of the fdc is steep, it indicates that a catchment has highly variable streamflow predominantly originating from direct runoff, and when the slope is relatively flat, it suggests the presence of surface or groundwater storage, which equalises flow. At the low end of the fdc (here fdc_{bl}), a flat slope points to the presence of long-term storage within the catchment, while a steep slope indicates that very little exists (Searcy, 1959). Similarly, baseflow is the portion of streamflow that is derived from groundwater and other delayed sources (Smakhtin, 2001), and a low BFI indicates a catchment for which a majority of streamflow comes from direct runoff. We have defined two baseflow metrics, BFI and DFI, a delayed flow metric over a longer time span (Gnann et al., 2021; Stoelzle et al., 2020), both of which are frequently important predictors of elasticity curve shape. Further, while snow fraction was not necessarily the most important predictor in cold months, temperature, latitude,

elevation, and the time gap between the most important precipitation and streamflow season, attributes which speak to precipitation type and snow dominance, were.

Storage components consist of anything ranging from surface waterbodies such as wetlands, to snow cover, and ground water influxes, all of which interact with fluvial systems on different timescales. Catchments with relatively flat elasticity curves in cold months (winter and fall), are typically those at high latitudes which receive higher percentages of precipitation as snow, or those in the semi-arid southwestern region which are predominantly fed by snow melt upstream (Li et al., 2017). These curves are flat and have actual elasticity estimates which are heavily skewed towards zero (Figure 5.3 winter: A; B; fall: A; B) because snow melt does not usually occur in winter or fall. However, at the annual timescale, the same catchments have actual elasticity values ranging from less than 1 for low flows to around 2 for the highest annual flows because the streamflow response is delayed, but occurs within the same year. In the fall, there are additionally catchments in Florida and scattered along the southern coast with relatively flat elasticity curves, potentially due to increased storage within the catchment area e.g. as wetlands. The seasonal elasticity estimates specifically capture the influence of in-season precipitation on streamflow (i.e. within that same season). Streamflow in many rivers is driven by out-of-season precipitation. Thus, while flat seasonal elasticity curves and low percentile-specific point estimates indicate a muted hydrologic response, they do not rule out the possibility that the timescale for response is merely longer than that which is considered. Further, as noted previously, Seasonal flow percentiles represent subsamples of annual flow. These may or may not directly correspond to the same section of the flow distribution. For instance, the 50th percentile of summer flow may relate to a much lower or higher annual flow percentile, depending on the temporal distribution of flow in the year.

Flat elasticity curves are present across the majority of the country during the summer (Figure 5.2 summer; Figure 5.3 C), indicating that the response of streamflow to summer precipitation is similar across all flow percentiles in these catchments. Similar to winter and fall, the flat elasticity curves tend to have higher BFI and DFI and lower fdc_{bl} values than type A or C curves. Many of these catchments have average actual elasticity values which approximate 0, indicating that in-season precipitation has little to no influence on seasonal streamflow, however, others have larger average actual elasticity values, often greater than one (Figure 5.3 summer: A; B), which in turn, implies summer precipitation has a substantial influence on summer streamflow, but that the influence is consistent across the distribution. This differs

from a majority of cases in other seasons and at the annual timescale, for which the influence of precipitation on streamflow is magnified in higher streamflow percentiles.

Evidence suggests that high flow magnitudes are driven by the combined influences of precipitation events and antecedent soil moisture (Ivancic & Shaw, 2015; Slater & Villarini, 2016a). Summer is a period of relative soil moisture deficit (Koehn et al., 2021) and high potential evaporation. It is plausible therefore that the non-zero magnitude flat elasticity curves in the majority of the study region during this period are emblematic of the relationship between antecedent wetness, precipitation, and streamflow. In other words, because of a soil moisture deficit, the precipitation changes are not typically magnified in higher streamflow percentiles in the majority of catchments (78%) during this period, especially in catchments where sources of delayed flow (e.g. groundwater) are large contributors across the flow distribution (Berghuijs & Slater, 2023).

This does not, however, explain the relative homogenisation of the elasticity curve structure in the spring, a period in which soil moisture recharge is likely to occur. Instead, it seems probable that the flatness of the elasticity curve shape, despite a persistently broad range of elasticity magnitudes in spring (Figure 5.3 spring: B), may be due to the fact that streamflow is the least variable on average in springtime compared to the other seasons, as determined by the coefficient of variation (CV) of the daily streamflow measurements, and that springtime is the low flow season in only 24 catchments. In other words, the lowest flows in spring may be more heavily driven by runoff from precipitation rather than storage as compared to other seasons. This hypothesis is further supported by the cluster-specific CV distributions at other timescales – where type B elasticity curves correspond to catchments with relatively low variability (Figure 5.4 spring). The shape may also reflect, in part, the climatic drivers dominant over different regions.
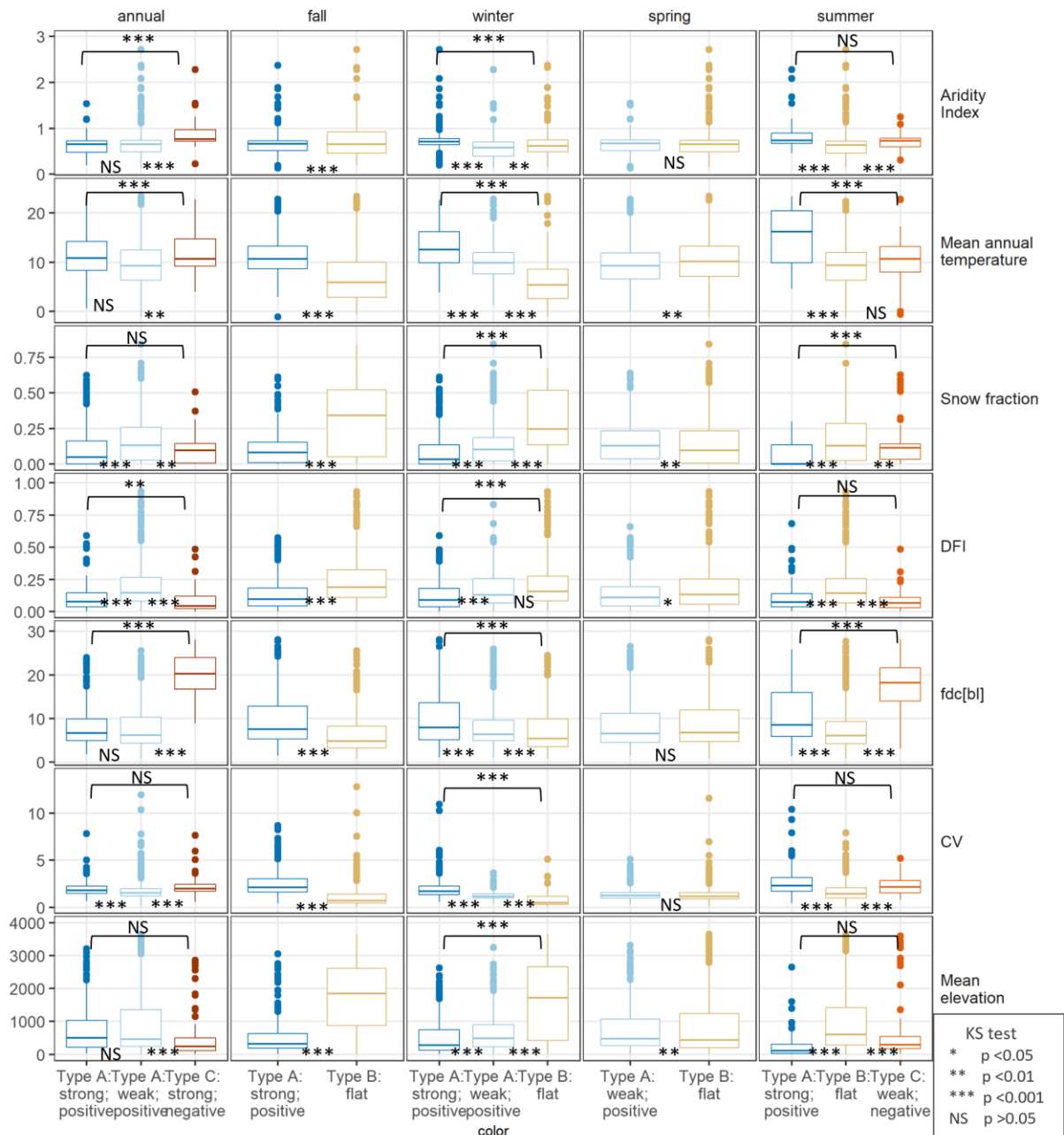
Figure 5.4 Boxplots showing the distributions of static catchment attributes split by time period and cluster membership.

Significance is shown between each box for neighbouring distribution plots and the significance of the difference between the first and last distribution in each time period is plotted at the top of the panel. Boxplots are included only for attributes which are important in the RF analysis, and which can be represented by continuous numeric values, so seasonality metrics are excluded here.

The range of type B elasticity curves which is present across the seasons is washed out at the annual scale, demonstrating that the catchment storage which leads to a uniform response

across the distribution of streamflow generally operates at a timescale of less than a year (Figure 5.2). Type A elasticity curves with a strong signal exist across temporal scales, in catchments which have relatively low BFI and DFI and steep middle sections of the flow duration curve, fdc_b, as compared to type B and weak type A signals (Figure 5.4). Interestingly, at the annual timescale, curve type C (negative) catchments are in some ways similar to those with strong curve type A (positive) signals, in that they both have low snow fraction, low BFI, and steep fdc_b slopes. They differ, however, in a number of other attributes, most notably, DFI and slope of the low end of the flow duration curve, fdc_{bl}. This difference indicates that while streamflow in catchments exhibiting both types of curves is predominantly rain-fed, those exhibiting strong type A curves are better able to sustain low flows as compared to type C catchments. Catchments with type C curves have very flashy low flow behaviour. We controlled for ephemeral streams in this study in order to simplify our methodology, but including those catchments may increase the prevalence of type C curves. The type C elasticity curves have wide interquartile ranges and wide confidence intervals when estimated with a panel regression model (Figure B1), indicating lower robustness in the estimation of this group overall (Figure 5.2). The strong type C cluster at the annual timescale also exhibits a positive slope above the 35th percentile of streamflow. While speculative, these results suggest that type C curves may differ from positive $\varepsilon_{c,P}$s predominantly in that they exhibit highly flashy low flow behaviour (Figure 5.4).

5.5.1   Example catchments and limitations

In order to contextualize the approach at individual locations, we examine the elasticity curves of three streamflow gauges. The non-normalized elasticity curves for Turnback Creek above Greenfield (gauge id: 06918460), Current River at Van Buren (gauge id: 07067000), and Reddies River at North Wilkesboro (gauge id: 02111500) are included in Figure 5.5. Despite being located near one another, gauge 07067000 lies over the Ozark aquifer, a more mature karstic environment, with comparatively more long-term storage and higher DFI (0.4) and BFI (0.7) as compared to gauge 06918460 (DFI: 0.1; BFI: 0.5) (Gnann et al., 2021). Conversely, gauge 02111500 is physically distant from the other two catchments and has a different geological profile (Zimmer & Gannon, 2018), but has substantial seasonal and stable storage components resulting in high DFI (0.4) and BFI (0.7) values compared to both of the Ozarks catchments. Catchment attributes for each of these sites are presented in Table 5.1.

Table 5.1 Attributes of example catchments:

Turnback Creek above Greenfield (gauge id: 06918460), Current River at Van Buren (gauge id: 07067000), and Reddies River at North Wilkesboro (gauge id: 02111500). Definitions of attributes are included in table B1. Max and Min P season are the most and least important precipitation seasons respectively, and Max and Min Q season are the most and least important flow seasons respectively.

STAID	BFI	DFI	fdc_b	fdc_{bl}	fdc_{bu}	RC	Aridity	LAT	SF	Average annual T (°C)	Annual PET (mm)	Annual P (mm)	Drainage area (km^2)
02111500	0.7	0.4	1.6	5.5	10.9	0.4	0.6	36.2	0	12.8	774.2	1335.8	233.7
06918460	0.5	0.1	3.5	8.3	14.4	0.3	0.7	37.4	0	13.4	843	1159.5	650.7
07067000	0.7	0.4	1.7	2.5	13.1	0.4	0.7	37	0	13.2	826.2	1183	4349

STAID	Max P season	Min P season	Max Q season	Min Q season
02111500	Summer	Fall	Spring	Fall
06918460	Spring	Winter	Spring	Fall
07067000	Spring	Fall	Spring	Fall

At the seasonal timescale, both of the Ozarks catchments (Figure 5.5; in purple) are consistently classified as the same curve type. However, several things are apparent: first, in a non-normalized format, as presented in panel A of Figure 5.5, it is clear that the catchment with young Karstic geology (06918460) and comparatively less long-term storage experiences a higher absolute magnitude of elasticity to precipitation (Figure 5.5 A) when compared to its counterpart. This is particularly clear in summer, where the curve shape is similar (Figure 5.5 B) but the estimated magnitude of elasticity differs by more than one percentage point. Second, despite having relatively similar curves at the seasonal timescale, these two catchments exhibit different behaviour at the annual timescale, where 06918460 has a strongly positive signal and 07067000 has a weakly positive signal, demonstrating the association between increased long-term storage and a less steeply sloping elasticity curve. At the annual timescale, the elasticity curves of these two catchments demonstrate the nuance required in interpreting the classification system – both curves span a similar total range of elasticity, however, the overall condition of the strongly positive curve (06918460) is steeper, as a large portion of the increase in the elasticity curve for 07067000 occurs between the 95th and 100th flow percentiles. Further, the more physically distant catchment (02111500; Figure 5.5, represented in green), has relatively similar characteristics to 07067000 (Figure 5.5; Table 5.1) and exhibits similar

curve structure at the annual and seasonal timescales, although with a slightly flatter overall condition.
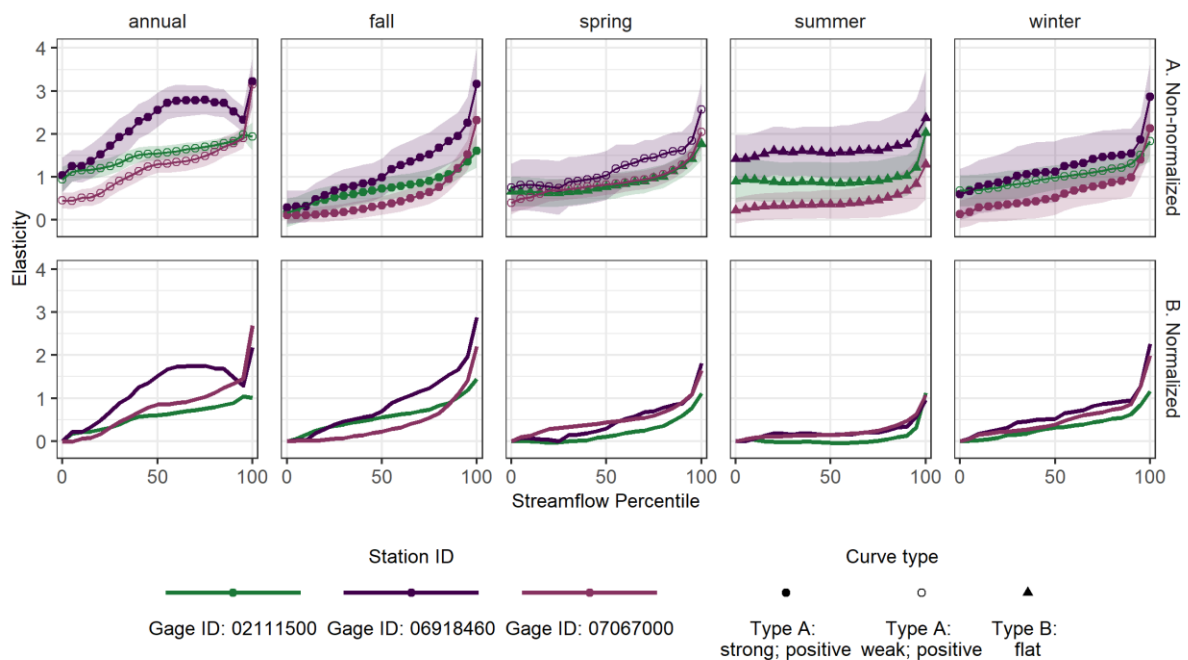
Figure 5.5 Examples of elasticity curves from three example catchments.

Turnback Creek above Greenfield (gauge id: 06918460), Current River at Van Buren (gauge id: 07067000), and Reddies River at North Wilkesboro (gauge id: 02111500). Panel A shows the non-normalized curves to demonstrate actual elasticity, and Panel B shows the normalized curves to demonstrate the similarity in curve form. Catchments located near one another geographically are both represented in shades of purple. Point shape represents the curve type and the ribbon represents the 95 percent confidence interval. Points and confidence intervals have been removed from Panel B to improve visibility, but the curve types and confidence bands are consistent across both panels.

Informative in the aggregate, the elasticity curve concept is limited in several ways, some of which are apparent in these examples. First, while curve shape is approximately consistent within the clusters, there is a margin of error around the groupings. The choice of the number of clusters per temporal scale was carefully considered in the interest of parsimony, so some catchments inevitably exhibit behaviour outside of the norm. Further, the shapes of the curves are not always smooth, as is evident in the example catchment 06918460, where a substantial decrease in elasticity is evident between the 80th and 95th percentiles at the annual timescale. The intention of this paper is to introduce the concept in a large-sample context and additional research is needed to determine the extent to which minor variations in shape may be due to

statistical noise or physical processes. Thus, the suitability of the concept for application to small scales remains to be established.

The lm-constructed curves or point estimates in individual catchments may deviate substantially from the cluster average, may comprise insignificant point estimates, or may violate assumptions of the regression approach used. For instance, depending on the streamflow percentile, the residuals of between 68 (ε_p^0) and 78 (ε_p^{100}) percent of the single catchment lms were normally distributed as estimated by a Shapiro-Wilks test with an alpha level of 0.01, and between 75 (ε_p^0) and 80 (ε_p^{100}) percent had a Durbin-Watson test statistic of greater than one, indicating that autocorrelation was not a serious concern in these sites. This means that the normality assumption was violated in around 20 to 30% of catchments and the non-autocorrelation assumption was violated in 20-25% of catchments. The fixed effects panel regression approach (Appendix 3 and Figure 8.2) helps to mitigate these concerns, lending credibility to the aggregated curves, but the reader is cautioned that application at the scale of a single catchment may carry substantial uncertainty. Further, the single catchment multivariate regression approach which we have taken here is a standard method for calculating point estimates of elasticity, however, this approach does not accommodate the possibility of non-linear elasticity, e.g. the possibility that a one percent and a 10 percent difference in precipitation are not linearly related. This work only considers the elasticity of streamflow magnitude, a singular component of streamflow which may not fully capture the influence of precipitation variability. Finally, the selected clusters depict whether curves are generally increasing or decreasing but do not account for the exact shape of the curves themselves, for instance, at which percentiles the slope begins to increase or decrease. In some instances, the curves for individual sites do not follow the precise curve types for which we have named the clusters. For instance, while the average curve in a cluster may be type A: strong, an individual curve may be type A: weak, etc. For this reason, we have presented the single catchment data with the interquartile ranges of curve estimates, and recommend caution when estimating elasticity curves or even elasticity magnitude for individual locations.

The work presented in this manuscript represents an introduction to elasticity curves. This concept may support further research into understanding of how changes in water storage might effect streamflow response over time (Saft et al., 2015, 2016), how groundwater contributes to flood-generation (Berghuijs & Slater, 2023), and provides insight into the implications of climate change for the hydrological cycle and the rainfall runoff relationship. Further, we include panel regression models as a tool for more robust elasticity estimation (Appendix 3) –

a method which may be well suited to regional calculation of elasticity and estimation in ungauged basins.

5.6 Conclusions

In this paper, we introduce a new concept for understanding and classifying streamflow response to precipitation. Representing streamflow elasticity to precipitation as a curve which reflects the range of responses across the distribution of streamflow within a given time period, instead of a single point estimate, provides a novel lens through which we can interpret hydrological behaviour. We have shown that ε_P estimated from the central summary of streamflow, e.g. the annual median, does not provide a complete picture of streamflow change. We have demonstrated that elasticity curve shape, i.e. the response of different flow percentiles relative to one another in a given catchment, can be understood separately from between-catchment variation in the magnitude of streamflow elasticity associated with a one percent change in precipitation.

We identify 3 typical elasticity curve shapes:

Type A – in which low flows are the least and high flows are the most responsive. The majority of catchments at the annual, winter, and fall timescales exhibit this behaviour.

Type B – in which the response is relatively consistent across the flow distribution. At the seasonal timescale, many catchments experience a consistent level of response across the flow regime. This is especially true in snow-fed catchments during cold months, when the actual elasticity skews towards zero for all flow percentiles while precipitation is held in storage. Consistent response is seen across the majority of the country during spring when streamflow is comparatively stable and rainfall driven, and in summer when evaporative demand is high and soil moisture is low.

Type C – where low flows are the most responsive to precipitation change. These catchments are dominated by highly flashy low flow behaviour.

Depending on the timescale examined, annual or seasonal, we predict elasticity curve type with fairly high accuracy, ranging from 95% in the fall to 63% in the spring, using catchment characteristics and other hydrologic signatures. The best predictors of curve type include the low end of the slope of the flow duration curve, mean annual temperature, seasonality, mean catchment elevation, and the baseflow index. All of these attributes relate to hydrological storage and release timing.

6 Streamflow sensitivity to precipitation shows large inter-annual and spatial variability

6.1 Abstract

Empirically-derived sensitivities of streamflow to precipitation are often assumed to be temporally unchanging. This assumption is unrealistic because changes in climate and storage will alter this relationship. We present a non-stationary regional regression applied to approximately 3000 catchments in the United States. We estimate long-term trends and variability in interannual streamflow elasticity to precipitation over a 39-year period. Elasticity is highly variable in water-limited catchments year to year, indicating high sensitivity to climate variability in arid regions, as compared to humid regions. These year-to-year variations in elasticity frequently correlate to the Standardized Precipitation Index (SPI) in the same and lagged years, suggesting that antecedent soil moisture, groundwater storage, and seasonality influence streamflow sensitivity. Finally, statistically significant long-term trends in elasticity exist in some regions, but trend magnitude is generally small. Total absolute changes in elasticity range from 0.28 to 0.60 over the study period, relative to long-term averages typically ranging from about 1-2.5. Assuming stationarity in long-term average elasticity may still be appropriate at the regional scale, but should be used cautiously.

6.2   Introduction

The rainfall-runoff relationship is moderated across multiple spatial and temporal scales by climatological and landscape properties. These control partitioning of precipitation between storage, runoff, and evaporative processes. 'Streamflow elasticity' is a simple metric which describes how responsive, or sensitive, streamflow is to precipitation or other environmental variables (Schaake, 1990). It is used to describe and classify catchments whilst also providing a useful empirical or theoretical metric for streamflow prediction. Usually estimated for the long-term average annual streamflow, elasticity is the expected proportional change in streamflow (Q) in response to a 1% increase in precipitation (Sankarasubramanian et al., 2001) and can be estimated using a range of approaches (Andréassian, Coron, et al., 2016). Typically, elasticity is calculated by taking the median of the inter-annual difference in annual streamflow divided by the inter-annual difference in annual precipitation (*P*) and normalized by the long-term runoff ratio (Sankarasubramanian et al., 2001). This value is often represented by a single number calculated for the period of record.

First order controls on the long-term average elasticity of streamflow to precipitation include the morphological and climatological characteristics of the catchment that are relatively constant in time (Figure 6.1 a). Topographic features include slope and geology, and long-term climatic norms may relate to the seasonality, phase, and quantity of precipitation (Chiew et al., 2006; Cooper et al., 2018; Sankarasubramanian et al., 2001) as well as the ratio of precipitation to potential evaporation (Berghuijs et al., 2017; Chiew, 2006; Chiew et al., 2006; Tang et al., 2019; Y. Zhang et al., 2022). Climatic forcing may influence elasticity through the storage components of the water balance, by controlling factors such as soil moisture, groundwater, and plant-water interactions. These features determine, to a large extent, a resulting gradient of streamflow elasticity across catchments, whereby some catchments are more responsive to changes in precipitation than others. On average, a 1% change in precipitation results in a 1-3% change in streamflow depending on the location and the methods used to estimate elasticity (Chiew, 2006; Chiew et al., 2006; Sankarasubramanian et al., 2001). Lower streamflow magnitudes are generally relatively less elastic (responsive) than higher streamflow magnitudes at the annual timescale (Anderson et al., 2023).

Many empirical and analytical studies on elasticity assume, implicitly or explicitly, that the relationship between streamflow and precipitation is linear and constant in time (Andréassian, Coron, et al., 2016; Bennett et al., 2018; Chiew et al., 2006; Konapala & Mishra, 2016; Y. Zhang et al., 2023). These assumptions are made either because the variability represents

uncertainty which is difficult to define, or because it is presumed that the long-term average relationship changes little over time (Y. Zhang et al., 2023), an assumption which is known to be physically unrealistic. Substituting for variation in space may be similarly unrealistic, as climate and landscape do not necessarily coevolve linearly (Berghuijs & Woods, 2016; Perdigão & Blöschl, 2014). Further, it is often assumed that at the annual timescale, terrestrial water storage has a negligible influence on streamflow (Arora, 2002; H. Yang & Yang, 2011), however, storage has been shown to have a strong influence on streamflow at the annual timescale (Donohue et al., 2010; Milly & Dunne, 2002; Tang et al., 2020; D. Wang, 2012; Wu et al., 2018) and has been shown to vary year to year (Krakauer & Temimi, 2011).

Natural variation in storage as a result of interannual climate fluctuations will influence how responsive streamflow is to precipitation on a year-to-year basis. For instance, antecedent moisture and groundwater, which may fluctuate annually, play important roles in controlling streamflow and elasticity across the flow distribution (Bennett et al., 2018; Berghuijs et al., 2016, 2019; Berghuijs & Slater, 2023; Fowler et al., 2022; Hughes et al., 2012; Neri et al., 2019; Safeeq et al., 2013; Saft et al., 2015; Slater & Villarini, 2017; Tague & Grant, 2009) with potentially greater effects for lower streamflow (Bennett et al., 2018; Cooper et al., 2018). Temperature changes can also have an effect on hydrologic response to precipitation through changes in evaporation (G. Fu et al., 2007), and snowfall quantity and timing (Donohue et al., 2011; Safeeq et al., 2013), with potentially greater effects for lower streamflow (Bennett et al., 2018; Cooper et al., 2018). A second mechanism by which the assumption of stationary elasticity may be flawed is in the presence of long-term climate change, or permanent shifts in the landscape, such as urbanisation, which may alter the division of precipitation between the streamflow, evaporation, and storage, but we do not explore this in detail in this study. The relationships described above are summarized in Figure 6.1, which outlines the causal pathways between precipitation and streamflow, showing long term controls on elasticity, as well as the hypothesized role of climate in determining interannual streamflow elasticity.

The simplifying assumption that streamflow elasticity is a temporally unchanging property of a catchment is a physical fallacy which has been pointed out for many years (Bassiouni et al., 2016; Harman et al., 2011). However, research explicitly into the non-linear or non-constant nature of streamflow elasticity has been somewhat limited. Recent work has aimed to address the problem by incorporating a storage component, represented by a proxy as baseflow or the residuals of the water balance closure, into elasticity estimation approaches (Cooper et al., 2018; Konapala & Mishra, 2016; Tang et al., 2020; Y. Zhang et al., 2023). The inclusion of

these parameters has led to the assumption that "elasticity changes little over time since it already represents the adaptation of catchment response to changes in its drivers" (Y. Zhang et al., 2023). This approach relies on the assumption that the baseflow index fully captures the effect of water storage within the same year, as well as the effect of water storage brought forward from previous years. While baseflow is associated with storage contributions to streamflow, for groundwater and other stores, it is unlikely to fully capture these storage components (Kalbus et al., 2006), or the variability which catchments experience on an interannual basis. Further, research directly into the interannual variability of streamflow elasticity to precipitation have argued that while some variability was present, particularly in arid regions, it was generally very small across the United States (Tang et al., 2020). These results represent an incomplete picture of hydrological sensitivity to precipitation: the catchments analysed were predominately concentrated in the Eastern U.S. (Tang et al., 2020), thus largely excluding water-limited regions of the western and central U.S.

Our principal hypothesis is that precipitation (in conjunction with temperature/evaporation) influences streamflow directly (Figure 6.1.b.), but also indirectly, by altering the available water storage within a catchment (Figure 6.1.c.). These relationships vary from catchment to catchment in line with its long-term average climatological conditions and landscape characteristics as has been shown in the existing elasticity literature (Figure 6.1.a.), but likely also vary in time. Patterns of temporal variation may be detectable in the observed record through the interannual elasticity calculated using a non-stationary regional scale model (Equation 6.1), designed to isolate the climatological effects.
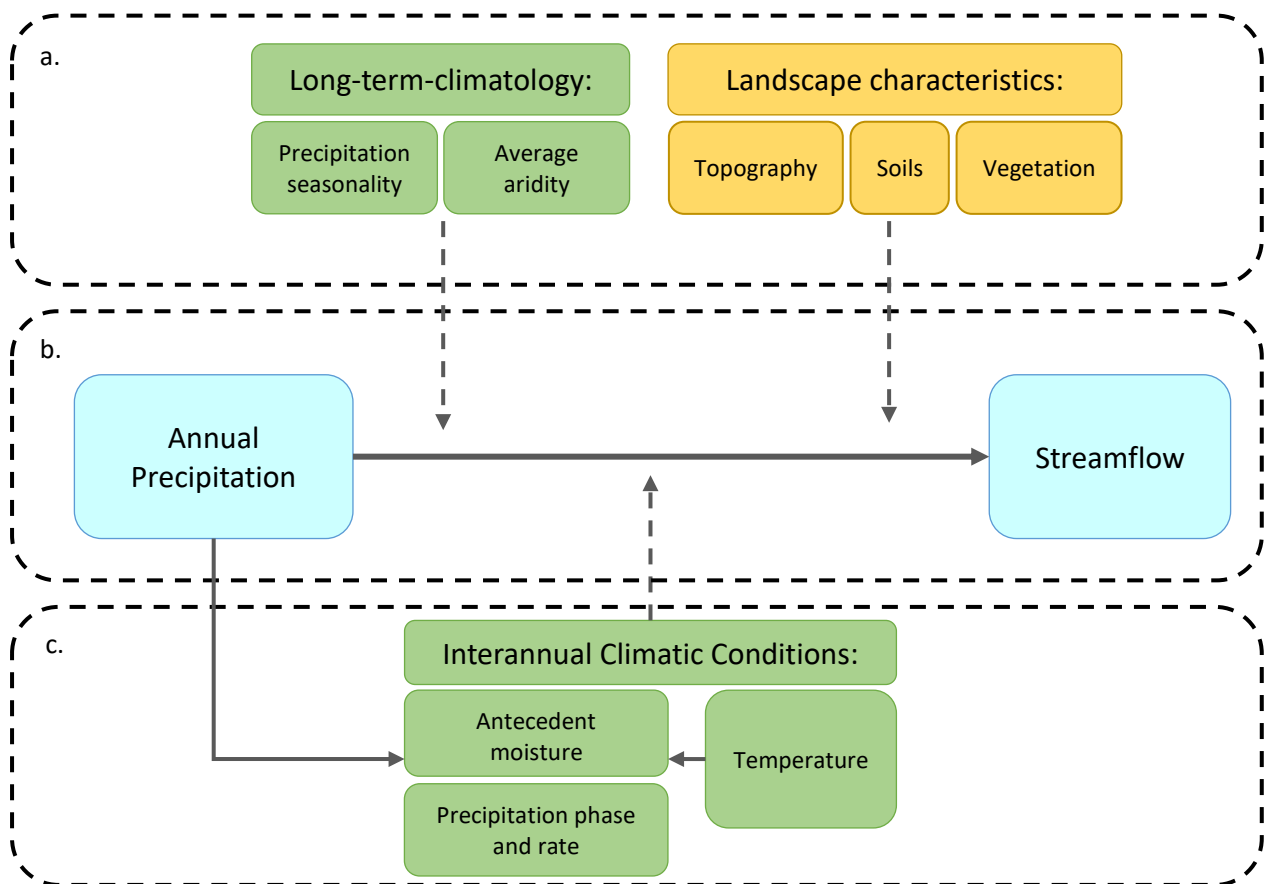
Figure 6.1 Conceptual diagram

(b) Describes the hypothesized pathways by which streamflow elasticity to annual precipitation is influenced by climate and catchment conditions as a long-term average (a) and interannually (c). Yellow rectangles indicate catchment scale processes, and green rectangles indicate regional scale processes. The relationship between the Blue rectangles is the relationship of interest in this study. Arrows represent conceptual relationships and not physical pathways. This diagram represents a necessary simplification of interactions between variables and some connections maybe excluded. For instance, in reality, boxes a and b would be connected.

As such, in this study, we aim to develop a more robust understanding of temporal change in streamflow elasticity to precipitation. By assessing interannual variability and long-term trends in interannual elasticity using a regional panel regression model, we hope to capture the influence of climate, a characteristic which largely exhibits regional-scale behaviour, on streamflow sensitivity to temporal changes in precipitation. Our regional regression-based approach is applied to nearly 3000 catchments in the United States to address the following:

How is interannual variability of streamflow elasticity to precipitation associated with climatic conditions at the regional scale? And, are long-term trends in interannual precipitation elasticity of streamflow present at the regional scale and if so, how do they relate to climate trends?

The approach used in this paper relies on a regional regression model in order to develop interannual elasticity estimates. In order to isolate climate effects and be able to capture interannual variability in sensitivity using observed data, the presented elasticities are calculated for climatologically similar regions, in which catchments may or may not exhibit hydrologically similar behaviour, but some analysis has been done to validate that they are reasonable similar. This is described in the methodology. We rely on the model design to isolate interannual and long-term variability which is the result of processes that occur at the regional scale, principally climate. In other words, the following results can be interpreted as regional averages, and variability in elasticity as an approximation of the relative sensitivity experienced in each region, rather than precise estimates of elasticity at individual locations. Further, spatial variation in elasticity is relatively poorly predictable (Addor et al., 2018), but to the best of our knowledge is most frequently attributed to long-term climate. Further, there is a strong connection between elasticity and the runoff ratio, which we feel lends credibility to the approach.

6.3 Methods

6.3.1 Observed hydrologic and climate data

The daily streamflow time series were obtained for the period 1981–2020 from the USGS website (https://waterdata.usgs.gov/nwis) using the R package dataRetrieval (DeCicco et al., 2024) and were selected from the Geospatial Attributes of Gages for Evaluating Streamflow (GAGES II) catchments (Falcone, 2011). These catchments were filtered so that streamflow records used in the analysis had at least 30 years of 95% complete consecutive daily streamflow data between January 1, 1981 and December 31, 2021. Daily streamflow records were converted to water years, defined here as October through September, so that snowfall seasons occurred within the same time step. Since water years require antecedent data, water years 1981 and 2021 were excluded from the final analysis because they are incomplete. The 10th, 50th, and 90th percentiles of daily streamflow were calculated for each water year.

Ephemeral streams, defined as having any zero flow days, were removed from the dataset because we rely on a log-log linear regression approach for this analysis. This resulted in a total sample size of 2967 catchments. Separately, a subsample of 830 catchments meeting the same

criteria but with minimal regulation, defined as in Anderson et al. (2023) and Blum et al. (2020), were selected and the same processes were applied in order to confirm that dam storage had little influence on the overall detected relationships. This had an effect on long-term trends as reported in the main text, but little influence on the overall patterns. Data used to estimate upstream dam storage was taken from GAGES II dataset, as were watershed boundaries (Falcone, 2017). Upstream dam storage was calculated by dividing total upstream dam storage by annual average inflow (Falcone, 2011) resulting in a value with units in days. Streamgauges with more than one day of upstream dam storage were eliminated (Blum et al., 2020; Hodgkins et al., 2019).

Gridded monthly precipitation and temperature (4 km resolution) were extracted from the Oregon State PRISM project (https://prism.nacse.org/recent/) using the R package prism (Edmund & Bell, 2015). We estimated average daily precipitation (mm/day) for each year from the monthly raster grid within each catchment boundary. We calculated average daily potential evaporation (mm/day) for each year in R using the Hamon equation (Hamon, 1963b; Lu et al., 2007b) with monthly temperature as previously described, and solar radiation estimated from latitude and Julian date (Equations 6.10-6.12). The Hamon equation was used to retain consistency with the GAGES II dataset and because this method has been shown to perform well relative to other approaches, despite its simple formulation (Lu et al., 2007b). Annual values for climatological variables were calculated for water years, to coincide with streamflow data and snowfall seasons.

Table 6.1 Number of catchments in each hydro-climatic region.

Bukovsky Regions	Number of catchments
a. Pacific Northwest	164
b. East	1090
c. South	336
d. Great Lakes	203
e. Mountain West	592
f. Prairie	269
g. Pacific Southwest	106
h. Great Plains	174
i. Desert	33

The Bukovsky regions (Bukovsky, 2011) were used to define the nine hydro-climatic regions used in this analysis. Table 6.1 contains the distribution of catchments per region. These regions were selected because they offer a balance between hydrologic significance and regional climatological consistency. To validate the use of these regions, we applied the classification approach used in Knoben et al. (2018), which uses hydrologic signatures to determine whether regions derived from climatological characteristics represent independent hydrologic behavior. The Bukovsky regions performed as well as this clustering approach overall, and are more spatially contiguous. Appendix figures 8.4 and 8.5 show the clusters derived using the method in Knoben et al. (2018) alongside the Bukovsky regions used in this analysis as well as an example of the differences in hydrologic signatures between Bukovsky regions. SPI (equation 6.6) and STA (Equation 6.7) were calculated from annual regional average precipitation and temperature, by taking the mean of the catchment averaged values in each region.

6.3.2 Panel model setup

We estimate average interannual elasticity using a fixed effects panel regression model, also referred to more generally as longitudinal regression models. These models resemble those which have been used for similar problems in hydrology (Anderson et al., 2022, 2023; Bassiouni et al., 2016; Blum et al., 2020). They allow for consideration of double indexed data, while controlling for time-invariant confounding variables at the catchment level, making them more robust than a typical cross sectional or time series regression model. The model is designed so that time invariant confounding variables which exist at the catchment scale are controlled for by the fixed effects, so that the captured variability in streamflow elasticity to precipitation is attributable to drivers which exist at the scale of the hydroclimatic regions. We assume that regional-scale changes predominantly relate to climate.

The regional average interannual elasticity estimates were calculated for three percentiles of streamflow: the 10th, 50th, and 90th using the following model:

$$ln\big(Q_{r(i),Y}\big) = \alpha_i + \varepsilon_{P(Y,r)}ln(P_{i,Y}) \cdot Y \cdot r + \varepsilon_{E(r)}ln\big(E_{i,Y}\big) \cdot r + \varepsilon_{BFI(r)}BFI_{i,Y} \cdot r + \eta_{i,Y} \quad 6.1$$

where $ln\big(Q_{r(i),Y}\big)$ is the natural logarithm of the annual streamflow percentiles, here either Q10, Q50, or Q90 calculated for year *(Y)* for catchment *(i)* in region *(r)*, α_i is the streamgauge-specific intercept, which controls for catchment-specific confounding variables, $ln\big(P_{i,Y}\big)$ is the natural logarithm of catchment averaged daily precipitation for each catchment in year *(Y)*,

$\ln(E_{i,Y})$ is the natural logarithm of catchment averaged daily potential evaporation, $BFI_{i,Y}$ is the baseflow index calculated annually.

Note that while the streamflow percentile varies for each model, the annual daily mean precipitation and potential evaporation values are used to predict all three streamflow percentiles. As such, our model estimates the sensitivity of low, median, and high flows respectively to a shift in the total annual precipitation. The interaction terms between the region variables and precipitation, potential evaporation and BFI are used to estimate different coefficients in each region and could be equally represented with an individual regression model applied to each region. The use of the interaction term here is to streamline the procedure for calculation. While this decision will influence the model adjusted R^2 and AIC, this choice should not affect the conclusions of our analysis.

Interannual variability in precipitation elasticity, the primary effect of interest, is represented by the regression coefficient: $\varepsilon_{P(Y,r)}$, the coefficients for potential evaporation and BFI are estimated as the long-term regional average only and are represented by $\varepsilon_{E(r)}$ and $\varepsilon_{BFI(r)}$ respectively. The error term is $\eta_{i,Y}$. BFI is included in the model because this approach has been taken in previous work on elasticity as a proxy for hydrologic storage. Note that the coefficient for BFI represents the regional average effect of the annual ratio of baseflow to total flow volume on streamflow. In reality, the influence of BFI on the estimated streamflow percentiles is likely to be specific to individual catchments because the proportion of streamflow which originates from groundwater sources, as well as most surface water storage is likely dependent on the geological and morphological characteristics of a catchment. Thus, BFI is included here for comparability with related models in other publications.

Equation 6.1 can be rearranged to solve for interannual precipitation elasticity in each region so that:

$$\varepsilon_{P(Y,r)} = \frac{(\ln(Q_{r(i),Y}) - \alpha_i - \varepsilon_{E(r)} \ln(E_{i,Y}) \cdot r - \varepsilon_{BFI(r)} BFI_{i,Y} \cdot r - \eta_{i,Y})}{(\ln(P_{i,Y}) \cdot Y \cdot r)} \qquad 6.2$$

The model is estimated in R using the plm package (Croissant & Millo, 2008) and the "within" estimator, indicating that fixed effects were used. The catchment level fixed effects control for time-invariant confounding variables at the catchment level, e.g. the physical characteristics of individual catchments, such as soil type and geology. By controlling for time-invariant confounders at the catchment level, we are able to isolate changes in $\varepsilon_{P(Y,r)}$ which occur at the

regional scale on a year-to-year basis, effectively limiting these effects to climatic changes. Thus, the interannual coefficients represent the effect of climate on the elasticity of streamflow to precipitation.

Autocorrelation in fixed effects panel models can lead to the underestimation of standard errors. We address this concern by clustering standard errors at the streamgauge level (Anderson et al., 2022, 2023; Bertrand et al., 2004; Blum et al., 2020). A generalized R script for the application of these approaches is available in the Supplementary Information (Text 8.4).

6.3.3   Model comparison

We fit two additional forms of this model for comparison. The first is a fixed effects panel regression model which excludes the interaction term for "year" as below:

$$ln\big(Q_{r(i),Y}\big) = \alpha_i + \varepsilon_{P(r)}ln\big(P_{i,Y}\big) + \varepsilon_{E(r)}ln\big(E_{i,Y}\big) \cdot r + \varepsilon_{BFI(r)}BFI_{i,Y} \cdot r + \eta_{i,Y} \qquad 6.3$$

where all terms are the same as in Equation 6.1. This model estimates the average elasticity across the period of record without considering the effects of temporal change in precipitation on streamflow.

The second model form is a version of Equation 6.1 which excludes the annual baseflow index (BFI), as in Equation 6.4 below.

$$ln\big(Q_{r(i),Y}\big) = \alpha_i + \varepsilon_{P(Y,r)}ln\big(P_{i,Y}\big) \cdot Y \cdot r + \varepsilon_{E(r)}ln\big(E_{i,Y}\big) \cdot r + \eta_{i,Y}) \qquad 6.4$$

where all terms are the same as in Equation 6.1. This model is estimated to determine if the inclusion of BFI substantially reduces the effect of interannual variability on the model.

The results of these two models are compared with the results of the time variant model in Equation 6.1 using the Akaike information criterion (AIC), and indicate that the model which accommodates temporal variation in the precipitation elasticity of streamflow as well as annual BFI (Equations 6.1 and 6.2) outperforms the long-term average model (Equation 6.3) and the temporally variant model which excludes BFI (Equation 6.4) for every flow percentile tested (Supplementary Materials Table 8.3).

6.3.4   Hydrologic signatures and other tests

AIC is estimated for the panel models manually in R as follows:

$$AIC = -2\,ln\left(\hat{L}\right) + 2k \qquad\qquad 6.5$$

where $ln\left(\hat{L}\right)$ is the log likelihood function and k is the number of estimated parameters. The model with the lower resulting AIC value is considered to be the best fitting model.

It is worth noting that the sample size within each region has an effect on the precision and accuracy of regression coefficient estimation in the panel regression model. As in Table 6.1, the Desert region contains substantially fewer catchments than other regions in the analysis, which may partially explain the relatively wider confidence intervals and reduced significance of interannual elasticity.

Trends in the interannual elasticity estimates were estimated using a trend-free pre-whitened version (Bayazit & Önöz, 2007) of the widely applied Mann Kendall trend test (Hamed, 2008; Kendall, 1948; Mann, 1945). This test was applied using R package modifiedmk (Jassby et al., 2022).

Additional variables used in the analysis were calculated as below:

Regional Standardized precipitation index (SPI):

$$SPI_r = \frac{P_{rY} - \overline{P_r}}{sd(P_r)} \qquad\qquad 6.6$$

where $\overline{P_r}$ = regional average (r) of annual mean daily precipitation (P) across all years; P_{rY} = as above in each year (Y); $sd(P_r)$ = standard deviation of regional average precipitation across all years.

Regional Standardized temperature anomaly (STA):

$$STA_r = \frac{T_{rY} - \overline{T_r}}{sd(T_r)} \qquad\qquad 6.7$$

As above for average annual daily mean temperature *(T)*

Snow fraction (SF):

$$SF_i = \frac{\sum P_{i,m}(T_{i,m} < 0)}{\sum P_i} \qquad\qquad 6.8$$

where $\sum P_i(T_{i,m} < 0)$ = the sum of monthly average daily *(m)* precipitation *(P)* in catchment *(i)* when the monthly average daily temperature $(T_{i,m})$ was below 0°C

Aridity index (AI):

$$AI_i = \frac{\bar{E}_i}{\bar{P}_i} \qquad\qquad 6.9$$

where $\bar{E}_i$ = long term average of the annual daily mean potential evaporation *(E)* in catchment *(i)* measured in mm/day; $\bar{P}_i$ = long term average of the annual daily mean precipitation *(P)* in catchment *(i)* measured in mmday^{-1}

Baseflow index (BFI):

The baseflow index is estimated using the UK Institute of Hydrology "smoothed minima" approach (Gustard et al., 1992) using a standard 5-day block size (Stoelzle et al., 2020) and the R package lfstat (Gauster et al., 2022).

Potential evaporation *(E)* in mm/day was calculated using the Hamon equation (Hamon, 1963b):

$$E = 0.1651 \cdot Ld \times \rho \cdot K \qquad\qquad 6.10$$

$$\rho = \frac{216.7 \cdot E_{sat}}{(T+237.3)} \qquad\qquad 6.11$$

$$E_{sat} = 6.108 \cdot exp\left(\frac{17.26939 \cdot T}{(T+237.3)}\right) \qquad\qquad 6.12$$

where Ld = length of day in units of 12 hours; ρ = saturated vapour density (gm^{-3}) at daily mean air temperature (°C) *(T)*; E_{sat} = saturated vapor pressure *(mb)* at a given T; K = calibration coefficient, here set to 1.

6.3.5 *Limitations*

As with other estimates of elasticity, it is important to recognise limitations of the approach. For instance, we focus on a single-parameter precipitation elasticity of streamflow, and do not consider the explicit interactions and subsequent non-linear nature of streamflow elasticity as precipitation and temperature (potential evaporation) vary simultaneously (G. Fu et al., 2007). The particular method employed here is limited in many ways. Primarily, the elasticity estimates presented in the study represent within-group averages for entire regions, and

therefore are not likely accurate estimates of streamflow sensitivity at individual locations. The fixed effects in the panel regression model control for time-invariant confounders at the catchment scale (ex. topography), and thus it is assumed the interannual and long-term variability detected here is the result of processes occurring at the regional scale, principally climate. In other words, these results can be interpreted as regional averages, and variability in elasticity as an approximation of the relative sensitivity experienced in each region, as noted in the introduction. The variations in elasticity which are presented in this study are robust at the regional level, with the caveat that sample sizes in each region influence the statistical significance of the annual estimates. This is most relevant to the estimates for the Desert region which contains fewer catchments than the others, the implications of which are visible in the breadth of the associated confidence intervals (Figure 6.3). Regardless, the majority (~80% for low and median and 75% for high flows) of interannual elasticity estimates are significant ($p < 0.05$) in the Desert region.

6.4 Results

6.4.1 Regional scale interannual variability

We estimate regional average interannual elasticity for nine hydro-climatic regions. The resultant estimates sometimes deviate greatly from elasticity estimated as a single value for the entire period of record (Table 6.2). For instance, the mean absolute difference between the long-term elasticity estimate and the interannual estimates for median streamflow in the Prairie region is approximately 0.5. This means that on average across the 39-year time period, the elasticity in each year deviates from the long-term mean elasticity estimate (2.66) by half a percentage point, with elasticity estimates in each year ranging between 1.4 and 2.9 (Table 6.2). Thus, a 10% change in precipitation would correspond to anywhere from a 14% to 29% change in streamflow, assuming changes magnify linearly.

We use the Akaike information criterion (AIC) to determine the best model fit. The inclusion of an interannual component in the regression model improves the model fit for all three percentiles examined (Supplementary Information Table 8.3). The interannual component improves model fit including when the annual baseflow index, as a representation of groundwater inputs, is included in the model, suggesting that BFI alone fails to capture all variability.

Table 6.2 . Mean absolute difference in the regional interannual elasticity estimates and the long-term elasticity estimate estimated using the whole period of record for low (Q10), median (Q50), and high (Q90) streamflow as estimated from the 39-year period of record.

Region	Q10			Q50			Q90		
	Mean absolute difference	Long-term estimate	Interannual range	Mean absolute difference	Long-term estimate	Interannual range	Mean absolute difference	Long-term estimate	Interannual range
Desert	0.29	0.77	0.28-1.55	0.30	0.63	-0.07-1.44	0.66	1.11	-0.41-2.55
East	0.10	1.76	1.5-2.05	0.11	1.74	1.26-1.92	0.09	1.60	1.39-1.82
Great Lakes	0.13	1.50	1.2-1.86	0.15	1.41	0.95-1.57	0.11	1.42	1.04-1.62
Great Plains	0.27	1.18	0.18-1.71	0.31	1.35	0.24-2.03	0.35	1.98	0.39-2.59
Mountain West	0.10	0.87	0.58-1.04	0.15	1.14	0.79-1.24	0.12	1.90	1.58-2.14
Pacific Northwest	0.05	0.69	0.57-0.86	0.05	1.42	1.29-1.55	0.04	1.56	1.48-1.69
Pacific Southwest	0.19	0.77	0.11-1.12	0.42	1.34	0.23-1.36	0.47	2.19	1.13-2.24
Prairie	0.46	2.25	1.1-2.61	0.48	2.66	1.41-2.87	0.27	2.54	1.33-2.88
South	0.22	1.45	0.97-1.54	0.14	2.22	1.77-2.43	0.11	2.29	2.08-2.58

We find clear statistically significant differences in the average interannual variability of streamflow elasticity to precipitation between the nine Bukovsky regions (Bukovsky, 2011) (Figure 6.2.a). The regions with the highest variability tend to be more water-limited (higher aridity index) and receive the majority of their precipitation as rain rather than snow, although not exclusively (Figure 6.2.b). The most variable regions also have the lowest runoff ratios, somewhat inline with existing literature (Chiew et al., 2006). Meanwhile the regions which experience the least interannual variability in streamflow elasticity to precipitation (e.g. the East and South) are also warm and rain-dominated, but climatic conditions in these regions are far more humid on average (Figure 6.2). The precise estimates of elasticity in each year are included in Figure 6.3, demonstrating how the range of hydrological responses can vary widely from year to year.
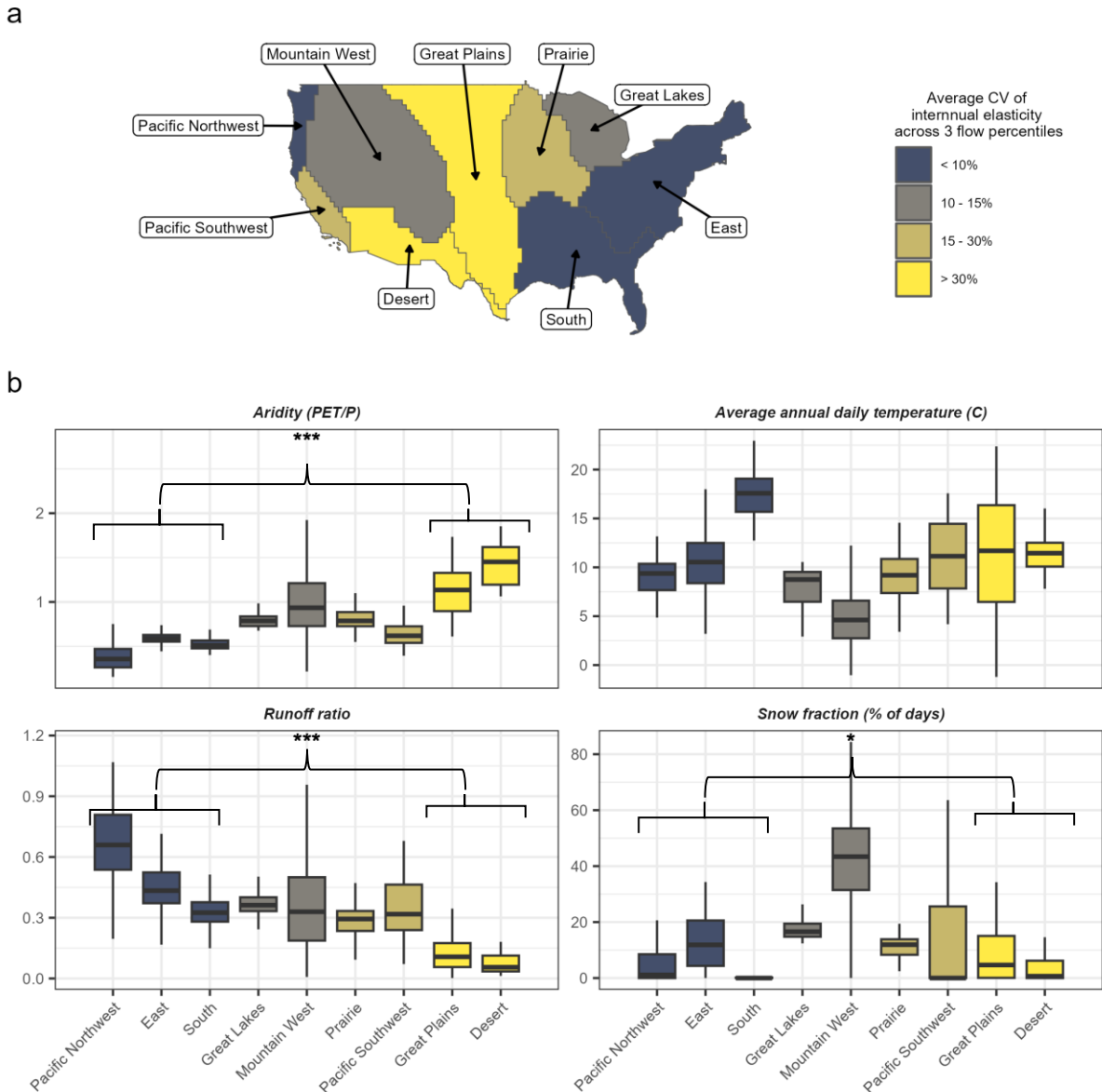
Figure 2 Average interannual variability of elasticity and catchment attributes by region.

a) map of hydro-climatologic regions coloured by average interannual variability in elasticity to precipitation, as determined by the mean coefficient of variation (CV) across three flow percentiles, rounded to the nearest whole number. b) Box plots showing the distribution of climatological attributes in each region. Centre line is the median, boxes show the 25th and 75th percentiles, and whiskers represent the interquartile range multiplied by 1.5 from the 25th and 75th percentiles. Regions are ordered on the x-axis from least to most variable based on average CV across three flow percentiles and coloured as in b. Stars indicate a significant difference in the median of each variable of interest between the groups with the least variable

6.4.2 Long-term trends in interannual elasticity

We conducted a Mann Kendall test to determine whether there was a statistically significant trend in the long term mean of the regional interannual elasticities (Figure 6.3; Table 6.2). Results indicate that some long-term trends are present in certain flow quantiles in the Desert, South, Great Plains, Mountain West, Pacific Northwest and Pacific Southwest regions (a, c, e, g, h, i). For the most part, the magnitude of the trends is very small (Table 6.2), so while statistically significant, they may not represent a substantial change in average streamflow sensitivity to precipitation over the 39-year study period.

Table 6.3 Results of the Mann Kendall Trend test on interannual elasticity and the coefficient of variation of interannual elasticity.

Statistically significant trends (p < 0.05) are presented in bold font. Total change is the Sen's slope rate of change multiplied by 39 years. This is a unitless metric representing the linearly-estimated difference in streamflow elasticity to precipitation in water year 1982 vs. water year 2020. Coefficient of variation in each region over each time period is presented as a percentage.

We reran the analysis with a sub-sample of data (Supplementary Information Table 8.5) from which sites with substantial dam storage had been removed. This resulted in failure to detect long-term trends in the Pacific Southwest (g) and for low flows in the Desert (i) region (Supplementary Information Table 8.5). Trends in the South, Mountain West, and for median flows in the Desert region (c, e, i) persisted in this smaller sub-sample, and a positive trend in low flows in the Great Plains region (h) is apparent in the smaller sample. These results suggest that long-term trends in these regions may be due to anthropogenic water use, with the caveat that this change resulted in substantially reduced sample sizes (Supplementary Information Table 8.5) and therefore confidence in the estimates. Interannual variability remained relatively consistent in models resulting from this subsample.

Table 6.4 Results of the Mann Kendall Trend test on interannual elasticity and the coefficient of variation of interannual elasticity.

Statistically significant trends (p < 0.05) are presented in bold font. Total change is the Sen's slope rate of change multiplied by 39 years. This is a unitless metric representing the linearly-estimated difference in streamflow elasticity to precipitation in water year 1982 vs. water year 2020. Coefficient of variation in each region over each time period is presented as a percentage.

Mann Kendall trend test					
Bukovsky Regions	Q10 total change	Q10 p value	Q50 total change	Q50 p value	Q90 total change	Q90 p value
a. Pacific Northwest	0.02	0.687	0.06	0.145	**0.07**	**0.042**
b. East	0.03	0.669	-0.04	0.669	-0.08	0.191
c. South	-0.03	0.669	**-0.28**	**0.003**	**-0.28**	**< 0.001**
d. Great Lakes	-0.02	0.763	-0.07	0.513	-0.04	0.669
e. Mountain West	0.06	0.436	0.04	0.706	**0.23**	**0.008**
f. Prairie	0.28	0.291	0.05	0.880	-0.29	0.108
g. Pacific Southwest	**0.35**	**0.002**	**0.33**	**0.013**	**0.38**	**0.006**
h. Great Plains	-0.05	0.615	-0.37	0.066	**-0.52**	**0.033**
i. Desert	**-0.60**	**0.012**	**-0.58**	**0.010**	-0.56	0.303

Coefficient of variation			
Bukovsky Regions	Q10 CV	Q50 CV	Q90 CV
a. Pacific Northwest	9%	4%	3%
b. East	7%	8%	7%
c. South	12%	8%	6%
d. Great Lakes	10%	12%	9%
e. Mountain West	14%	13%	7%
f. Prairie	21%	17%	14%
g. Pacific Southwest	34%	27%	17%
h. Great Plains	30%	33%	26%
i. Desert	45%	58%	69%
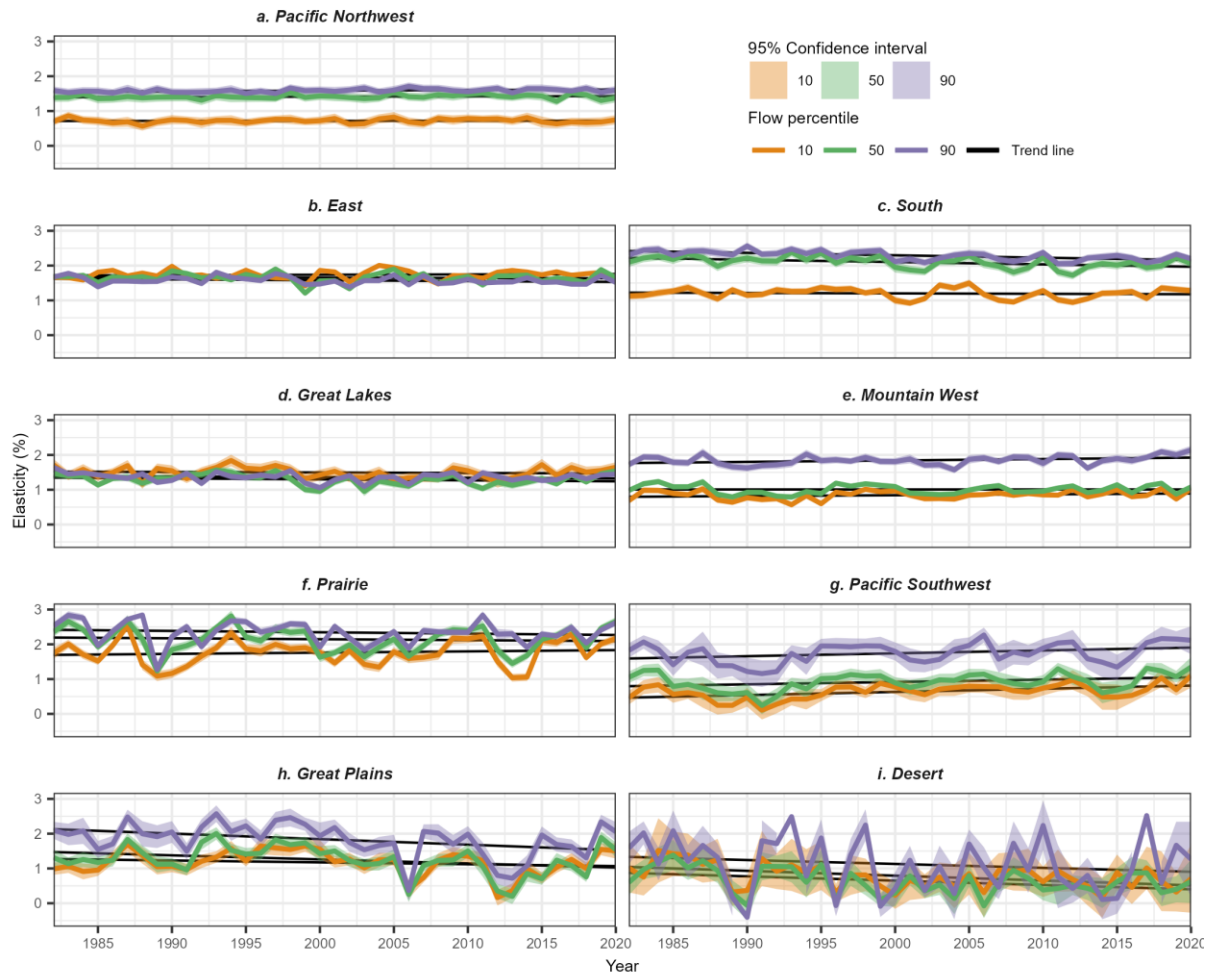
Figure 6.3 Interannual variability and long-term trends in interannual elasticity estimates for the 10th, 50th, and 90th percentile of streamflow by hydro-climatological region.

The linear Mann Kendall trend lines are presented in black. Panels are ordered on the x-axis as in Figure 6.2 and all share the same axis limits and breaks in order to facilitate comparison.

6.4.3 Comparison of interannual variability to climate

We are interested in capturing how year-to-year climate variability influences streamflow sensitivity to precipitation. Therefore, we next compared the interannual elasticity estimates to contemporary and lagged climatic variables, namely, the regional standardized precipitation index (SPI) (Figure 6.4) and the regional standardized temperature anomaly (STA) (Figure 6.5). SPI is highly ($R > 0.5$ and $p < 0.05$) or moderately ($0.5 > R > 0.3$ and $p < 0.05$) positively correlated with interannual variability in elasticity in the previous, present, or both years in a majority of regions: 89% of regions for low flows, 89% for median flows, and 56% for high flows. The five regions with the most variable interannual elasticity (e-i) were correlated with

SPI in the same year or first lag for all examined flow percentiles (Figure 6.4 e-i). Thus, in a majority of regions, streamflow is more responsive to average annual precipitation if that year, the previous year, or both are wetter than average.

Figure 6.4 Pearson correlation between streamflow elasticity to precipitation and the regional average SPI in the same (lag_0), the previous (lag_1), or two water years prior (lag_2) for the 10th, 50th, and 90th percentiles of streamflow.
*Significance is rounded to two digits and indicated by stars where p < 0.05 is represented by * and p < 0.01 is represented by **.*

SPI correlation with the first lagged time step is most prevalent for low to median flows. The majority of discharge for the 10th percentile of streamflow comes from delayed sources (Table 8.7). The correlation between delayed SPI and elasticity is likely due in part to the seasonality of precipitation and streamflow (Supplementary Information Table 8.4). For example, the low flow season for the majority of catchments in the desert region is summer, thus occurring within the same water year as the high precipitation season (typically winter or summer). In the majority of catchments in the Pacific Southwest region, the low flow season is fall while the largest proportion of precipitation falls in winter which occurs after fall in the water year (October-September). Thus, it follows that antecedent precipitation in the previous year could have a larger effect on low flows than those in the same year.

Baseflow, the proportion of streamflow originating from delayed sources, also represents more than 10% of median streamflow in the Pacific Southwest, Mountain West, Pacific Northwest, East, and Great Lakes regions (a, b, d, e, g), based on the regional median ratio of annual average daily baseflow to the annual median streamflow (Supplementary Information Table 8.6). This contribution may explain some of the pattern of delayed responses in median

116

streamflow. This proportion is even larger in catchments without substantial dam storage (Supplementary Information Table 8.6). The precipitation elasticity of high streamflow magnitude shows fewer correlations with SPI overall. Results indicate a correlation only in the five most variable regions.

Correlations between interannual elasticity and temperature, characterized using STA, are weaker and less consistent overall. However, STA is moderately negatively correlated ($p < 0.05$) with elasticity in the same year for low, median, and high flows in the most arid regions, the Desert and Great Plains (h, i), and for low flows in the Great Lakes (d) region (Figure 6.5). At the lag-1 timestep, STA is negatively correlated with interannual elasticity in the Desert, Great Plains, Prairie, and Great Lakes regions (d, f, h, i) for low flows, the Great Plains, Prairie, and Great Lakes regions (d, f, h) for median flows, and the Great Plains and South regions (d, h) for high flows. These results indicate that increased temperatures exert some negative influence on streamflow elasticity on a year-to-year basis, likely as a result of increased potential evaporation and drying soils (Sharma et al., 2018). In the Mountain West region (e), low flow elasticity is positively correlated with temperature, indicating that warmer temperatures result in low flows which are more responsive to precipitation. This is likely related to snowfall, as warmer temperatures will result in more precipitation as rain during winter as well as earlier snowmelt times, eliciting a more rapid streamflow response (Berghuijs et al., 2014b).
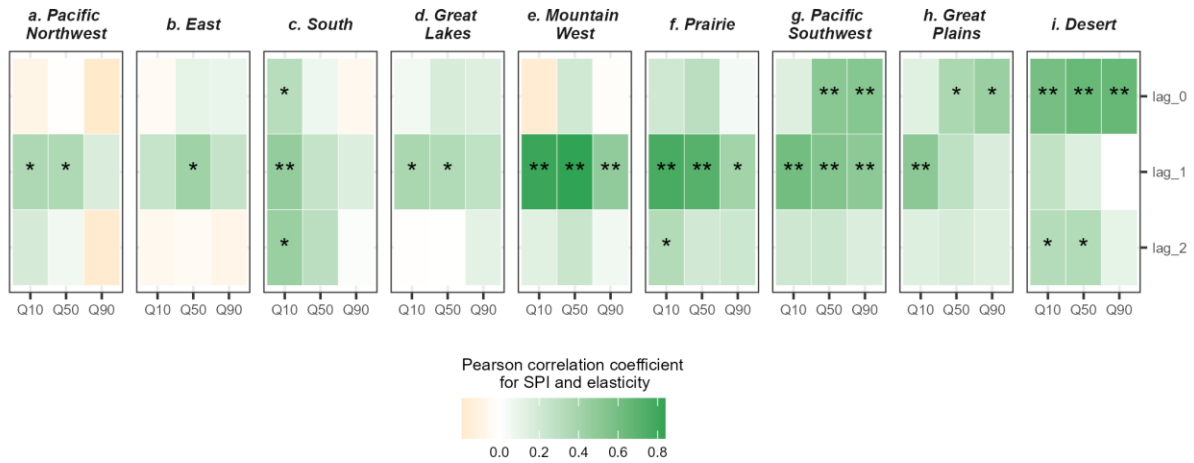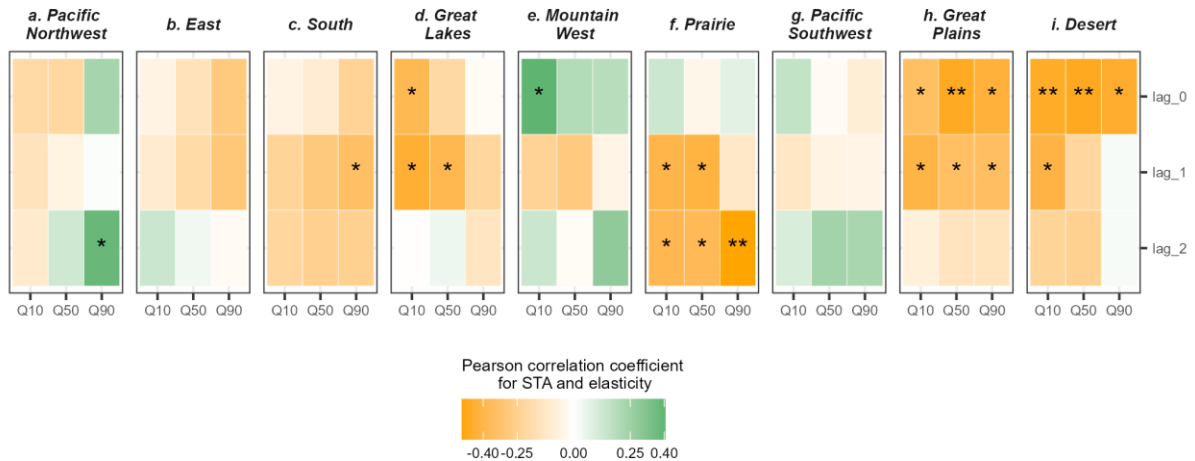
Figure 6.5 Pearson correlation between streamflow elasticity to precipitation and the regional average STA in the same (lag_0), the previous (lag_1), or two water years prior (lag_2) for the 10th, 50th, and 90th percentiles of streamflow.

Significance is rounded to two digits and indicated by stars where p < 0.05 is represented by * and p < 0.01 is represented by **.

6.5   Discussion

Building on our foundational hypothesis that *streamflow elasticity is unlikely to be stationary in time*, the objective of this paper is to quantify the extent to which streamflow sensitivity to precipitation, as characterized by elasticity, may change over time. We show that elasticity varies over multiple time scales – interannually and in some regions on average -- over the 39-year study period. Interannual variation is not fully captured by inclusion of the baseflow index and long-term trends, although small, may be indicative of non-stationarity.

Our findings align with prior literature regarding the variability of flow in arid and semi-arid catchments (Farquharson et al., 1992), and the importance of antecedent moisture conditions for generating flow in these regions (Ivancic & Shaw, 2015). For example, a recent analysis presented the concept of an "antecedent effect ratio" (AER) which quantifies the difference in catchment response to heavy precipitation dependent on antecedent moisture conditions (Bennett et al., 2018). They found that more arid catchments had high AER values, indicating that antecedent conditions were more important for flood flow generation in drier catchments than wetter ones. The authors speculated that the difference in response was related to the range of antecedent moisture conditions in drier catchments (i.e. arid catchments become much drier than humid ones, so relative moisture makes a bigger difference in response) (Bennett et al.,

2018). Rivers in the Great Plains, Prairie, and Desert regions of our analysis likely have a very weak seasonality and flows are dominated by the occurrence of short precipitation events such as storms (Brunner et al., 2020). The fast and slow flow thresholds, a parameter that indicates the amount of precipitation required to generate storm runoff and baseflow respectively, are likely to consume a large proportion of annual precipitation in arid catchments (Harman et al., 2011; Sivapalan et al., 2011). Thus, arid catchments more susceptible to fluctuations in precipitation which fall below or rise above those thresholds than catchments in humid regions.

In other words, streamflow which is driven predominantly by weather events, rather than longer-term storage, is likely to be more sensitive at short timescales to both antecedent and contemporary climatic conditions. Therefore, it is unsurprising that these regions experience larger shifts in elasticity on an interannual basis. In fact, previous research has suggested that drier regions experience higher, and much more spatially variable elasticity (Harman et al., 2011; Sankarasubramanian et al., 2001). Many of these studies have relied on the MOPEX dataset, which is geographically limited in the especially arid Desert and southern portion of the Great Plains. Contrary to these studies, we find relatively low long-term average elasticities in the arid Desert region. However, highly variable interannual elasticities here indicate that previously identified high spatial variability in dry regions may also be reflected in temporal variability, with highly elastic, and minimally elastic years interspersed through the historical record.

We detect some significant long-term trends in the regional interannual elasticity estimates, indicating that overall streamflow sensitivity to precipitation may be changing in some regions and for some parts of the flow distribution. Previous research has indicated that the elasticity is correlated with the Horton index, a water balance-based indicator of catchment humidity which relates evaporation to catchment wetting (precipitation-surface runoff) (Harman et al., 2011) Shifts in this index at long timescales may cause changes in catchment wetting and vaporization thresholds, eventually leading to changes in elasticity as a new climatic equilibrium is achieved (Harman et al., 2011). Further, it is possible that prolonged drought conditions may influence the long-term relationship between precipitation and streamflow (Fowler et al., 2022; Saft et al., 2015), and it is the case that significant increasing trends in drought (Ficklin et al., 2015), as well as large scale decreases in terrestrial water storage (Slater & Villarini, 2016a) have been detected in the Desert region and in large parts of the South. These patterns may carry some responsibility for the decreasing trends in elasticity in these regions, providing an avenue for future research.

That said, the statistically significant trends which we detect across all regions are small in magnitude, showing a total absolute estimated linear change over the 39-year period of between 0.28 and 0.60. In some regions these changes may be explained by regulation. Thus, these results imply that large-scale shifts in elasticity due to climate may be occurring in some places, but that so far, the geographic extent and magnitude of these changes is limited. Because of the regional design of our study, we focus here on climate impacts, however, other anthropogenic activity such as land cover changes are also likely to have an effect on elasticity at smaller spatial scales – potentially outsizing the average climate-driven patterns we find here. Long-term trends in a number of regions appear to be explained by dam regulation, as removal of these catchments from the data sample eliminated trends in several regions. In these cases, it may be reasonable to assume that long-term trends are due to management of hydrological resources rather than changes in climatology.

Improving knowledge of temporal variability in elasticity is an important first step towards understanding how, and if, elasticity can be used to project streamflow changes into the future, or to validate simulations. These questions lend themselves to larger considerations which are not addressed here, regarding how to manage hydrological non-stationarity, and the lengths of time series necessary to adequately estimate streamflow sensitivity from the observed record. For instance, large interannual variability in arid regions suggests that using short periods of record may strongly influence elasticity estimates in these areas. Meanwhile in the Pacific Northwest, for example, where interannual variability in elasticity is marginal, a short period of record may have little influence.

The results of this study imply that the time period used to estimate elasticity may have an effect on the resultant value, because large year-to-year variability exists in some regions. We demonstrate that climate conditions in recent years play an important role in determining streamflow sensitivity to precipitation at the annual timescale. The difference in model performance indicates that the baseflow index may fail to fully capture the influence of changes in storage on streamflow sensitivity, and highlights this as an important avenue for future research. Further, we show that long-term elasticity estimates may systematically under or over-estimate this relationship, particularly if storage changes and climatic conditions are not considered. While long-term changes in elasticity are currently small, further climate change and anthropogenic activity may alter this relative equilibrium. Further exploration, for instance, into the effects of land cover changes and groundwater abstraction would be instructive.

7 Discussion

The overarching aims of this thesis were to advance understanding of the relationships between potentially changing environmental variables, namely land cover changes and precipitation, and streamflow magnitude using observed data from thousands of catchments. I sought to assess the statistical robustness of a series of regression approaches for the quantification of the sensitivity of streamflow magnitude to land cover changes across the U.S; to determine the average influence of tree cover change and urbanisation on downstream flow magnitude in U.S. rivers; to develop an elasticity-based classification system for U.S. rivers which demonstrated the variable nature of elasticity across the annual and seasonal flow distributions; to assess the likely drivers of intra-catchment variability in elasticity; and finally, using the approaches established through the previous objectives, to assess the efficacy of an assumption of stationarity in the streamflow-precipitation elasticity relationship.

In this thesis, I have sought to understand the drivers of these changes, their extent and impact, and as well as methods for detecting and attributing their causes. Once embedded in this literature, my research questions took on a technical, and often methodological focus, often leaning towards understanding the potential flaws in existing methodologies, as well as the robustness and reliability of alternate approaches.

The remainder of this discussion and conclusions section is structured as follows:

First, I provide summaries of each of the research chapters. Then, I discuss the binding themes of the research, before exploring some future directions and presenting my final remarks.

7.1 Chapter summaries

7.1.1 Statistical attribution of the influence of urban and tree cover change on streamflow: A comparison of large sample statistical approaches

In the first part of the research comprising this thesis, I conducted a large-sample analysis exploring the relationships between land cover changes and streamflow magnitude for low, average, and high flows. I used 729 catchments which fit a strict set of criteria and causal diagrams to guide the creation of two kinds of statistical models. I aimed to discern the influence of tree cover changes and urbanisation on streamflow in the United States between 1992 and 2020. Panel regression models were used as a tool for causal inference. The idea with using these models was that by pooling data across many locations, more robust relationships could be detected. This aim is typically challenging with single-location time series or cross-sectional models given noisy data and the small effects which land cover changes may often

be expected to have on streamflow. I compared the panel model results to single-catchment time series models (GLMs).

Overall, I found small but statistically significant effects of urbanisation on median and high streamflow using the panel regression models. Tree cover changes did not have statistically significant effects on flow and urbanisation did not have a statistically significant effect on low flows on average. The results of the single-catchment models measuring the relationship between land cover changes and streamflow were very wide-ranging and inconsistent, demonstrating the improved robustness of the panel regression models over single-location timeseries approaches.

This work builds upon limited previous research demonstrating the improved robustness of panel regression models for applications in hydrology (Bassiouni et al., 2016; Steinschneider et al., 2013). Thus, this is an important contribution towards the use of panel regression techniques as a tool for hydrological analysis. Further, our results contribute to the debate around the generalizability of rules of thumb regarding the effects of urbanisation and tree cover changes on streamflow. For instance, while small sample studies have frequently shown that increases in tree cover decrease streamflow, I show no statistically significant effect when tree types, stand age, drainage area, etc. are averaged. This indicates that tree cover changes may only have a clearly discernible effect on streamflow under certain circumstances. This empirical result corroborates results from some modelling studies and reviews (Farley et al., 2005; Goeking & Tarboton, 2020).

In the first research chapter, precipitation and temperature were included in the panel regression models to represent the climatic variables which drive streamflow. This model produced elasticity estimates for both temperature and precipitation which were examined relative to the land cover coefficients. The precipitation elasticity coefficients were highly significant, relatively consistent even in the GLMs and differed between low, average, and high flows when considered in aggregate at the national scale. These results led to the concept which is addressed in the second research paper (Chapter 5).

7.1.2 *Elasticity curves: A novel lens for interpreting the variable nature of the streamflow-precipitation relationship*

In the second research chapter, I estimated streamflow elasticity to precipitation across the different segments of the flow distribution simultaneously using data from 805 perennial rivers in the United States. Calculated for every 5th percentile of flow using a GLM, elasticity

coefficients were aggregated graphically so that streamflow responsiveness to a change in total annual precipitation could be visualized as a curve. I then clustered the curves based on similarity of their shapes and show that there are clear regional patterns in some seasons. The patterns correspond to some extent with hydrologic signatures and catchment characteristics, including the baseflow index, slope of the flow duration curve, and the aridity index, among others. I posit that the shape of the curve corresponds to water storage capacity within a catchment.

The concept of an elasticity curve is a novel one. To our knowledge, only one other study has compared the sensitivity of streamflow to precipitation across different flow percentiles at all, and in that case they compared point estimates only to state that low flows were less elastic than average and high flows (Harman et al., 2011). Therefore, the paper offers a new theoretical and methodological concept, and a new analytical tool to diagnose catchment response to external climatic drivers.

The idea for the final research chapter was born out of the work conducted on elasticity curves. Initially, the idea was that the curve shape would be likely to change over time, if available water storage were to increase or decrease substantially. Given the limited period of record, climatological changes are unlikely to have occurred at an extent so as to cause large, statistically significant shifts in elasticity curve shape. However, interannual climate variability and long-term climate trends are likely to have some influence on elasticity.

7.1.3 *Streamflow sensitivity to precipitation shows large inter-annual and spatial variability*

In the final research chapter, I explored the effects of interannual climate variability and long-term trends in climate on streamflow elasticity to precipitation for low, median, and high flows. Using a sample of 2967 catchments in the United States, I fit a fixed effects panel regression model to estimate interannual changes in elasticity at the scale of hydroclimatic regions. The regional approach was taken so that interannual shifts in elasticity could be attributed to regional scale processes, namely climate. The questions addressed in this paper arose from the work done in chapter 5, and from the underlying assumptions of stationarity which are present in many empirical studies on streamflow elasticity.

I found large interannual variability in elasticity in some regions and very limited interannual elasticity in others. The regions with more variable elasticity were those with drier climates, and less seasonally consistent flow regimes. I found some statistically significant long-term trends in elasticity, but the total changes were small. Further, many of these trends could

potentially be explained by dam regulation, rather than climatology. It is well accepted in hydrology that contemporary climatic conditions effect runoff rate and thus should have some influence on the sensitivity of streamflow to climatological variables. However, empirical studies which use, estimate, or assess elasticity frequently assume that a single estimate averaged over the period of record is sufficient to capture hydrological response. This is often the case even using single-site timeseries models with periods of record which are quite short. This work represents an important attempt to capture temporal variability in streamflow elasticity to precipitation both through interannual variability and long-term trends. I highlight the importance of considering climate variability and long-term trends when assessing streamflow sensitivity.

7.2 Binding themes and key takeaways

The three research chapters presented in this thesis and summarized above all seek to assess the potential drivers of hydrologic non-stationarity. This work represents an empirical attempt to quantify the effects of two main categories of drivers, precipitation and land cover change, on streamflow magnitude in an average, generalizable context.

7.2.1 Large-sample hydrology, panel regression, and causal inference

The research components share a series of methodological approaches, as well as a common conceptual underpinning. As described by Addor et al. (2020) and in the introductory chapters of this thesis, large-sample hydrology seeks to leverage large datasets to learn from the similarities and differences of many catchments, ultimately with the goal of drawing generalizable conclusions about hydrological processes. Conceptually, my work here is aligned with these goals. I use statistical tools and significance testing to try to derive these relationships as suggested by Gupta et al. (2014).

The results of my work indicate that some generalizability is possible, particularly relating to relationships between climate and streamflow. For instance, in chapter 5, I found that in aggregate, the vast majority of variation in elasticity curve shape could be explained by the characteristics of the catchments and hydrologic signatures which related to storage and delayed sources of flow. In chapter 6, using an aggregated regional spatial scale allowed us to explore the average interannual effects of climate on streamflow elasticity to precipitation. However, the work presented in this thesis also highlights the uniqueness of place and the necessity of nuance (Beven, 2000a) when aiming to understand more variable processes such as land cover change, or how, specifically, these processes may effect individual catchments.

For example, the research presented in chapter 4 is heavily caveated, because the nature of the relationships explored requires a high degree of nuance – it is likely that land cover change will affect streamflow differently in different locations. For instance, while I find a consistent and statistically significant *average* effect of urbanisation on streamflow across the entire United States, the actual effect at the scale of individual catchments is potentially quite different from the average. Some characteristics, like impervious surface cover, are common across all urbanizing areas, leading to a significant positive relationship. However, this relationship is still dependent to some degree, on the specific characteristics of the landscape in question. In other words, depending on anthropogenically driven traits such as water management practices, infrastructure, or population density, as well as climate characteristics like aridity, the effects of urbanisation on runoff could reasonably differ in magnitude. The site-specific variation is less influential for climate impacts, which typically occur at a large scale and have a fairly large and consistent influence on flow. These features do interact with the landscape and effect streamflow differently depending on contemporary and lagged conditions, as demonstrated in chapter 6, however, and it would be instructive to perform a deeper analysis into the interactions between land cover, groundwater storage changes, or other catchment scale changes with precipitation elasticity of streamflow.

All three research chapters share one common methodological element: the application of panel regression models. This approach, which is explored in detail in chapter 4, is relatively new in hydrology; the first application in the discipline, to our knowledge, occurred in 2013 (Steinschneider et al., 2013). Steinschneider et al. (2013) explored fixed effects and random effects panel regression models for estimating the influence of urbanisation on flow in comparison to more typical cross sectional and timeseries regression models using a very small case study of 19 catchments. Bassiouni et al. (2016) also compared panel regression models of different types with cross-sectional "space-for-time" models for estimating the sensitivity of low streamflow to precipitation in 86 Hawaiian catchments. Both highlight the capacity for panel regression models to overcome some of the limitations of traditional regression, especially in regards to managing omitted variable bias and multicollinearity. I explore this concept in depth in chapter 4, expanding the conclusions of previous papers to larger, more variable, and noisy datasets, demonstrating in a similar fashion the robustness of the methodology, as well as highlighting its potential weaknesses and areas where it may be unnecessary. In chapter 6, I apply a non-stationary regional version of this model type. Overall,

my work demonstrates how the panel regression model approach carries enormous potential for applications in large sample hydrology.

Finally, throughout this thesis, I explore the concept of causal inference. In each of the papers which comprise this thesis, I use a series of regression approaches to mathematically model the relationships between streamflow magnitude and different environmental variables. I use panel regression models which are designed to be causally interpretable, and single-site regression models which are not. Because panel regression models facilitate the control of various confounding relationships, they can be used in conjunction with subject-specific expertise and strict assumptions to isolate causal relationships. If an omitted variable can be represented by a proxy, panel regression models can be used for causal inference (Nichols, 2007). The idea is to use an individual (a hydrological catchment in the case of my thesis) as its own "control group" by including temporally variant data. Beyond this, panel models require confounding variables to either be measured, or to be invariant across one dimension, e.g. time.

While the approaches taken throughout this thesis offer a degree of robustness and a sincere attempt at causal inference, the approaches are not without limitations. For instance, Ferraro et al. (2019) note that in the context of coupled human-natural systems, the validity of a causal claim relies on the assumptions of excludability and no interference. Excludability refers to the idea that the drivers of variability in the "treatment" have no effect on the outcome (change in streamflow) other than through variability in the treatment itself. Treatment refers to the causal variable of interest e.g. precipitation variation or land cover change. The "no interference" assumption, also known as the stable-unit-treatment-value assumption (Nichols, 2007), boils down to an assumption of independence between samples -- e.g. changes in precipitation in one catchment do not affect streamflow in a second catchment. It is easy to imagine instances in which both of these assumptions, but especially the excludability assumption, may be violated within the remit of the questions asked by this thesis. Further, in the context of research which is based upon observed data and not experimental results, claims of causality are only as good as the assumptions guiding the models used to assess these relationships. For these reasons, among others, causal inference using secondary observed data can be difficult, and requires both explicit assumptions and an abundance of caution. Further, it is worth considering how connected causal chains could be examined in this context, as the hydrologic cycle is an interconnected dynamical system.

7.2.2 *Streamflow sensitivity*

In this thesis I focus primarily on assessing the sensitivity of streamflow to environmental changes, namely precipitation and land cover change. In addition to methodological advancements, my work contributes to the bank of empirically-based knowledge on streamflow sensitivity, offering new insights into these relationships. As pointed out by Andréassian et al. (2016), most studies on streamflow sensitivity are "theoretical" in that they are based on the simulated outputs of hydrologic models, or analytical in that they are based off of water balance equations.

Theoretical approaches such as these have many advantages. For instance, they can circumvent the issue of covariation of precipitation and temperature, can model rare and extreme scenarios, and allow for controlled variation of drivers. In contrast, hydrological models are constructed based on simplified representations of our current understanding of reality. They require empirical data for validation (Andréassian, Coron, et al., 2016), and meaningful causal interpretation can be difficult, as similar outputs can be achieved with very different model parameterization. The sensitivities explored in my work are based on simple, but relatively robust, empirical tools and widely available data. In this way they are accessible and useful to support policy and management decision making. Further, because of their empirical basis, they might additionally serve as benchmarking and calibration tools for theoretical approaches. The panel regression approach and potential for regional models also means that future work based on this thesis could have the potential for applicability to ungauged basins as in e.g. Bassiouni et al. (2016), although this application was not explored herein. These results are also limited in some ways, including that my method only allows for estimation of elasticity in perennial rivers, observed data must be available, and that despite best efforts, it is still possible that some omitted variable bias or noise in the data may affect results (as discussed in Ch. 4). Analytical approaches such as those based on water balance equations may present an opportunity to circumvent some of these issues.

My exploration of streamflow sensitivity across the flow distribution and over time in this thesis begins to capture some fundamental principles of hydrology in an empirical context. For instance, the elasticity curve concept captures some of the variation in the ways in which rivers intersect with the landscape through the processes of flow generation and precipitation partitioning, in a predictable way, regardless of the sources of hydrologic storage. Further, within hydrology, we have known for many years that streamflow sensitivity is not constant. In fact, in one of the foundational papers on streamflow elasticity, Sankarsubramanian et al.

(2001) stated that, "the sensitivity of streamflow to climate is itself a dynamic quantity which may change as climate changes." However, despite this, many empirical studies make assumptions of stationarity in the interest of simplicity. My work in chapter 6 begins to address these physical inconsistencies.

7.3 Outlook and future work

While this work offers new insights into streamflow sensitivity, it is far from a comprehensive assessment of these relationships. Many questions remain. One of the primary questions, is how information regarding non-stationarity and variable streamflow elasticity can be incorporated into modelling and decision-making processes. Bayazit (2015) outlined a series of problems with non-stationarity in engineering practice. Their questions boiled down to essentially: How do we detect non-stationarity, and once detected, what can we do about it?

My work has focussed entirely on understanding the drivers of change. This an essential step towards developing the ability to project changes into the future, because without understanding of the physical drivers of change, resorting to a non-stationary model may reduce a model's predictive performance and increase uncertainty; a stationary model may be preferable to a non-stationary one when the change in time cannot be reliably predicted (Bayazit, 2015).

Assuming that trends will continue linearly into the future is problematic. It is instead possible that non-stationary streamflow regimes might reach a new equilibrium at some point in the future, or that some non-stationarity visible in the historical record is the product of long-term persistence, thus representing a prolonged natural fluctuation in climate norms, but not a change which will exist in perpetuity (Bayazit, 2015). In other words, stationarity may be re-established relative to a different mean than we have historically relied upon. Our ability to project these changes and their effects with some degree of reliability requires a solid physical explanation for the drivers of change. For these reasons, the non-stationary model explored in chapter 6 could be developed further. My approach captures the importance of annual conditions for variability in streamflow response, but relates these to physical processes in a manner which is incomplete – the drivers of the changes cannot be incorporated directly into the model its current form. For this reason, despite being informative, any projection based on this model would be difficult. Further, my models assume linearity in the historical relationship, as well as in trend projection, a reasonable, but physically unrealistic, assumption. Thus, the research presented in chapter 6 offers many avenues for future research. Further, my

focus in this chapter was on climate and a regional approach was taken; however, changes in land surface characteristics such as groundwater (Berghuijs & Slater, 2023) should also be explored as drivers of long-term shifts in streamflow sensitivity to precipitation.

My research suggests that elasticity curve shape is determined in large part by the availability of hydrologic inputs from stored sources. Deeper exploration of this concept is needed to develop a comprehensive picture of the role of storage for hydrologic responsiveness to precipitation. Flat curves (where all flow quantiles respond to precipitation in a similar way) appear to occur when precipitation and streamflow are out of season with one another, or when there may be a substantial soil moisture deficit, although these relationships require additional investigation. For example, in this work, I relate elasticity curve shape predominantly to hydrologic signatures, many of which are poorly linked to underlying physical properties. A deeper investigation into the connections of physical properties e.g. soil moisture deficit, and the elasticity curve, could provide an improved understanding of hydrological sensitivity, and increased predictability of change.

The work on land cover change presented in Chapter 4 represents a simplified approach to a series of extremely challenging questions. Some limitations and future directions were discussed in the chapter itself; however, one primary challenge is that I estimated only national average effect of the land cover changes on flow. Considering that evidence suggests that effects may vary depending on climatology, and that characteristics such as forest cover type may have a strong moderating effect on hydrologic response, it would be instructive to examine these relationships in a more comprehensive manner.

7.4 Conclusions

The work presented in this thesis has helped to advance understanding of the relationships between potentially changing environmental variables. I have employed methodological approaches which are relatively new to hydrology, and assessed the statistical robustness of panel regression models relative to a single-site timeseries approach. My work comparing these approaches in the context of small effect sizes and noisy data demonstrates the improved robustness of panel regression, while highlighting some of the downsides of the method.

My research on the influence of land cover changes on streamflow captures the variability of hydrologic responses to changes in tree cover across the flow regime, something which has been noted in previous research, but is still slowly making its way into the realm of common understanding in the field. These results highlight the need for caution and nuance when aiming

to use forest cover as a strategy for flood risk management. My results on urbanisation corroborate the common expectation that urbanisation increases streamflow for higher flow percentiles, and offer another line of evidence that response for low flows may be inconsistent.

In the final two research chapters, I explore precipitation effects on streamflow instead of land cover. The elasticity curve is an entirely novel concept and represents one of very few explorations of elasticity across multiple flow regimes in the literature. Broadly speaking, this work reveals the utility of using simple empirical approaches for investigating hydrological behaviour at a large scale. The regional cohesion and explainability of the elasticity curve shapes is encouraging regarding the potential for further development and applicability of the concept.

My work on non-stationarity of elasticity is a first step towards acknowledging and addressing a key issue with hydrologic projection based on empirically driven sensitivity estimates. My application of a non-stationary regression model captures regional differences in interannual variability and highlights the need for more work in this area.

Overall, this thesis has been an ode to using statistical tools to understand the natural world; to the marriage of some of the things that bring me the greatest joy: the intersections of science, mathematics and art. This work has challenged existing paradigms in the field, offered new lines of evidence for long-held hypotheses, and generated pathways for future research into the complex dynamics at play between hydrology and the environment.

8   Appendices

8.1   Appendix 1: Water cycle diagram
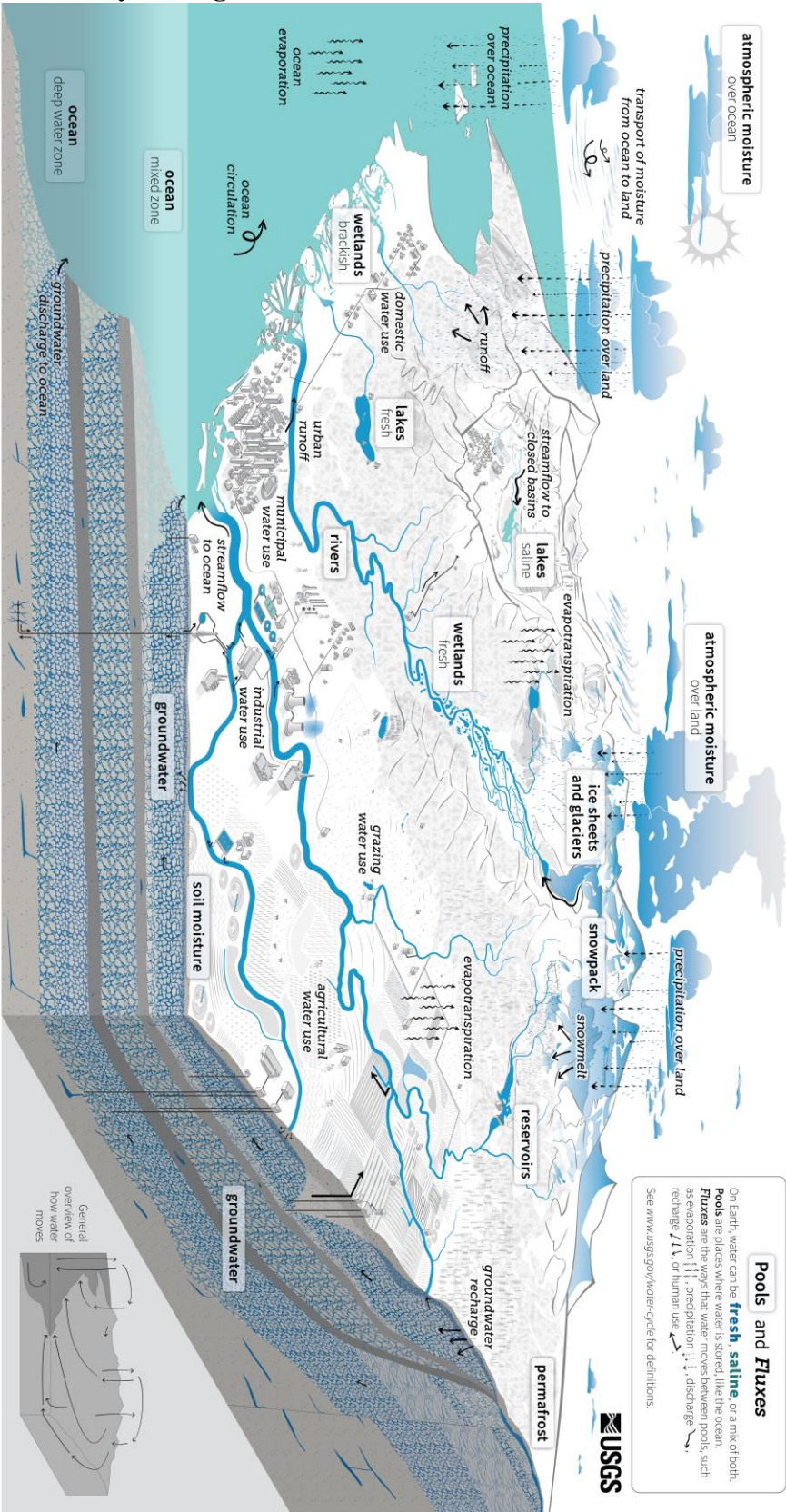
The Water Cycle

The water cycle describes where water is found on Earth and how it moves. Water can be stored in the atmosphere, on Earth's surface, or below the ground. It can be in a liquid, solid, or gaseous state. Water moves between the places it is stored at large scales and at very small scales. Water moves naturally and because of human interaction, both of which affect where water is stored, how it moves, and how clean it is.

Liquid water can be fresh, saline (salty), or a mix (brackish). Ninety-six percent of all water is saline and stored in **oceans**. Places like the ocean, where water is stored, are called **pools**. On land, saline water is stored in liquid form in **saline lakes**, whereas fresh water is stored in liquid form in **freshwater lakes**, artificial **reservoirs**, **rivers**, **wetlands**, and in soil as **soil moisture**. Deeper underground, liquid water is stored as **groundwater** in aquifers, within the cracks and pores of rock. The solid, frozen form of water is stored in **ice sheets**, **glaciers**, and **snowpack** at high elevations or near the Earth's poles. Frozen water is also found in the soil as **permafrost**. Water vapor, the gaseous form of water, is stored as **atmospheric moisture** over the ocean and land.

As it moves, water can transform into a liquid, a solid, or a gas. The different ways in which water moves between pools are known as **fluxes**. **Circulation** mixes water in the oceans and transports water vapor in the atmosphere. Water moves between the atmosphere and the Earth's surface through **evaporation**, **evapotranspiration**, and **precipitation**. Water moves across the land surface through **snowmelt**, **runoff**, and **streamflow**. Through infiltration and **groundwater recharge**, water moves into the ground. When underground, groundwater flows within aquifers and can return to the surface through **springs** or from natural **groundwater discharge** into rivers and oceans.

Humans alter the water cycle. We redirect rivers, build dams to store water, and drain water from wetlands for development. We use water from rivers, lakes, reservoirs, and groundwater aquifers. We use that water (1) to supply our **homes** and **communities**, (2) for **agricultural** activities like **grazing** livestock, and (3) in **industrial** activities like thermoelectric power generation, mining, and aquaculture. The amount of available water depends on how much water is in each pool (water quantity). Water availability also depends on when and how fast water moves (water timing), and how much water is used (water use), and how clean the water is (water quality).

Human activities affect **water quality**. In agricultural and urban areas, irrigation and precipitation wash fertilizers and pesticides into rivers and groundwater. Power plants and factories return heated and contaminated water to rivers. Runoff carries chemicals, sediment, and sewage into rivers and lakes. Downstream from these types of sources, contaminated water can cause harmful algal blooms, spread diseases, and harm habitats. **Climate change** is also affecting the water cycle. It affects water quality, quantity, timing, and use. Understanding these impacts can allow progress toward sustainable water use.

Pools and Fluxes

On Earth, water can be **fresh**, **saline**, or a mix of both. **Pools** are places where that water is stored, like the ocean. **Fluxes** are the ways that water moves between pools, such as evaporation, precipitation, discharge. See www.usgs.gov/water-cycle for definitions.

Labels: atmospheric moisture over ocean; transport of moisture from ocean to land; precipitation over ocean; ocean evaporation; ocean circulation; ocean mixed zone; ocean deep water zone; groundwater discharge to ocean; wetlands brackish; domestic water use; urban runoff; runoff; lakes fresh; streamflow to closed basins; lakes saline; precipitation over land; municipal water use; rivers; wetlands fresh; evapotranspiration; streamflow to ocean; industrial water use; groundwater; atmospheric moisture over land; ice sheets and glaciers; grazing water use; snowpack; snowmelt; precipitation over land; reservoirs; agricultural water use; evapotranspiration; soil moisture; groundwater; groundwater recharge; permafrost; General overview of how water moves.

Figure 8.1 Water cycle diagram from literature review presented in larger format. Source:

(Corson-Dosch et al., 2022)

Figure Appendix 2

8.1.1 Supplementary materials for Chapter 4

Supplementary information is presented in the format which it was submitted to the journal.

Introduction

This document includes text and tables which are intended for reference and to improve reproducibility.

Text 8.1.

The estimated effect of urbanisation, tree cover change, and temperature on streamflow in % is calculated $100*(e^\beta - 1)$ where β is the estimated average effect either tree cover change, urbanisation or temperature on a given streamflow quantile. The 90% confidence intervals are calculated as: $100*(e^{\beta \pm 1.645*SE} - 1)$ where SE is the robust standard error clustered by streamgauge for the panel models, and the standard error in the GLMs.

Text 8.2.

We estimate the standard errors in the panel regression models by clustering them at the streamgauge level using the following R script:

coeftest(Model, vcov=vcovHC(Model,type="HC0",cluster="group"))

Where "Model" is the panel model. "HC0" uses White's estimator to estimate the heteroscedasticity consistent covariance matrix of the coefficient estimates in the regression model, and "group" clusters by streamgauge.

Text 8.3.

In addition to the Panel model formulation, we also tested a version of the urbanisation panel models which did not include any climatological variables because they are not confounders urbanisation. The following equation was tested:

$$ln(Y_{r(i),t}) = \alpha_i + \beta_1 urban_{i,t} + \delta_t D_t + \gamma_{t,r} D_t D_r + \varepsilon_{i,t} \qquad 8.1$$

where all coefficients and variables are the same as those in the main text except that precipitation and temperature have been removed. The resulting coefficients (displayed in table S2) for urban area are effectively the same as those for the main text model which includes climate variables. This demonstrates the point that precipitation and temperature are not confounders for urbanisation in this context.

Table 8.1 Supplementary information: Description of aggregation of land cover classes from the ESA-CCI global land cover dataset (ESA CCI, 2017)

ID	Original Classification	New Classification
0	No Data	No Data
10	Cropland, rainfed	Cropland
11	herbaceous cover	Cropland
12	Tree or shrub cover	Cropland
20	Cropland, irrigated or post-flooding	Cropland
30	Mosaic cropland (> 50%) / natural vegetation (tree, shrub, herbaceous cover) (< 50%)	Cropland
40	Mosaic natural vegetation (tree, shrub, herbaceous cover) (> 50%) / cropland (< 50	Cropland
50	Tree cover, broadleaved, evergreen, closed to open (> 15%)	Tree cover
60	Tree cover, broadleaved, deciduous, closed to open (> 15%)	Tree cover
61	Tree cover, broadleaved, deciduous, closed (> 40%)	Tree cover
62	Tree cover, broadleaved, deciduous, open (15-40%)	Tree cover
70	Tree cover, needleleaved, evergreen, closed to open (> 15%)	Tree cover
71	Tree cover, needleleaved, evergreen, closed (> 40%)	Tree cover
72	Tree cover, needleleaved, evergreen, open (15-40%)	Tree cover
80	Tree cover, needleleaved, deciduous, closed to open (> 15%)	Tree cover
81	Tree cover, needleleaved, deciduous, closed (> 40%)	Tree cover
82	Tree cover, needleleaved, deciduous, open (15-40%)	Tree cover
90	Tree cover, mixed leaf type (broadleaved and needleleaved)	Tree cover
100	Mosaic tree and shrub (> 50%) / herbaceous cover (< 50%)	Tree cover
110	Mosaic herbaceous cover (> 50%) / tree and shrub (< 50%)	Cropland
120	Shrubland	Grassland
121	Evergreen shrubland	Grassland
122	Deciduous shrubland	Grassland
130	Grassland	Grassland
140	Lichens and mosses	Grassland
150	Sparse vegetation (tree, shrub, herbaceous cover) (< 15%)	Grassland
151	Sparse tree (< 15%)	Grassland
152	Sparse shrub (< 15%)	Grassland
153	Sparse herbaceous cover (< 15%)	Grassland
160	Tree cover, flooded, fresh or brackish water	Tree cover
170	Tree cover, flooded, saline water	Tree cover
180	Shrub or herbaceous cover, flooded, fresh/saline/brackish water	Wetland
190	Urban areas	Urban areas
200	Bare areas	Bare
201	Consolidated bare areas	Bare
202	Unconsolidated bare areas	Bare

210	Water bodies	Water bodies
220	Permanent snow and ice	Bare

Table 8.2 Supplementary information: Comparison of the coefficients for urbanisation in the panel model inin the main text, which includes climate variables, and the panel model described in text 8.3, which does not include climate variables.

Q	Model	Variable	Estimated change (%)	P value	Lower	Upper
0.01	Main model	Urban area	0.269	0.5366	-0.446	0.989
Qmean	Main model	Urban area	0.575	0.0003	0.314	0.837
0.99	Main model	Urban area	0.741	0.0014	0.359	1.125
0.01	NO climate	Urban area	0.208	0.6252	-0.491	0.913
Qmean	NO climate	Urban area	0.491	0.0019	0.231	0.752
0.99	NO climate	Urban area	0.650	0.004	0.278	1.024

8.2 Appendix 3

8.2.1 Appendices from chapter 5
8.2.1.1 Panel model design

In order to further validate the elasticity estimates, we constructed a fixed-effects panel regression model (Equation 8.2) for each timescale ($\varepsilon_{c,P}^{g,q}$). The panel models were designed to control for confounding variables, and the clusters established from the lm results were included as interaction terms to help explain variation in elasticity curve shape. A confounding variable is an attribute of a catchment or group of catchments which could influence both the dependent variable and independent variable, causing a spurious association.

Time-invariant confounders at the catchment scale are controlled for by the stream gauge-specific intercept α_i. At the timescale of this study (30-39 years of data per site), the majority of confounding variables at the catchment scale may be reasonably expected to be time-invariant (e.g. topography). While some land cover changes are likely over the time period, a minority of catchments are likely to have experienced large percentages of detectable land cover change, and, when considered jointly in a panel model, the effects of land cover changes on streamflow are likely to be small relative to climatic effects (Anderson et al., 2022). Variables such as temperature and actual evapotranspiration are partially or fully considered

through the calculation or inclusion of other variables. More complex formulations of the panel model, which explicitly included eco-regions and/or a control for time varying confounders at the national scale were considered, however, the resulting curves were not substantially different from one another, and thus the simplest model (Equation 8.2) is used. The panel model is represented by:

$$ln\left(Q_{i,t}^q\right) = \alpha_{i,t} + \beta_1 \, ln\left(P_{i,t}\right) + \beta_2 \, ln\left(E_{i,t}\right) + \varepsilon_P^{g,q} \, ln\left(P_{i,t}\right) g_i + \varepsilon_E^{g,q} \, ln\left(E_{i,t}\right) g_i + \eta_{i,t}^q \qquad 8.2$$

where $ln\left(Q_{i,t}^q\right)$ is the natural logarithm of a streamflow percentile (q) calculated for time period (t) for catchment (i), $\alpha_{i,t}$ is the streamgauge-specific intercept, $ln\left(P_{i,t}\right)$ is the logarithm of catchment averaged daily precipitation, and $ln\left(E_{i,t}\right)$ is the logarithm of catchment averaged daily potential evaporation. The elasticity curve cluster for each catchment is represented by a categorical variable (g), and $ln\left(P_{i,t}\right)g_i$ and $ln\left(E_{i,t}\right)g_i$ are interaction terms between the assigned cluster and precipitation or potential evaporation. Precipitation elasticity, the effect measured by this model, is represented by the regression coefficient: $\varepsilon_P^{g,q}$ and potential evaporation elasticity is represented by $\varepsilon_E^{g,q}$. The error term is $\eta_{i,t}^q$. Autocorrelation in fixed effects panel models can lead to the underestimation of standard errors. We address this concern by clustering standard errors at the streamgauge level as in Anderson et al. (2022). The panel regression results are normalized following the same procedure as the lms – by subtracting $\varepsilon_P^{g,0}$ from each $\varepsilon_P^{g,q}$ value.

The panel regression models are included as a more robust method of estimation and as a tool for confirming the results of the individual regression models. The results of these models were not included in the main text because they do not differ substantially from the simpler regression approach. They are included here as appendices because longitudinal regression approaches such as panel regression models are substantially more robust when averages are of interest, and lend credibility to the outcomes of the analysis.

The curves in Figure B1 were produced using the panel regression approach (Equation 8.2) and are plotted with the normalized 95% confidence intervals of the panel model. The panel regression model results in one estimate elasticity value for each percentile, and allows for easy calculation of statistical uncertainty.

The point estimates, $\varepsilon_P^{g,q}$, are all significant at the 99.99% confidence level. The interactions are also significant at the 99.99% confidence level, except for annual streamflow above the

65th percentile, where all interactions are significant at the 95% confidence level, at least, except for the highest annual flow (100th percentile) for which the interaction is not significant. This means that the $\varepsilon_P^{g,q}$ estimates are statistically significantly different from one another for each of the clusters in every temporal scale and every percentile, with the exception of the highest annual streamflow. The actual magnitude of the elasticity estimates for the maximum annual streamflow is not statistically different across the groups (Figure 3 annual: B).

8.2.1.2 Figures and tables in appendices for journal submission of the research presented in Chapter 5

Table 8.3 Description of catchment attributes considered in the explanatory analysis.

Variable	Method	Description
DFI	Smoothed minima method, 90-day window	Calculated using the R package delayed flow: https://modche.github.io/delayedflow/
BFI	Smoothed minima method, 5-day window	Calculated using the R package delayed flow: https://modche.github.io/delayedflow/
Snow fraction		Proportion of precipitation falling in months when the average temperature is below 0
Permeability		Average catchment permeability (mm/hr) (Falcone, 2017).
Aridity index	$\text{Aridity} = \left(\dfrac{\overline{PET}}{\overline{P}}\right) * 100$	The aridity index as a percentage of mean potential evaporation ($\overline{PET}$) divided by mean precipitation ($\overline{P}$).
Runoff Coefficient	$\text{RC} = \dfrac{\overline{Q}/D}{\overline{P}} * 100$	Runoff coefficient estimated as a percentage, where $\overline{Q}$ is mean annual streamflow across the whole time series, D is the drainage area, and $\overline{P}$ is mean precipitation.
fdc$_b$	$\text{fdc}_b = \dfrac{\ln(Q_{33}) - \ln(Q_{66})}{(0.66 - 0.33)}$	Slope of the annual flow duration curve calculated with daily flow between the 33rd and 66th flow exceedance probabilities
fdc$_{bu}$	$\text{fdc}_{bu} = \dfrac{\ln(Q_0) - \ln(Q_{32})}{0.32}$	Slope of the annual flow duration curve calculated with daily flow between the 0th and 32nd flow exceedance probabilities
fdc$_{bl}$	$\text{fdc}_{bl} = \dfrac{\ln(Q_{67}) - \ln(Q_{100})}{(1 - 0.67)}$	Slope of the annual flow duration curve calculated with daily flow between the 67th and 100th flow exceedance probabilities
Annual temperature		Mean annual temperature

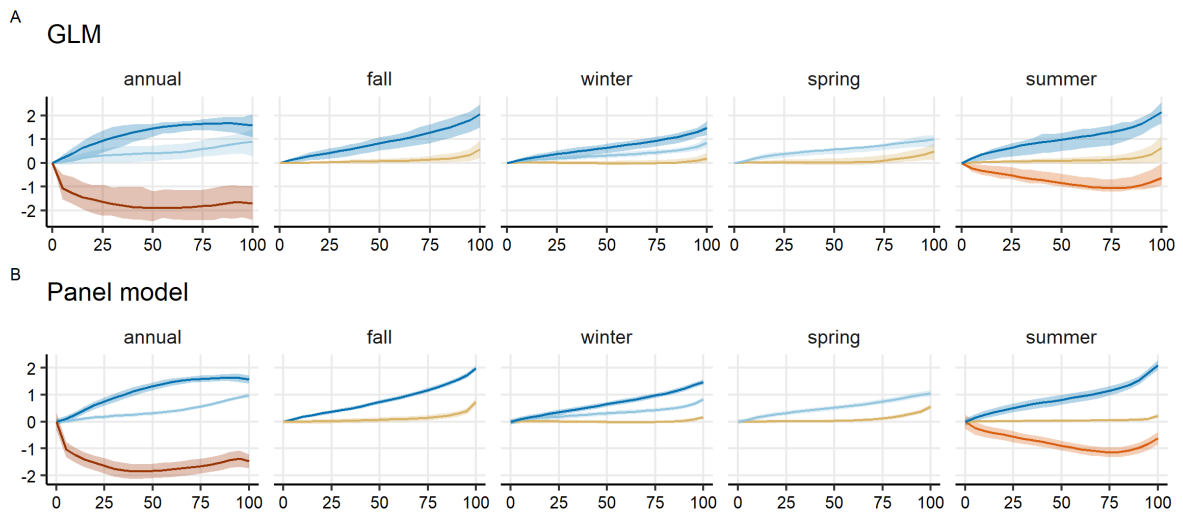|		
Mean catchment elevation		In meters (Falcone, 2017)
Latitude		Latitude at gage site (Falcone, 2017)
Drainage area		In Km^2 (Falcone, 2017)
Average catchment slope		In degrees (Falcone, 2017)
Coefficient of variation	$CV = \dfrac{sd(Q)}{\overline{Q}}$	CV of streamflow -- Calculated in each time step using daily streamflow

Figure 8.2 Elasticity curves as estimated using the single-site regression models (Panel A) and the aggregated panel regression models described in appendix 3 (Panel B). GLMs are presented with the interquartile range of all estimates and panel models are presented with the 95% confidence intervals. Panel A is duplicated from Figure 5.2 here to facilitate comparison.

8.3 Appendix 4

8.3.1 Supplementary materials for Chapter 6

Supplementary information is presented in the format which it was submitted to the journal.

Table 8.4 Supplementary information: Akaike information criterion (AIC) for three different model parameterisations as described in the methods section.
The smallest AIC value (presented in bold font) indicates the best model fit.

Model form	Flow Percentile	AIC
Equation 6.2	**10**	**111897.0**
Equation 6.4	10	137549.3
Equation 6.5	10	119130.2
Equation 6.2	**50**	**79076.41**
Equation 6.4	50	115706.75
Equation 6.5	50	88989.64
Equation 6.2	**90**	**86699.06**
Equation 6.4	90	111952.64
Equation 6.5	90	92972.77

Table 8.5. Supplementary information: Precipitation and streamflow seasonality. Percentage of catchments in each region for which the largest fraction of precipitation, falls in each season and the season with the largest total streamflow, calculated as the seasonal sum of mean daily streamflow.

Bukovsky Region	Precipitation season								Flow season							
	Fall		Winter		Spring		Summer		Fall		Winter		Spring		Summer	
	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
Pacific Northwest	1	0	99	0	0	0	0	100	0	17	83	2	13	0	4	81
Pacific Southwest	0	4	100	0	0	0	0	96	0	69	43	3	43	0	13	28
Mountain West	4	14	66	10	24	3	6	72	1	32	4	59	50	1	45	8
Desert	3	42	70	15	0	33	27	9	6	12	61	6	27	3	6	79
Great Plains	5	6	2	86	37	0	55	9	6	29	1	66	55	1	39	5
Prairie	7	1	0	98	7	0	87	1	0	28	1	71	86	0	13	1
South	14	57	24	19	28	20	34	4	4	62	56	3	35	6	5	29
Great Lakes	50	0	0	96	1	3	48	0	0	48	1	21	98	0	0	31
East	31	32	10	55	23	8	35	4	0	37	13	0	87	0	0	63

Table 8.6 Trend free prewhitened Mann Kendall trend test results for subsample of catchments excluding those with substantial dam storage.

Statistically significant trends (p < 0.05) are presented in bold font. The proportion removed is one minus the number of catchments remaining after dam storage removal divided by the total number of catchments used in the main analysis. Total sites is the absolute number of sites remaining in each region following the dam storage removal process.

			Mann Kendall trend test – Dams removed						
Bukovsky Region	Proportion removed	Total sites	Q10 total change	Q10 p value	Q50 total change	Q50 p value	Q90 total change	Q90 p value
a. Pacific Northwest	0.44	92	-0.04	0.48	0.04	0.19	0.05	0.08
b. East	0.72	300	0.05	0.45	-0.04	0.6	-0.04	0.41
c. South	0.70	101	0.00	1.00	**-0.25**	**<0.01**	**-0.21**	**<0.01**
d. Great Lakes	0.70	61	-0.05	0.74	-0.03	0.78	-0.04	0.66
e. Mountain West	0.68	189	-0.02	0.78	0.01	0.96	**0.17**	**<0.01**
f. Prairie	0.87	35	0.40	0.14	0.13	0.56	-0.11	0.61
g. Pacific Southwest	0.75	26	0.08	0.62	0.13	0.25	0.12	0.27
h. Great Plains	0.90	18	**0.38**	**0.03**	0.01	0.94	-0.03	0.82
i. Desert	0.76	8	-0.12	0.53	**-0.39**	**<0.01**	-0.50	0.29

Table 8.7 Supplementary information: Comparison of baseflow and flow percentiles. Baseflow is estimated using the smoothed minima approach using the R package lfstat (Gauster et al., 2022). This is calculated for the data sample used in the main text (Original data sample) and the sub-sample from which site with substantial dam storage have been removed. Values show the ratio of baseflow (BF) to a given streamflow percentile (Q10/Q50/Q90), indicating how much of the flow hypothetically originates from delayed sources.

Bukovsky Region	Dams Removed			Original data sample		
	BF/Q10	BF/Q50	BF/Q90	BF/Q10	BF/Q50	BF/Q90
a. Pacific Northwest	1.60	0.28	0.08	0.67	0.17	0.04
b. East	1.20	0.37	0.11	0.50	0.14	0.04
c. South	0.51	0.14	0.03	0.21	0.06	0.01
d. Great Lakes	0.71	0.32	0.09	0.25	0.10	0.03
e. Mountain West	0.99	0.49	0.09	0.31	0.14	0.03
f. Prairie	0.37	0.13	0.03	0.11	0.04	0.01
g. Pacific Southwest	2.78	0.80	0.16	0.59	0.22	0.05
h. Great Plains	0.49	0.23	0.10	0.08	0.03	0.01
i. Desert	0.27	0.18	0.07	0.08	0.03	0.01

Text 8.4. The majority of the analysis for this study was carried out in R. The following is a sample code demonstrating how the panel regression model was fit using the plm package(Croissant & Millo, 2008) and how standard errors were clustered at the streamgauge level. This is included in order to facilitate reproducibility. Note that similar, mixed effects models could be fitted using the nlme package(José Pinheiro et al., 2023). Code is indented to indicate where it begins and ends.

```
#load library

library(plm); library(lmtest)

#fit panel model

plm_test <- plm(Q50 ~ P:region:Year + PET: region + BFI:region, data = DF, model = "within", index = c("STAID", "Year"))
```

In this example, we fit the panel regression model in Equation 6.2 for the 50th percentile of flow. *Q50* is the annual timeseries of the 50th percentile of streamflow, *P* is the annual mean daily catchment-averaged precipitation in all catchments, *region* is a factor variable containing the Bukovsky regions (each catchment falls within only one region), *Year* is a factor variable for each year of the timeseries, *PET i*s the annual mean daily catchment-averaged potential evaporation in all catchments, and *BFI* is the baseflow index for every catchment estimated annually. *DF* is a data table or dataframe containing all information used in the model, and STAID is a unique identifier for each catchment.

Once the model is fit, we recalculate the standard errors clustered at the streamgauge level using the *vcovHC* function from the plm package as below:

```
#calculate robust standard errors

test_plm_robust <- coeftest(plm_test,
vcov=vcovHC(plm_test,type="HC0",shift="group"))

sum_test <- summary(plm_test)

#replace the original standard errors in the model results

sum_test$coefficients[,2:4] <- test_plm_robust[,2:4]
```
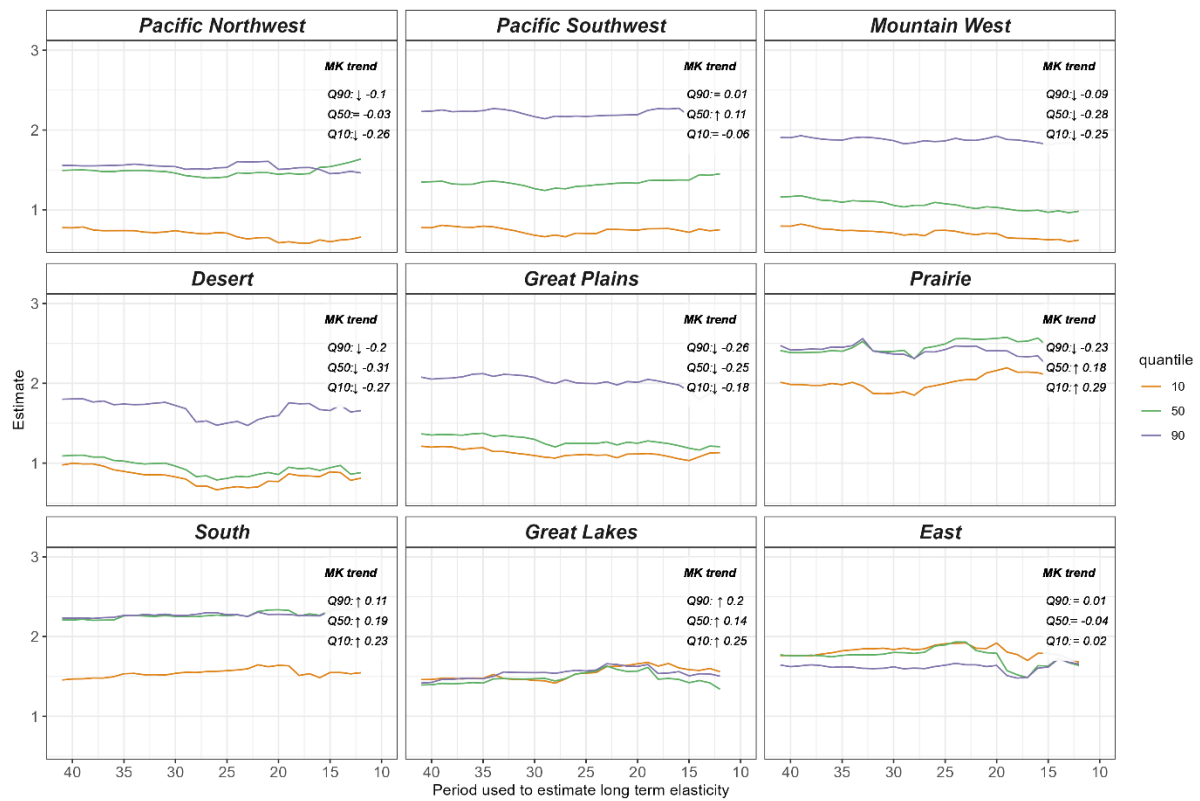
Additional supplement to chapter 6.

Figure 8.3 Regional average precipitation elasticity of streamflow estimating for incrementally shorter periods of time
ranging between 41 and 11 years. MK trend shows whether a statistically significant trend in elasticity occurs as increasingly shorter time periods are used. Results suggest that even when using robust regional regression approaches, elasticity estimates may become unstable when too few years of data are used, depending on the region. The instability is possibly related to the number of catchments in each region of the analysis.
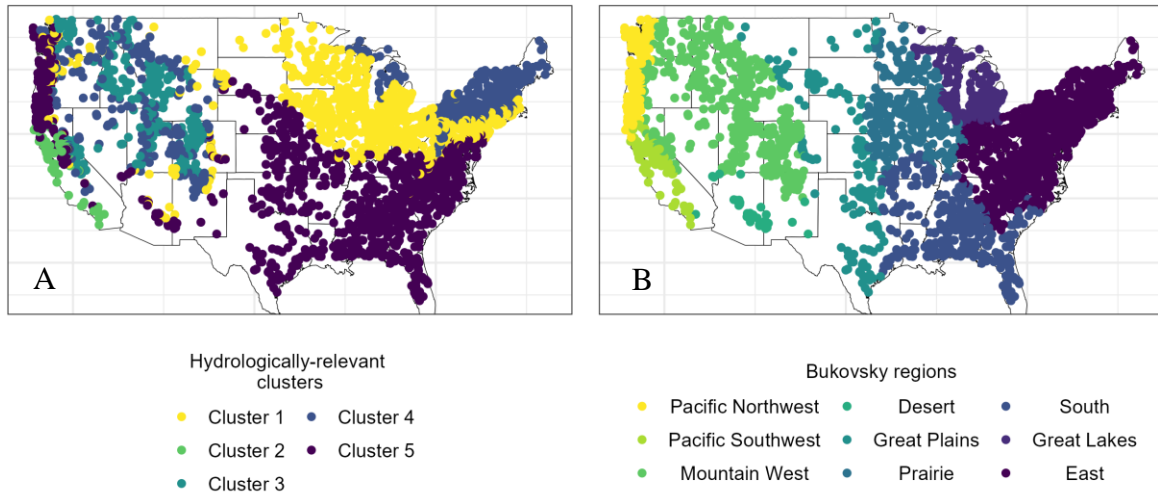
Figure 8.4 Bukovsky regions and "hydrologically-relevant" clusters

Hydrologically-relevant clusters (A) were derived using the method in Knoben et al. (2018). They are calculated based on the snow fraction, aridity index, and seasonality of aridity. These were used to help validate the hydrologic relevance of the Bukovsky regions (B) which are used in this study.
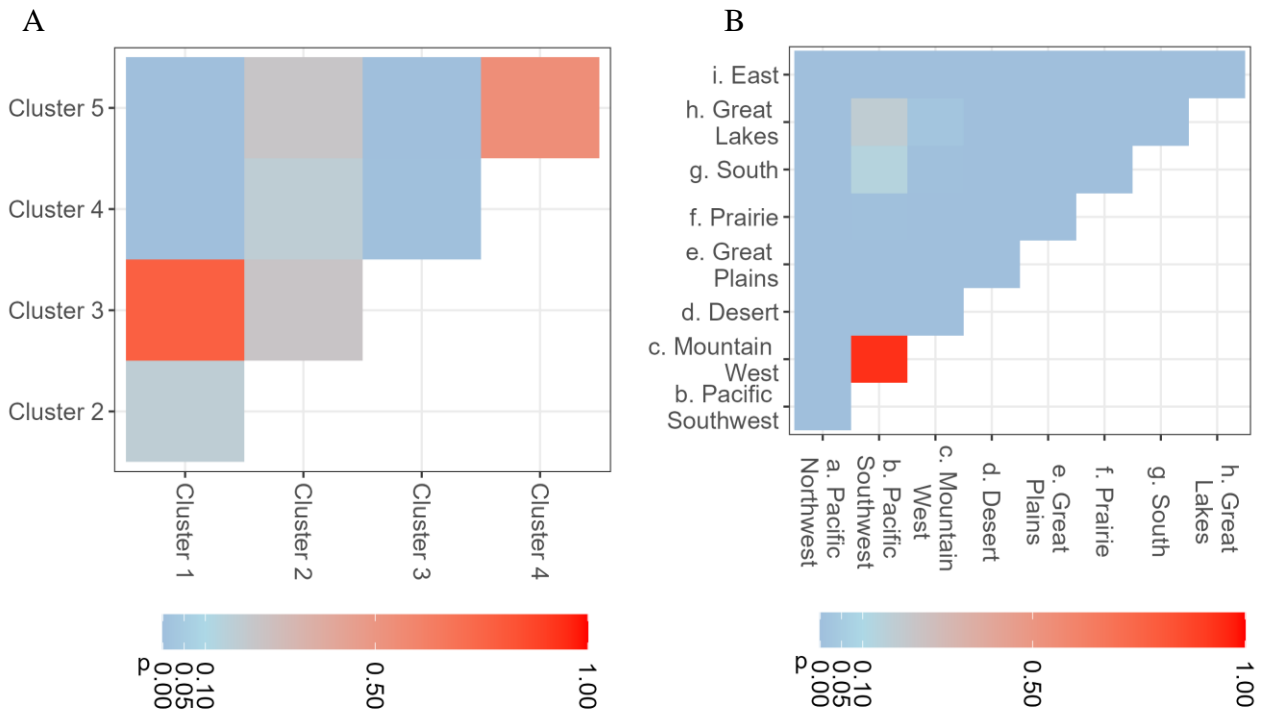
Figure 8.5 Example of differences in hydrologic signatures (runoff coefficient) between Bukovsky regions and "hydrologically-relevant" clusters

A. P-values of pairwise t-tests comparing the runoff ratios of catchments in each cluster estimated using the approach in Knoben et al. (2018). B. P-values of pairwise t-tests comparing the runoff ratios of catchments in each Bukovsky region. The runoff ratio is used as an example here because it is known to be strongly related to elasticity. All hydrologic signatures used in Knoben et al. (2018) were tested and the bukovsky regions consistently performed as well or better than the "hydrologically-relevant" clusters.

8.4 Appendix 5

8.4.1 Combined data availability statement for Chapters 4-6.

All data used in this thesis are publicly available. The land cover data used in chapter 4 are available through ESA-CCI at http://maps.elie.ucl.ac.be/CCI/viewer/download.php (ESA CCI, 2017); climatological data used in all three chapters are available through PRISM at https://prism.oregonstate.edu/recent/ where both mean annual precipitation and temperature can be downloaded; Stream flow data is available from the USGS at http://waterdata.usgs.gov/nwis/ (United States Geological Survey., 2020), and can be bulk downloaded using the R package dataRetrieval. More information on the package is available here: https://code.usgs.gov/water/dataRetrieval ; catchment boundaries used in chapter 4 are available from the USGS and USDA at https://prd-

tnm.s3.amazonaws.com/index.html?prefix=StagedProducts/Hydrography/WBD/National/GDB/ (United States Geological Survey and United States Department of Agriculture., 2020) and catchment boundaries used in chapters 5 and 6 are available through the GAGES II dataset (Falcone, 2011) at https://doi.org/10.3133/70046617; information regarding dam storage is available at https://doi.org/10.5066/F7HQ3XS4 (Falcone 2017). Physiographic regions used in chapter 4 are available at https://water.usgs.gov/lookup/getspatial?physio (Fenneman & Johnson, 1946). Bukovsky regions (Bukovsky, 2011) used in chapter 6 are available from the North American Regional Climate Change Assessment Program at:

https://www.narccap.ucar.edu/data/access.html

8.5 Appendix 6

8.5.1 Co-author attestations

I, Annalise Blum, certify that Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

Statistical Attribution of the Influence of Urban and Tree Cover Change on Streamflow: A Comparison of Large Sample Statistical Approaches

Print Name:

Annalise Blum

Signature: Date: July 24, 2023

I, Ilaria Prosdocimi, certify that Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

Statistical Attribution of the Influence of Urban and Tree Cover Change on Streamflow: A Comparison of Large Sample Statistical Approaches

Print Name: Ilaria Prosdocimi

Signature: *[signature]*   Date: 24 July 2023

I, Jessica Rapson, certify that, to my knowledge, Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

1. Streamflow sensitivity to precipitation shows large inter-annual and spatial variability

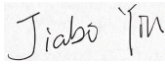Print Name: Jessica Rapson

Signature:

Date: 2023-09-07

xx

I, Jiabo Yin, certify that, to my knowledge, Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

1. Streamflow sensitivity to precipitation shows large inter-annual and spatial variability

Print Name: Jiabo Yin

Signature: *Jiabo Yin*   Date: 2023/9/5

I, Manuela Brunner, certify that, to my knowledge, Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

1. *Elasticity curves describe streamflow sensitivity to precipitation across the entire flow distribution*

2. *Streamflow sensitivity to precipitation shows large inter-annual and spatial variability*

Print Name: Manuela Brunner

Signature: Date: 05.09.2023

I, Marcus Buechel, certify that, to my knowledge, Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

1. Streamflow sensitivity to precipitation shows large inter-annual and spatial variability

Print Name: MARCUS BUECHEL

Signature: Date: 11/09/23

I, Louise Slater, certify that, to my knowledge, Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

1. *Statistical Attribution of the Influence of Urban and Tree Cover Change on Streamflow: A Comparison of Large Sample Statistical Approaches*

2. *Elasticity curves: a novel lens for interpreting the variable nature of the streamflow-precipitation relationship*

3. *Streamflow sensitivity to precipitation shows large inter-annual and spatial variability*

Print Name:

Louise Slater

Signature:

Date: 19 September 2023

I, Simon Dadson, certify that, to my knowledge, Bailey Anderson completed the majority of the work in the following journal articles, which form part of her DPhil thesis:

Statistical Attribution of the Influence of Urban and Tree Cover Change on Streamflow: A Comparison of Large Sample Statistical Approaches

Elasticity curves: a novel lens for interpreting the variable nature of the streamflow-precipitation relationship

Streamflow sensitivity to precipitation shows large inter-annual and spatial variability

Print Name: SIMON DADSON

Signature:

Date: 29/9/23

8.6   Appendix 7: Acronyms and notation

AER -- Antecedent effect ratio

AI – Aridity index

AIC – Akaike information criterion

AMV -- the Atlantic Multi-decadal Variability

BACI – Before After Control Impact

BF or Qb -- Baseflow

BFI – Baseflow index

BIC – Bayesian information criterion

CCI -- Climate Change Initiative (ESA)

CV – coefficient of variation

DAG – Directed Acyclic Graph

DFI – Delayed flow index

E – Actual evaporation

Ep -- Potential evaporation

ESA -- European Space Agency

FDC – Flow duration curve

GAGES II -- Geospatial Attributes of Gages for Evaluating Streamflow version II

GAMLSS – Generalized additive models for location scale and shape

GLM – Generalized linear model

GRACE - Gravity Recovery and Climate Experiment

lm – Linear model

MERIS SR -- Medium resolution imaging spectrometer surface reflectance

NAO -- North Atlantic Oscillation

NbS – Nature based solutions

NHDv1 -- National Hydrography Dataset version 1

NSO -- El Niño–Southern Oscillation

P – Precipitation

PRISM -- Parameter-elevation Regressions on Independent Slopes Model

Q – Streamflow

Qf – Quickflow

RMSE – Root mean squared error

SDG -- Sustainable Development Goals

SF – Snow fraction

SPI – Standardized precipitation index

STA – Standardized temperature anomaly

SWAT -- Soil Water Assessment Tool

T – Temperature

TWS – Total water storage

U.S. – United States

UNCCD -- United Nations Convention to Combat Desertification

V – Vaporisation

VIC -- Variable Infiltration Capacity

VIF -- Variance inflation factor

W – Wetting

WBD -- Watershed Boundary Dataset

Zp – wetting or vaporisation potentials

9 References

Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., Hannah, D. M., Conner, L., Ellison, D., Godsey, S. E., Plont, S., Marçais, J., Kolbe, T., Huebner, A., Frei, R. J., Hampton, T., Gu, S., Buhman, M., Sara Sayedi, S., … Pinay, G. (2019). Human domination of the global water cycle absent from depictions and perceptions. *Nature Geoscience*, *12*(7), Article 7. https://doi.org/10.1038/s41561-019-0374-y

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A Ranking of Hydrological Signatures Based on Their Predictability in Space. *Water Resources Research*, *54*(11), 8792–8812. https://doi.org/10.1029/2018WR022606

Ahn, K.-H., & Merwade, V. (2017). The effect of land cover change on duration and severity of high and low flows. *Hydrological Processes*, *31*(1), 133–149. https://doi.org/10.1002/hyp.10981

Allaire, M. C., Vogel, R. M., & Kroll, C. N. (2015). The hydromorphology of an urbanizing watershed using multivariate elasticity. *Advances in Water Resources*, *86*, 147–154. https://doi.org/10.1016/j.advwatres.2015.09.022

Allan, R. P., Barlow, M., Byrne, M. P., Cherchi, A., Douville, H., Fowler, H. J., Gan, T. Y., Pendergrass, A. G., Rosenfeld, D., Swann, A. L. S., Wilcox, L. J., & Zolina, O. (2020). Advances in understanding large-scale responses of the water cycle to climate change. *Annals of the New York Academy of Sciences*, *1472*(1), 49–75. https://doi.org/10.1111/nyas.14337

Anderson, B. J., Brunner, M. I., Slater, L. J., & Dadson, S. J. (2023). Elasticity curves describe streamflow sensitivity to precipitation across the entire flow distribution. *Hydrology and Earth System Sciences Discussions*, 1–27. https://doi.org/10.5194/hess-2022-407

Anderson, B. J., Slater, L. J., Dadson, S. J., Blum, A. G., & Prosdocimi, I. (2022). Statistical Attribution of the Influence of Urban and Tree Cover Change on Streamflow: A Comparison of Large Sample Statistical Approaches. *Water Resources Research*, *58*(5), e2021WR030742. https://doi.org/10.1029/2021WR030742

Andréassian, V., Coron, L., Lerat, J., & Le Moine, N. (2016). Climate elasticity of streamflow revisited – an elasticity index based on long-term hydrometeorological records. *Hydrology and Earth System Sciences*, *20*(11), 4503–4524. https://doi.org/10.5194/hess-20-4503-2016

Andréassian, V., Mander, Ü., & Pae, T. (2016). The Budyko hypothesis before Budyko: The hydrological legacy of Evald Oldekop. *Journal of Hydrology*, *535*, 386–391. https://doi.org/10.1016/j.jhydrol.2016.02.002

Andréassian, V., & Perrin, C. (2012). On the ambiguous interpretation of the Turc-Budyko nondimensional graph. *Water Resources Research*, *48*(10). https://doi.org/10.1029/2012WR012532

Archer, D. R. (2007). The use of flow variability analysis to assess the impact of land use change on the paired Plynlimon catchments, mid-Wales. *Journal of Hydrology*, *347*(3), 487–496. https://doi.org/10.1016/j.jhydrol.2007.09.036

Archfield, S. A., Hirsch, R. M., Viglione, A., & Blöschl, G. (2016). Fragmented patterns of flood change across the United States. *Geophysical Research Letters*, *43*(19), 10,232-10,239. https://doi.org/10.1002/2016GL070590

Arellano, M. (1987). Computing Robust Standard Errors for Within-groups Estimators. *Oxford Bulletin of Economics & Statistics*, *49*(4), 431–434. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=5173664&site=ehost-live&authtype=ip,uid

Arora, V. K. (2002). The use of the aridity index to assess climate change effect on annual runoff. *Journal of Hydrology*, *265*(1), 164–177. https://doi.org/10.1016/S0022-1694(02)00101-4

Arrigoni, A. S., Greenwood, M. C., & Moore, J. N. (2010). Relative impact of anthropogenic modifications versus climate change on the natural flow regimes of rivers in the Northern Rocky Mountains, United States. *Water Resources Research*, *46*(12). https://doi.org/10.1029/2010WR009162

Baker, D. B., Richards, R. P., Loftus, T. T., & Kramer, J. W. (2004). A New Flashiness Index: Characteristics and Applications to Midwestern Rivers and Streams1. *JAWRA Journal of the American Water Resources Association*, *40*(2), 503–522. https://doi.org/10.1111/j.1752-1688.2004.tb01046.x

Barlow, P. M., & Leake, S. A. (2012). *Streamflow depletion by wells: Understanding and managing the effects of groundwater pumping on streamflow* (Vol. 1376). US Geological Survey Reston, VA.

Barros, A. P., Duan, Y., Brun, J., & Medina, M. A. (2014). Flood Nonstationarity in the Southeast and Mid-Atlantic Regions of the United States. *Journal of Hydrologic Engineering*, *19*(10), 05014014. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000955

Bart, R. R., Tague, C. L., & Moritz, M. A. (2016). Effect of Tree-to-Shrub Type Conversion in Lower Montane Forests of the Sierra Nevada (USA) on Streamflow. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0161805

Bassiouni, M., Vogel, R. M., & Archfield, S. A. (2016). Panel regressions to estimate low-flow response to rainfall variability in ungaged basins. *Water Resources Research*, *52*(12), 9470–9494. https://doi.org/10.1002/2016WR018718

Bastin, J.-F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C. M., & Crowther, T. W. (2019). The global tree restoration potential. *Science*, *365*(6448), 76–79. https://doi.org/10.1126/science.aax0848

Bayazit, M. (2015). Nonstationarity of Hydrological Records and Recent Trends in Trend Analysis: A State-of-the-art Review. *Environmental Processes*, *2*(3), 527–542. https://doi.org/10.1007/s40710-015-0081-7

Bayazit, M., & Önöz, B. (2007). To prewhiten or not to prewhiten in trend analysis? *Hydrological Sciences Journal*, *52*(4), 611–624. https://doi.org/10.1623/hysj.52.4.611

Beck, H. E., van Dijk, A. I. J. M., Miralles, D. G., de Jeu, R. A. M., (Sampurno) Bruijnzeel, L. A., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, *49*(12), 7843–7863. https://doi.org/10.1002/2013WR013918

Bennett, B., Leonard, M., Deng, Y., & Westra, S. (2018). An empirical investigation into the effect of antecedent precipitation on flood volume. *Journal of Hydrology*, *567*, 435–445. https://doi.org/10.1016/j.jhydrol.2018.10.025

Berghuijs, W. R., Gnann, S. J., & Woods, R. A. (2020). Unanswered questions on the Budyko framework. *Hydrological Processes*, *34*(26), 5699–5703. https://doi.org/10.1002/hyp.13958

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The Relative Importance of Different Flood-Generating Mechanisms Across Europe. *Water Resources Research*, *55*(6), 4582–4593. https://doi.org/10.1029/2019WR024841

Berghuijs, W. R., Larsen, J. R., Emmerik, T. H. M. van, & Woods, R. A. (2017). A Global Assessment of Runoff Sensitivity to Changes in Precipitation, Potential Evaporation, and

Other Factors. *Water Resources Research*, *53*(10), 8475–8486.

https://doi.org/10.1002/2017WR021593

Berghuijs, W. R., Luijendijk, E., Moeck, C., Velde, Y. van der, & Allen, S. T. (2022). Global

Recharge Data Set Indicates Strengthened Groundwater Connection to Surface Fluxes.

Geophysical Research Letters, *49*(23), e2022GL099010.

https://doi.org/10.1029/2022GL099010

Berghuijs, W. R., & Slater, L. J. (2023). Groundwater shapes North American river floods.

Environmental Research Letters, *18*(3), 034043. https://doi.org/10.1088/1748-9326/acbecc

Berghuijs, W. R., & Woods, R. A. (2016). Correspondence: Space-time asymmetry

undermines water yield assessment. *Nature Communications*, *7*(1), Article 1.

https://doi.org/10.1038/ncomms11603

Berghuijs, W. R., Woods, R. A., & Hrachowitz, M. (2014a). A precipitation shift from snow

towards rain leads to a decrease in streamflow. *Nature Climate Change*, *4*(7), Article 7.

https://doi.org/10.1038/nclimate2246

Berghuijs, W. R., Woods, R. A., & Hrachowitz, M. (2014b). A precipitation shift from snow

towards rain leads to a decrease in streamflow. *Nature Climate Change*, *4*(7), Article 7.

https://doi.org/10.1038/nclimate2246

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood

generating mechanisms across the United States. *Geophysical Research Letters*, *43*(9), 4382–

4390. https://doi.org/10.1002/2016GL068070

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust

Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, *119*(1), 249–

275. https://doi.org/10.1162/003355304772839588

Beven, K. (2000a). On uniqueness of place and process representations in hydrology. *Hydrology and Earth Systems Science*, *4*.

Beven, K. (2000b). Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, *4*(2), 203–213. https://doi.org/10.5194/hess-4-203-2000

Biederman, J. A., Harpold, A. A., Gochis, D. J., Ewers, B. E., Reed, D. E., Papuga, S. A., & Brooks, P. D. (2014). Increased evaporation following widespread tree mortality limits streamflow response. *Water Resources Research*, *50*(7), 5395–5409. https://doi.org/10.1002/2013WR014994

Biederman, J. A., Somor, A. J., Harpold, A. A., Gutmann, E. D., Breshears, D. D., Troch, P. A., Gochis, D. J., Scott, R. L., Meddens, A. J. H., & Brooks, P. D. (2015). Recent tree die-off has little effect on streamflow in contrast to expected increases from historical studies. *Water Resources Research*, *51*(12), 9775–9789. https://doi.org/10.1002/2015WR017401

Bladon, K. D., Bywater-Reyes, S., LeBoldus, J. M., Keriö, S., Segura, C., Ritóková, G., & Shaw, D. C. (2019). Increased streamflow in catchments affected by a forest disease epidemic. *Science of The Total Environment*, *691*, 112–123. https://doi.org/10.1016/j.scitotenv.2019.07.127

Bloomfield, J. P., Gong, M., Marchant, B. P., Coxon, G., & Addor, N. (2021). How is Baseflow Index (BFI) impacted by water resource management practices? *Hydrology and Earth System Sciences*, *25*(10), 5355–5379. https://doi.org/10.5194/hess-25-5355-2021

Blöschl, G. (2013). *Runoff prediction in ungauged basins: Synthesis across processes, places and scales*. University Press.

Blöschl, G., Ardoin-Bardin, S., Bonell, M., Dorninger, M., Goodrich, D., Gutknecht, D., Matamoros, D., Merz, B., Shand, P., & Szolgay, J. (2007). At what scales do climate variability and land cover change impact on flooding and low flows? *Hydrological Processes*, *21*(9), 1241–1247. https://doi.org/10.1002/hyp.6669

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., … Zhang, Y. (2019). Twenty-three unsolved problems in hydrology (UPH) – a community perspective. *Hydrological Sciences Journal*, *64*(10), 1141–1158. https://doi.org/10.1080/02626667.2019.1620507

Blöschl, G., & Montanari, A. (2010). Climate change impacts—Throwing the dice? *Hydrological Processes: An International Journal*, *24*(3), 374–381.

Blum, A. G., Ferraro, P. J., Archfield, S. A., & Ryberg, K. R. (2020). Causal Effect of Impervious Cover on Annual Flood Magnitude for the United States. *Geophysical Research Letters*, *47*(5), e2019GL086480. https://doi.org/10.1029/2019GL086480

Boggs, J., Sun, G., & McNulty, S. (2016). Effects of Timber Harvest on Water Quantity and Quality in Small Watersheds in the Piedmont of North Carolina. *Journal of Forestry*, *114*(1), 27–40. https://doi.org/10.5849/jof.14-102

Booij, M. J., Schipper, T. C., & Marhaento, H. (2019). Attributing Changes in Streamflow to Land Use and Climate Change for 472 Catchments in Australia and the United States. *Water*, *11*(5), Article 5. https://doi.org/10.3390/w11051059

Bosch, J. M., & Hewlett, J. D. (1982). A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. *Journal of Hydrology*, *55*(1), 3–23. https://doi.org/10.1016/0022-1694(82)90117-2

Bradshaw, C. J. A., Sodhi, N. S., Peh, K. S.-H., & Brook, B. W. (2007). Global evidence that deforestation amplifies flood risk and severity in the developing world. *Global Change Biology*, *13*(11), 2379–2395. https://doi.org/10.1111/j.1365-2486.2007.01446.x

Brady, A., Faraway, J., & Prosdocimi, I. (2019). Attribution of long-term changes in peak river flows in Great Britain. *Hydrological Sciences Journal*, *64*(10), 1159–1170. https://doi.org/10.1080/02626667.2019.1628964

Brandes, D., Cavallo, G. J., & Nilson, M. L. (2005). Base Flow Trends in Urbanizing Watersheds of the Delaware River Basin. *JAWRA Journal of the American Water Resources Association*, *41*(6), 1377–1391. https://doi.org/10.1111/j.1752-1688.2005.tb03806.x

Brown, A. E., Western, A. W., McMahon, T. A., & Zhang, L. (2013). Impact of forest cover changes on annual streamflow and flow duration curves. *Journal of Hydrology*, *483*, 39–50. https://doi.org/10.1016/j.jhydrol.2012.12.031

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *Journal of Hydrology*, *310*(1), 28–61. https://doi.org/10.1016/j.jhydrol.2004.12.010

Brunner, M. I., Melsen, L. A., Newman, A. J., Wood, A. W., & Clark, M. P. (2020). Future streamflow regime changes in the United States: Assessment using functional classification. *Hydrology and Earth System Sciences*, *24*(8), 3951–3966. https://doi.org/10.5194/hess-24-3951-2020

Brunner, M. I., Swain, D. L., Gilleland, E., & Wood, A. W. (2021). Increasing importance of temperature as a contributor to the spatial extent of streamflow drought. *Environmental Research Letters*, *16*(2), 024038. https://doi.org/10.1088/1748-9326/abd2f0

Brutsaert, W. (2008). Long-term groundwater storage trends estimated from streamflow records: Climatic perspective. *Water Resources Research*, *44*(2). https://doi.org/10.1029/2007WR006518

Budyko, M. (1951). On climatic factors of runoff. *Prob. Fiz. Geogr*, *16*.

Budyko, M. (1974). *Climate and life*. Academic press.

Buechel, M., Slater, L., & Dadson, S. (2022). Hydrological impact of widespread afforestation in Great Britain using a large ensemble of modelled scenarios. *Communications Earth & Environment*, *3*(1), Article 1. https://doi.org/10.1038/s43247-021-00334-0

Buechel, M., Slater, L., & Dadson, S. (2023). Afforestation impacts on terrestrial hydrology insignificant compared to climate change in Great Britain. *Hydrology and Earth System Sciences Discussions*, 1–31. https://doi.org/10.5194/hess-2023-138

Bukovsky, M. S. (2011). Masks for the Bukovsky regionalization of North America. *Regional Integrated Sciences Collective*.

Chappell, N. A., & Tych, W. (2012). Identifying step changes in single streamflow and evaporation records due to forest cover change. *Hydrological Processes*, *26*(1), 100–116. https://doi.org/10.1002/hyp.8115

Chen, J., Theller, L., Gitau, M. W., Engel, B. A., & Harbor, J. M. (2017). Urbanization impacts on surface runoff of the contiguous United States. *Journal of Environmental Management*, *187*, 470–481. https://doi.org/10.1016/j.jenvman.2016.11.017

Chiew, F. (2006). Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Science Journal*, *51*, 612–625. https://doi.org/10.1623/hysj.51.4.613

Chiew, F., Peel, M., McMahon, T., & Siriwardena, L. (2006). Precipitation elasticity of streamflow in catchments across the world. *Climate Variability and Change—Hydrological*

Impacts (Proceedings of the Fifth FRIEND World Conference Held at Havana, Cuba, November 2006), *308*, 256–262. https://api.semanticscholar.org/CorpusID:59492611

Chiew, F., Potter, N. J., Vaze, J., Petheram, C., Zhang, L., Teng, J., & Post, D. A. (2014). Observed hydrologic non-stationarity in far south-eastern Australia: Implications for modelling and prediction. *Stochastic Environmental Research and Risk Assessment*, *28*(1), 3–15. https://doi.org/10.1007/s00477-013-0755-5

Clarke, R. T. (2013). How should trends in hydrological extremes be estimated? *Water Resources Research*, *49*(10), 6756–6764. https://doi.org/10.1002/wrcr.20485

Cohen-Shacham, E., Walters, G., Janzen, C., & Maginnis, S. (2016). Nature-based solutions to address global societal challenges. *IUCN: Gland, Switzerland*, *97*.

Collins, M. J., Hodgkins, G. A., Archfield, S. A., & Hirsch, R. M. (2022). The Occurrence of Large Floods in the United States in the Modern Hydroclimate Regime: Seasonality, Trends, and Large-Scale Climate Associations. *Water Resources Research*, *58*(2), e2021WR030480. https://doi.org/10.1029/2021WR030480

Cooper, M. G., Schaperow, J. R., Cooley, S. W., Alam, S., Smith, L. C., & Lettenmaier, D. P. (2018). Climate Elasticity of Low Flows in the Maritime Western U.S. Mountains. *Water Resources Research*, *54*(8), 5602–5619. https://doi.org/10.1029/2018WR022816

Corson-Dosch, H., Nell, C., Volentine, R., Archer, A. A., Bechtel, E., Bruce, J. L., Felts, N., Gross, T. A., Lopez-Trujillo, D., Riggs, C. E., & Read, E. K. (2022). *The Water Cycle*.

Croissant, Y., & Millo, G. (2008). Panel Data Econometrics in R: The plm Package. *Journal of Statistical Software*, *27*(2). https://doi.org/10.18637/jss.v027.i02

Croissant, Y., & Millo, G. (Eds.). (2018). Endogeneity. In *Panel Data Econometrics with R* (pp. 139–159). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119504641.ch6

Cuo, L. (2016). Land Use/Cover Change Impacts on Hydrology in Large River Basins. In

Terrestrial Water Cycle and Climate Change (pp. 103–134). American Geophysical Union

(AGU). https://doi.org/10.1002/9781118971772.ch6

Dadson, S. J., Hall, J. W., Murgatroyd, A., Acreman, M., Bates, P., Beven, K., Heathwaite,

L., Holden, J., Holman, I. P., Lane, S. N., O'Connell, E., Penning-Rowsell, E., Reynard, N.,

Sear, D., Thorne, C., & Wilby, R. (2017). A restatement of the natural science evidence

concerning catchment-based 'natural' flood management in the UK. *Proceedings of the Royal*

Society A: Mathematical, Physical and Engineering Sciences, *473*(2199), 20160706.

https://doi.org/10.1098/rspa.2016.0706

Daly, C., & Bryant, K. (2013). The PRISM climate and weather system—An introduction.

Corvallis, OR: PRISM Climate Group.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J.,

& Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature

and precipitation across the conterminous United States. *International Journal of*

Climatology, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

Daly, E., Calabrese, S., Yin, J., & Porporato, A. (2019). Linking parametric and water-

balance models of the Budyko and Turc spaces. *Advances in Water Resources*, *134*, 103435.

https://doi.org/10.1016/j.advwatres.2019.103435

Davenport, F. V., Herrera-Estrada, J. E., Burke, M., & Diffenbaugh, N. S. (2020). Flood Size

Increases Nonlinearly Across the Western United States in Response to Lower Snow-

Precipitation Ratios. *Water Resources Research*, *56*(1), e2019WR025571.

https://doi.org/10.1029/2019WR025571

De Niel, J., & Willems, P. (2019). Climate or land cover variations: What is driving observed changes in river peak flows? A data-based attribution study. *Hydrology and Earth System Sciences*, *23*(2), 871–882. https://doi.org/10.5194/hess-23-871-2019

DeCicco, L., Hirsch, R., Lorenz, D., Watkins, D., & Johnson, M. (2024). *dataRetrieval: Retrieval Functions for USGS and EPA Hydrologic and Water Quality Data* (2.7.15) [R]. U.S. Geological Survey. 10.5066/P9X4L3GE

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*(3), 527–546. https://doi.org/10.1007/s00382-010-0977-x

Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K., & Arnold, J. G. (2008). Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. *Journal of Applied Meteorology and Climatology*, *47*(2), 475–497. https://doi.org/10.1175/2007JAMC1356.1

Dijk, A. I. J. M. V., Noordwijk, M. V., Calder, I. R., Bruijnzeel, S. L. A., Schellekens, J., & Chappell, N. A. (2009). Forest–flood relation still tenuous – comment on 'Global evidence that deforestation amplifies flood risk and severity in the developing world' by C. J. A. Bradshaw, N.S. Sodi, K. S.-H. Peh and B.W. Brook. *Global Change Biology*, *15*(1), 110–115. https://doi.org/10.1111/j.1365-2486.2008.01708.x

Dixon, S. J., Sear, D. A., Odoni, N. A., Sykes, T., & Lane, S. N. (2016). The effects of river restoration on catchment scale flood risk and flood hydrology. *Earth Surface Processes and Landforms*, *41*(7), 997–1008. https://doi.org/10.1002/esp.3919

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2010). Can dynamic vegetation information improve the accuracy of Budyko's hydrological model? *Journal of Hydrology*, *390*(1), 23–34. https://doi.org/10.1016/j.jhydrol.2010.06.025

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2011). Assessing the differences in sensitivities of runoff to changes in climatic conditions across a large basin. *Journal of Hydrology*, *406*(3), 234–244. https://doi.org/10.1016/j.jhydrol.2011.07.003

Dooge, J. C. I. (1992). Sensitivity of Runoff to Climate Change: A Hortonian Approach. *Bulletin of the American Meteorological Society*, *73*(12), 2013–2024. https://doi.org/10.1175/1520-0477(1992)073<2013:SORTCC>2.0.CO;2

Douglas, E. M., Vogel, R. M., & Kroll, C. N. (2000). Trends in floods and low flows in the United States: Impact of spatial correlation. *Journal of Hydrology*, *240*(1), 90–105. https://doi.org/10.1016/S0022-1694(00)00336-X

Douville, H., Raghavan, K., Renwick, J. A., Allan, R. P., Arias, P. A., Barlow, M., Cerezo Mota, R., Cherchi, A., Gan, T. Y., Gergis, J., Jiang, D., Khan, A., Pokam Mba, W., Rosenfeld, D., Tierney, J., & Zolina, O. (2021). Water cycle changes. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, Ö. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1055–1210). Cambridge University Press. https://doi.org/10.1017/9781009157896.001

Dudley, R. W., Hirsch, R. M., Archfield, S. A., Blum, A. G., & Renard, B. (2019). Low streamflow trends at human-impacted and reference basins in the United States. *Journal of Hydrology*, 124254. https://doi.org/10.1016/j.jhydrol.2019.124254

Dudley, R. W., Hirsch, R. M., Archfield, S. A., Blum, A. G., & Renard, B. (2020). Low streamflow trends at human-impacted and reference basins in the United States. *Journal of Hydrology*, *580*, 124254. https://doi.org/10.1016/j.jhydrol.2019.124254

Edmund, H., & Bell, K. (2015). *prism: Access Data from the Oregon State Prism Climate Project* (0.0.6) [R]. 10.5281/zenodo.33663

ESA CCI, L. C. (2017). Product user guide version 2.0. *UCL-Geomatics: London, UK*.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A., & Dech, S. (2013). Urban Footprint Processor—Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. *IEEE Geoscience and Remote Sensing Letters*, *10*(6), 1617–1621. https://doi.org/10.1109/LGRS.2013.2272953

Espey, W. H., Morgan, C. W., & Masch, F. D. (1965). A study of some effects of urbanization on storm runoff from a small watershed. *CWE Online Reports*.

Falcone, J. A. (2011). GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. In *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow* [USGS Unnumbered Series]. U.S. Geological Survey. https://doi.org/10.3133/70046617

Falcone, J. A. (2017). *U.S. Geological Survey GAGES-II time series data from consistent sources of land use, water use, agriculture, timber activities, dam removals, and other historical anthropogenic influences* [dataset]. U.S. Geological Survey. https://doi.org/10.5066/F7HQ3XS4

Farley, K. A., Jobbágy, E. G., & Jackson, R. B. (2005). Effects of afforestation on water yield: A global synthesis with implications for policy. *Global Change Biology*, *11*(10), 1565–1576. https://doi.org/10.1111/j.1365-2486.2005.01011.x

Farmer, D., Sivapalan, M., & Jothityangkoon, C. (2003). Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes:

Downward approach to water balance analysis. *Water Resources Research*, *39*(2). https://doi.org/10.1029/2001WR000328

Farquharson, F. A. K., Meigh, J. R., & Sutcliffe, J. V. (1992). Regional flood frequency analysis in arid and semi-arid areas. *Journal of Hydrology*, *138*(3–4), 487–501. https://doi.org/10.1016/0022-1694(92)90132-F

Fenneman, N. M., & Johnson, D. W. (1946). *Physiographic divisions of the conterminous United States.* [dataset]. https://water.usgs.gov/lookup/getspatial?physio

Ferraro, P. J., Sanchirico, J. N., & Smith, M. D. (2019). Causal inference in coupled human and natural systems. *Proceedings of the National Academy of Sciences*, *116*(12), 5311–5318. https://doi.org/10.1073/pnas.1805563115

Ferreira, S., & Ghimire, R. (2012). Forest cover, socioeconomics, and reported flood frequency in developing countries. *Water Resources Research*, *48*(8). https://doi.org/10.1029/2011WR011701

Ficklin, D. L., Maxwell, J. T., Letsinger, S. L., & Gholizadeh, H. (2015). A climatic deconstruction of recent drought trends in the United States. *Environmental Research Letters*, *10*(4), 044009. https://doi.org/10.1088/1748-9326/10/4/044009

Filoso, S., Bezerra, M. O., Weiss, K. C. B., & Palmer, M. A. (2017). Impacts of forest restoration on water yield: A systematic review. *PLOS ONE*, *12*(8), e0183210. https://doi.org/10.1371/journal.pone.0183210

Fischer, E. M., & Knutti, R. (2014). Detection of spatially aggregated changes in temperature and precipitation extremes. *Geophysical Research Letters*, *41*(2), 547–554. https://doi.org/10.1002/2013GL058499

Fouad, G., Skupin, A., & Tague, C. L. (2016). Regional regression models of percentile flows for the contiguous US: Expert versus data-driven independent variable selection. *Hydrology and Earth System Sciences Discussions*, 1–33. https://doi.org/10.5194/hess-2016-639

Fowler, K., Peel, M., Saft, M., Peterson, T. J., Western, A., Band, L., Petheram, C., Dharmadi, S., Tan, K. S., Zhang, L., Lane, P., Kiem, A., Marshall, L., Griebel, A., Medlyn, B. E., Ryu, D., Bonotto, G., Wasko, C., Ukkola, A., … Nathan, R. (2022). Explaining changes in rainfall–runoff relationships during and after Australia's Millennium Drought: A community perspective. *Hydrology and Earth System Sciences*, *26*(23), 6073–6120. https://doi.org/10.5194/hess-26-6073-2022

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression* (Second). Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

François, B., Schlef, K. E., Wi, S., & Brown, C. M. (2019). Design considerations for riverine floods in a changing climate – A review. *Journal of Hydrology*, *574*, 557–573. https://doi.org/10.1016/j.jhydrol.2019.04.068

Frans, C., Istanbulluoglu, E., Mishra, V., Munoz-Arriola, F., & Lettenmaier, D. P. (2013). Are climatic or land cover changes the dominant cause of runoff trends in the Upper Mississippi River Basin? *Geophysical Research Letters*, *40*(6), 1104–1110. https://doi.org/10.1002/grl.50262

Fu, B. P. (1981). On the calculation of the evaporation from land surface. *Scientia Atmospherica Sinica*, *5*(1), 23.

Fu, G., Charles, S. P., & Chiew, F. H. S. (2007). A two-parameter climate elasticity of streamflow index to assess climate change effects on annual streamflow. *Water Resources Research*, *43*(11). https://doi.org/10.1029/2007WR005890

Gauster, T., Laaha, G., & Koffler, D. (2022). *lfstat: Calculation of Low Flow Statistics for Daily Stream Flow Data* (0.9.12) [Computer software]. https://cran.r-project.org/web/packages/lfstat/index.html

Girotto, M., & Rodell, M. (2019). Chapter Two—Terrestrial water storage. In V. Maggioni & C. Massari (Eds.), *Extreme Hydroclimatic Events and Multivariate Hazards in a Changing Environment* (pp. 41–64). Elsevier. https://doi.org/10.1016/B978-0-12-814899-0.00002-X

Gnann, S. J., McMillan, H. K., Woods, R. A., & Howden, N. J. K. (2021). Including Regional Knowledge Improves Baseflow Signature Predictions in Large Sample Hydrology. *Water Resources Research*, *57*(2), e2020WR028354. https://doi.org/10.1029/2020WR028354

Gnann, S. J., Woods, R. A., & Howden, N. J. K. (2019). Is There a Baseflow Budyko Curve? *Water Resources Research*, *55*(4), 2838–2855. https://doi.org/10.1029/2018WR024464

Goeking, S. A., & Tarboton, D. G. (2020). Forests and Water Yield: A Synthesis of Disturbance Effects on Streamflow and Snowpack in Western Coniferous Forests. *Journal of Forestry*, *118*(2), 172–192. https://doi.org/10.1093/jofore/fvz069

Grafton, R. Q., Williams, J., Perry, C. J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., & Allen, R. G. (2018). The paradox of irrigation efficiency. *Science*, *361*(6404), 748–750. https://doi.org/10.1126/science.aat9314

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Guardiola-Claramonte, M., Troch, P. A., Breshears, D. D., Huxman, T. E., Switanek, M. B., Durcik, M., & Cobb, N. S. (2011). Decreased streamflow in semi-arid basins following drought-induced tree die-off: A counter-intuitive and indirect climate impact on hydrology. *Journal of Hydrology*, *406*(3), 225–233. https://doi.org/10.1016/j.jhydrol.2011.06.017

Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2016). The sensitivity of water availability to changes in the aridity index and other factors—A probabilistic analysis in the Budyko space. *Geophysical Research Letters*, *43*(13), 6985–6994. https://doi.org/10.1002/2016GL069763

Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Correspondence: Flawed assumptions compromise water yield assessment. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms14795

Guo, J., Li, H.-Y., Leung, L. R., Guo, S., Liu, P., & Sivapalan, M. (2014). Links between flood frequency and annual water balance behaviors: A basis for similarity and regionalization. *Water Resources Research*, *50*(2), 937–953. https://doi.org/10.1002/2013WR014374

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: A need to balance depth with breadth. *Hydrology and Earth System Sciences*, *18*(2), 463–477. https://doi.org/10.5194/hess-18-463-2014

Gustard, A., Bullock, A., & Dixon, J. M. (1992). *Low flow estimation in the United Kingdom.* Institute of Hydrology.

Hall, F. R. (1968). Base-Flow Recessions—A Review. *Water Resources Research*, *4*(5), 973–983. https://doi.org/10.1029/WR004i005p00973

Hamed, K. H. (2008). Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, *349*(3), 350–363. https://doi.org/10.1016/j.jhydrol.2007.11.009

Hamon, W. R. (1963a). Computation of direct runoff amounts from storm rainfall. *International Association of Scientific Hydrology Publication*, *63*, 52–62.

Hamon, W. R. (1963b). Estimating Potential Evapotranspiration. *Transactions of the American Society of Civil Engineers*, *128*(1), 324–338. https://doi.org/10.1061/TACEAT.0008673

Han, S., Slater, L., Wilby, R. L., & Faulkner, D. (2022). Contribution of urbanisation to non-stationary river flow in the UK. *Journal of Hydrology*, *613*, 128417. https://doi.org/10.1016/j.jhydrol.2022.128417

Harman, C. J., Troch, P. A., & Sivapalan, M. (2011). Functional model of water balance variability at the catchment scale: 2. Elasticity of fast and slow runoff components to precipitation change in the continental United States. *Water Resources Research*, *47*(2), 1–12. https://doi.org/10.1029/2010WR009656

Hart, E., & Bell, K. (2015a). *prism: Download data from the Oregon prism project* [R package version 0.0.6]. https://github.com/ropensci/prism

Hart, E., & Bell, K. (2015b). *Prism: Access Data From The Oregon State Prism Climate Project* [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.33663

Hejazi, M. I., & Markus, M. (2009). Impacts of Urbanization and Climate Variability on Floods in Northeastern Illinois. *Journal of Hydrologic Engineering*, *14*(6), 606–616. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000020

Hellwig, J., & Stahl, K. (2018). An assessment of trends and potential future changes in groundwater-baseflow drought based on catchment response times. *Hydrology and Earth System Sciences*, *22*(12), 6209–6224. https://doi.org/10.5194/hess-22-6209-2018

Hibbert, A. R. (1965). *Forest treatment effects on water yield*. Citeseer.

Hidalgo, H. G., Das, T., Dettinger, M. D., Cayan, D. R., Pierce, D. W., Barnett, T. P., Bala, G., Mirin, A., Wood, A. W., Bonfils, C., Santer, B. D., & Nozawa, T. (2009). Detection and

Attribution of Streamflow Timing Changes to Climate Change in the Western United States. *Journal of Climate*, *22*(13), 3838–3855. https://doi.org/10.1175/2009JCLI2470.1

Hodgkins, G. A., Dudley, R. W., Archfield, S. A., & Renard, B. (2019). Effects of climate, regulation, and urbanization on historical flood trends in the United States. *Journal of Hydrology*, *573*, 697–709. https://doi.org/10.1016/j.jhydrol.2019.03.102

Hollis, G. E. (1975). The effect of urbanization on floods of different recurrence interval. *Water Resources Research*, *11*(3), 431–435. https://doi.org/10.1029/WR011i003p00431

Hopkins, K. G., Morse, N. B., Bain, D. J., Bettez, N. D., Grimm, N. B., Morse, J. L., Palta, M. M., Shuster, W. D., Bratt, A. R., & Suchy, A. K. (2015, February 9). *Assessment of Regional Variation in Streamflow Responses to Urbanization and the Persistence of Physiography* (world) [Research-article]. ACS Publications; American Chemical Society. https://doi.org/10.1021/es505389y

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., … Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB)—A review. *Hydrological Sciences Journal*, *58*(6), 1198–1255. https://doi.org/10.1080/02626667.2013.803183

Hsiao, C. (2007). Panel Data Analysis—Advantages and Challenges. *TEST*, *16*, 1–22. https://doi.org/10.1007/s11749-007-0046-x

Hughes, J. D., Petrone, K. C., & Silberstein, R. P. (2012). Drought, groundwater storage and stream flow decline in southwestern Australia. *Geophysical Research Letters*, *39*(3). https://doi.org/10.1029/2011GL050797

Huntington-Klein, N. (2021). *The effect: An introduction to research design and causality*. CRC Press.

Ivancic, T. J., & Shaw, S. B. (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. https://doi.org/10.1007/s10584-015-1476-1

Jacobson, C. R. (2011). Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review. *Journal of Environmental Management*, *92*(6), 1438–1448. https://doi.org/10.1016/j.jenvman.2011.01.018

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 103). Springer New York. https://doi.org/10.1007/978-1-4614-7138-7

Jasechko, S., Kirchner, J. W., Welker, J. M., & McDonnell, J. J. (2016). Substantial proportion of global streamflow less than three months old. *Nature Geoscience*, *9*(2), Article 2. https://doi.org/10.1038/ngeo2636

Jassby, A., Cloern, J., & Stachelek, J. (2022). *wql: Exploring Water Quality Monitoring Data* (1.0.0) [Computer software]. https://cran.r-project.org/package=wql

Jenicek, M., Seibert, J., Zappa, M., Staudinger, M., & Jonas, T. (2016). Importance of maximum snow accumulation for summer low flows in humid catchments. *Hydrology and Earth System Sciences*, *20*(2), 859–874. https://doi.org/10.5194/hess-20-859-2016

José Pinheiro, Bates, D., & R Core Team. (2023). *nlme: Linear and Nonlinear Mixed Effects Models* (3.1-163) [Computer software]. https://cran.r-project.org/web/packages/nlme/index.html

Kalbus, E., Reinstorf, F., & Schirmer, M. (2006). Measuring methods for groundwater – surface water interactions: A review. *Hydrology and Earth System Sciences*, *10*(6), 873–887. https://doi.org/10.5194/hess-10-873-2006

Kelleher, C., Wagener, T., & McGlynn, B. (2015). Model-based analysis of the influence of catchment properties on hydrologic partitioning across five mountain headwater subcatchments. *Water Resources Research*, *51*(6), 4109–4136. https://doi.org/10.1002/2014WR016147

Kendall, M. G. (1948). *Rank correlation methods*. Griffin.

Kendy, E., & Bredehoeft, J. D. (2006). Transient effects of groundwater pumping and surface-water-irrigation returns on streamflow. *Water Resources Research*, *42*(8). https://doi.org/10.1029/2005WR004792

Kirchner, J. W. (2003). A double paradox in catchment hydrology and geochemistry. *Hydrological Processes*, *17*(4), 871–874. https://doi.org/10.1002/hyp.5108

Klaus, J., & McDonnell, J. J. (2013). Hydrograph separation using stable isotopes: Review and evaluation. *Journal of Hydrology*, *505*, 47–64. https://doi.org/10.1016/j.jhydrol.2013.09.006

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data. Water Resources Research, 54(7), 5088–5109. https://doi.org/10.1029/2018WR022913

Koehn, C. R., Petrie, M. D., Bradford, J. B., Litvak, M. E., & Strachan, S. (2021). Seasonal Precipitation and Soil Moisture Relationships Across Forests and Woodlands in the Southwestern United States. *Journal of Geophysical Research: Biogeosciences*, *126*(4), e2020JG005986. https://doi.org/10.1029/2020JG005986

Konapala, G., & Mishra, A. K. (2016). Three-parameter-based streamflow elasticity model: Application to MOPEX basins in the USA at annual and seasonal scales. *Hydrology and Earth System Sciences*, *20*(6), 2545–2556. https://doi.org/10.5194/hess-20-2545-2016

Kormos, P. R., Luce, C. H., Wenger, S. J., & Berghuijs, W. R. (2016). Trends and sensitivities of low streamflow extremes to discharge timing and magnitude in Pacific Northwest mountain streams. *Water Resources Research*, *52*(7), 4990–5007. https://doi.org/10.1002/2015WR018125

Koutsoyiannis, D. (2006). Nonstationarity versus scaling in hydrology. *Journal of Hydrology*, *324*(1–4), 239–254.

Krakauer, N. Y., & Temimi, M. (2011). Stream recession curves and storage variability in small watersheds. *Hydrology and Earth System Sciences*, *15*(7), 2377–2389. https://doi.org/10.5194/hess-15-2377-2011

Kustu, M. D., Fan, Y., & Rodell, M. (2011). Possible link between irrigation in the U.S. High Plains and increased summer streamflow in the Midwest. *Water Resources Research*, *47*(3). https://doi.org/10.1029/2010WR010046

Ladson, A. R., Brown, R., Neal, B., & Nathan, R. (2013). A Standard Approach to Baseflow Separation Using The Lyne and Hollick Filter. *Australasian Journal of Water Resources*, *17*(1), 25–34. https://doi.org/10.7158/13241583.2013.11465417

Levy, M. C., Lopes, A. V., Cohn, A., Larsen, L. G., & Thompson, S. E. (2018). Land Use Change Increases Streamflow Across the Arc of Deforestation in Brazil. *Geophysical Research Letters*, *45*(8), 3520–3530. https://doi.org/10.1002/2017GL076526

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the

future? *Geophysical Research Letters*, *44*(12), 6163–6172.

https://doi.org/10.1002/2017GL073551

Lins, H. F., & Cohn, T. A. (2002). *40. Floods in the greenhouse:spinning the right tale*. 7.

Lins, H. F., & Slack, J. R. (1999). Streamflow trends in the United States. *Geophysical*

Research Letters, *26*(2), 227–230. https://doi.org/10.1029/1998GL900291

Lins, H. F., & Slack, J. R. (2005). Seasonal and Regional Characteristics of U.S. Streamflow

Trends in the United States from 1940 to 1999. *Physical Geography*, *26*(6), 489–501.

https://doi.org/10.2747/0272-3646.26.6.489

Litvak, E., Manago, K. F., Hogue, T. S., & Pataki, D. E. (2017). Evapotranspiration of urban

landscapes in Los Angeles, California at the municipal scale. *Water Resources Research*,

53(5), 4236–4252. https://doi.org/10.1002/2016WR020254

Lombard, P. J., & Holtschlag, D. J. (2018). Estimating Lag to Peak between Rainfall and

Peak Streamflow with a Mixed-Effects Model. *JAWRA Journal of the American Water*

Resources Association, *54*(4), 949–961. https://doi.org/10.1111/1752-1688.12653

Loon, A. F. V., Rangecroft, S., Coxon, G., Breña Naranjo, J. A., Ogtrop, F. V., & Lanen, H.

A. J. V. (2019). Using paired catchments to quantify the human influence on hydrological

droughts. *Hydrology and Earth System Sciences*, *23*(3), 1725–1739.

https://doi.org/10.5194/hess-23-1725-2019

López, J., & Francés, F. (2013). Non-stationary flood frequency analysis in continental

Spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and*

Earth System Sciences, *17*(8), 3189–3203. https://doi.org/10.5194/hess-17-3189-2013

Lu, J., Sun, G., McNulty, S. G., & Amatya, D. M. (2007a). A Comparison of Six Potential

Evapotranspiration Methods for Regional Use in the Southeastern United States1. *JAWRA*

Journal of the American Water Resources Association, *41*(3), 621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x

Lu, J., Sun, G., McNulty, S. G., & Amatya, D. M. (2007b). A Comparison of Six Potential Evapotranspiration Methods for Regional Use in the Southeastern United States1. *JAWRA Journal of the American Water Resources Association*, *41*(3), 621–633. https://doi.org/10.1111/j.1752-1688.2005.tb03759.x

Lundquist, J. D., & Cayan, D. R. (2002). Seasonal and Spatial Patterns in Diurnal Cycles in Streamflow in the Western United States. *Journal of Hydrometeorology*, *3*(5), 591–603. https://doi.org/10.1175/1525-7541(2002)003<0591:SASPID>2.0.CO;2

Luthy, R. G., Sedlak, D. L., Plumlee, M. H., Austin, D., & Resh, V. H. (2015). Wastewater-effluent-dominated streams as ecosystem-management tools in a drier climate. *Frontiers in Ecology and the Environment*, *13*(9), 477–485. https://doi.org/10.1890/150038

Lyne, V., & Hollick, M. (1979). Stochastic time-variable rainfall-runoff modelling. *Institute of Engineers Australia National Conference*, *79*(10), 89–93.

Maidment, D. R. (1992). Handbook of Hydrology, McGraw-Hil. *Inc., New York, NY*.

Mallakpour, I., & Villarini, G. (2016). A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. *Hydrological Sciences Journal*, *61*(2), 245–254. https://doi.org/10.1080/02626667.2015.1008482

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245–259. https://doi.org/10.2307/1907187

Matalas, N. C. (2012). Comment on the Announced Death of Stationarity. *Journal of Water Resources Planning and Management*, *138*(4), 311–312. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000215

Matheussen, B., Kirschbaum, R. L., Goodman, I. A., O'Donnell, G. M., & Lettenmaier, D. P. (2000). Effects of land cover change on streamflow in the interior Columbia River Basin (USA and Canada). *Hydrological Processes*, *14*(5), 867–885. https://doi.org/10.1002/(SICI)1099-1085(20000415)14:5<867::AID-HYP975>3.0.CO;2-5

McCabe, G. J., Wolock, D. M., Pederson, G. T., Woodhouse, C. A., & McAfee, S. (2017). Evidence that Recent Warming is Reducing Upper Colorado River Flows. *Earth Interactions*, *21*(10), 1–14. https://doi.org/10.1175/EI-D-17-0007.1

McMillan, H. (2020). Linking hydrologic signatures to hydrologic processes: A review. *Hydrological Processes*, *34*(6), 1393–1409. https://doi.org/10.1002/hyp.13632

McMillan, H. (2021). A review of hydrologic signatures and their applications. *WIREs Water*, *8*(1), e1499. https://doi.org/10.1002/wat2.1499

McNamara, J. P., Tetzlaff, D., Bishop, K., Soulsby, C., Seyfried, M., Peters, N. E., Aulenbach, B. T., & Hooper, R. (2011). Storage as a Metric of Catchment Comparison. *Hydrological Processes*, *25*(21), 3364–3371. https://doi.org/10.1002/hyp.8113

McPhillips, L. E., Earl, S. R., Hale, R. L., & Grimm, N. B. (2019). Urbanization in Arid Central Arizona Watersheds Results in Decreased Stream Flashiness. *Water Resources Research*, *55*(11), 9436–9453. https://doi.org/10.1029/2019WR025835

Meyer, S. C. (2002). Investigation of Impacts of Urbanization on Base Flow and Recharge Rates, Northeastern Illinois: Summary of Year 2 Activities. *Proceedings, 12th Annual Illinois Groundwater Consortium Symposium*. https://hdl.handle.net/2142/55237

Meyer, S. C. (2005). Analysis of base flow trends in urban streams, northeastern Illinois, USA. *Hydrogeology Journal*, *13*(5), 871–885. https://doi.org/10.1007/s10040-004-0383-8

Mezentsev, V. (1955). Back to the computation of total evaporation. *Meteorologia i Gidrologia*, *5*, 24–26.

Miller, J. D., Kim, H., Kjeldsen, T. R., Packman, J., Grebby, S., & Dearden, R. (2014). Assessing the impact of urbanization on storm runoff in a peri-urban catchment using historical change in impervious cover. *Journal of Hydrology*, *515*, 59–70. https://doi.org/10.1016/j.jhydrol.2014.04.011

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? *Science*, *319*(5863), 573–574. https://doi.org/10.1126/science.1151915

Milly, P. C. D., & Dunne, K. A. (2002). Macroscale water fluxes 2. Water and energy supply control of their interannual variability. *Water Resources Research*, *38*(10), 24-1-24–29. https://doi.org/10.1029/2001WR000760

Milly, P. C. D., & Dunne, K. A. (2020). Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation. *Science*, *367*(6483), 1252–1255. https://doi.org/10.1126/science.aay9187

Milly, P. C. D., Kam, J., & Dunne, K. A. (2018a). *Annual Streamflow Sensitivity to Air Temperature Worldwide, 1901-2013* [dataset]. U.S. Geological Survey. https://doi.org/10.5066/F7SN085V

Milly, P. C. D., Kam, J., & Dunne, K. A. (2018b). On the Sensitivity of Annual Streamflow to Air Temperature. *Water Resources Research*, *54*(4), 2624–2641. https://doi.org/10.1002/2017WR021970

Moore, D. T., & Wondzell, S. M. (2005). Physical Hydrology and the Effects of Forest Harvesting in the Pacific Northwest: A Review1. *JAWRA Journal of the American Water Resources Association*, *41*(4), 763–784. https://doi.org/10.1111/j.1752-1688.2005.tb03770.x

Müller, M. F., & Levy, M. C. (2019). Complementary Vantage Points: Integrating Hydrology and Economics for Sociohydrologic Knowledge Generation. *Water Resources Research*, *55*(4), 2549–2571. https://doi.org/10.1029/2019WR024786

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: An overview. *WIREs Data Mining and Knowledge Discovery*, *2*(1), 86–97. https://doi.org/10.1002/widm.53

Neal, C., & Rosier, P. T. W. (1990). Chemical studies of chloride and stable oxygen isotopes in two conifer afforested and moorland sites in the British uplands. *Journal of Hydrology*, *115*(1), 269–283. https://doi.org/10.1016/0022-1694(90)90209-G

Neri, A., Villarini, G., Slater, L. J., & Napolitano, F. (2019). On the statistical attribution of the frequency of flood events across the U.S. Midwest. *Advances in Water Resources*, *127*, 225–236. https://doi.org/10.1016/j.advwatres.2019.03.019

Nichols, A. (2007). Causal Inference with Observational Data. *The Stata Journal*, *7*(4), 507–541. https://doi.org/10.1177/1536867X0800700403

Nijssen, B., O'Donnell, G. M., Hamlet, A. F., & Lettenmaier, D. P. (2001). Hydrologic Sensitivity of Global Rivers to Climate Change. *Climatic Change*, *50*(1), 143–175. https://doi.org/10.1023/A:1010616428763

O'Driscoll, M., Clinton, S., Jefferson, A., Manda, A., & McMillan, S. (2010). Urbanization Effects on Watershed Hydrology and In-Stream Processes in the Southern United States. *Water*, *2*(3), 605–648. https://doi.org/10.3390/w2030605

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. (2020). Evaluation of Methods for Causal Discovery in Hydrometeorological Systems. *Water Resources Research*, *56*(7). https://doi.org/10.1029/2020WR027251

Otsuka, J. (2023). *Thinking About Statistics: The Philosophical Foundations*. Routledge. https://doi.org/10.4324/9781003319061

Oudin, L., Salavati, B., Furusho-Percot, C., Ribstein, P., & Saadi, M. (2018a). Hydrological impacts of urbanization at the catchment scale. *Journal of Hydrology*, *559*, 774–786. https://doi.org/10.1016/j.jhydrol.2018.02.064

Oudin, L., Salavati, B., Furusho-Percot, C., Ribstein, P., & Saadi, M. (2018b). Hydrological impacts of urbanization at the catchment scale. *Journal of Hydrology*, *559*, 774–786. https://doi.org/10.1016/j.jhydrol.2018.02.064

Page, T., Chappell, N. A., Beven, K. J., Hankin, B., & Kretzschmar, A. (2020). Assessing the significance of wet-canopy evaporation from forests during extreme rainfall events for flood mitigation in mountainous regions of the United Kingdom. *Hydrological Processes*, *34*(24), 4740–4754. https://doi.org/10.1002/hyp.13895

Parry, S., Wilby, R. L., Prudhomme, C., & Wood, P. J. (2016). A systematic assessment of drought termination in the United Kingdom. *Hydrology and Earth System Sciences*, *20*(10), 4265–4281. https://doi.org/10.5194/hess-20-4265-2016

Partington, D., Brunner, P., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., & Dandy, G. C. (2012). Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model. *Journal of Hydrology*, *458–459*, 28–39. https://doi.org/10.1016/j.jhydrol.2012.06.029

Patil, S., & Stieglitz, M. (2011). Hydrologic similarity among catchments under variable flow conditions. *Hydrology and Earth System Sciences*, *15*(3), 989–997. https://doi.org/10.5194/hess-15-989-2011

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*(4), 669–688. https://doi.org/10.1093/biomet/82.4.669

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*(0), 96–146. https://doi.org/10.1214/09-SS057

Pei, L., Moore, N., Zhong, S., Kendall, A. D., Gao, Z., & Hyndman, D. W. (2016). Effects of Irrigation on Summer Precipitation over the United States. *Journal of Climate*, *29*(10), 3541–3558. https://doi.org/10.1175/JCLI-D-15-0337.1

Perdigão, R. A. P., & Blöschl, G. (2014). Spatiotemporal flood sensitivity to annual precipitation: Evidence for landscape-climate coevolution. *Water Resources Research*, *50*(7), 5492–5509. https://doi.org/10.1002/2014WR015365

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Carneiro Freire Sergio, M., Halkia, S., Julea Andreea, M., Kemper, T., Soille, P., & Syrris, V. (2016). *Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014* (JRC97705) [EUR - Scientific and Technical Research Reports]. Publications Office of the European Union. https://publications.jrc.ec.europa.eu/repository/handle/111111111/40182

Peters-Lidard, C. D., Clark, M., Samaniego, L., Verhoest, N. E. C., Emmerik, T. van, Uijlenhoet, R., Achieng, K., Franz, T. E., & Woods, R. (2017). Scaling, similarity, and the fourth paradigm for hydrology. *Hydrology and Earth System Sciences*, *21*(7), 3701–3713. https://doi.org/10.5194/hess-21-3701-2017

Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(2), 126–135. JSTOR. https://doi.org/10.2307/2346729

Pfister, L., & Kirchner, J. W. (2017). Debates—Hypothesis testing in hydrology: Theory and practice. *Water Resources Research*, *53*(3), 1792–1798. https://doi.org/10.1002/2016WR020116

Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., & McDonnell, J. J. (2017). Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. *Hydrological Processes*, *31*(10), 1828–1845. https://doi.org/10.1002/hyp.11134

Pike, J. G. (1964). The estimation of annual run-off from meteorological data in a tropical climate. *Journal of Hydrology*, *2*(2), 116–123. https://doi.org/10.1016/0022-1694(64)90022-8

Poff, N. L., Bledsoe, B. P., & Cuhaciyan, C. O. (2006). Hydrologic variation with land use across the contiguous United States: Geomorphic and ecological consequences for stream ecosystems. *Geomorphology*, *79*(3), 264–285. https://doi.org/10.1016/j.geomorph.2006.06.032

Poff, N. L., & Zimmerman, J. K. H. (2010). Ecological responses to altered flow regimes: A literature review to inform the science and management of environmental flows. *Freshwater Biology*, *55*(1), 194–205. https://doi.org/10.1111/j.1365-2427.2009.02272.x

Ponce, V. M., & Shetty, A. V. (1995). A conceptual model of catchment water balance: 1. Formulation and calibration. *Journal of Hydrology*, *173*(1–4), 27–40.

Potter, N. J., Petheram, C., & Zhang, L. (2011). Sensitivity of streamflow to rainfall and temperature in south-eastern Australia during the Millennium drought. 19th International Congress on Modelling and Simulation, Perth, Dec, 3636–3642.

Price, K. (2011). Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review: *Progress in Physical Geography*, *35*(4), 465–492. https://doi.org/10.1177/0309133311402714

Prosdocimi, I., Kjeldsen, T. R., & Miller, J. D. (2015). Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. *Water Resources Research*, *51*(6), 4244–4262. https://doi.org/10.1002/2015WR017065

Prudhomme, C., Crooks, S., Kay, A. L., & Reynard, N. (2013). Climate change and river flooding: Part 1 classifying the sensitivity of British catchments. *Climatic Change*, *119*(3), 933–948.

Pyrce, R. (2004). *Hydrological Low Flow Indices and their Uses WSC. Report No.04-2004*. 33. http://files.faksnes.webnode.se/200004594-b3995b5902/LowFlowOntRpt2004.pdf

Rice, J. S., Emanuel, R. E., Vose, J. M., & Nelson, S. A. C. (2015). Continental U.S. streamflow trends from 1940 to 2009 and their relationships with watershed spatial characteristics. *Water Resources Research*, *51*(8), 6262–6275. https://doi.org/10.1002/2014WR016367

Richter, B. D., Baumgartner, J. V., Powell, J., & Braun, D. P. (1996). A Method for Assessing Hydrologic Alteration within Ecosystems. *Conservation Biology*, *10*(4), 1163–1174. https://doi.org/10.1046/j.1523-1739.1996.10041163.x

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized Additive Models for Location, Scale and Shape. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *54*(3), 507–554. https://doi.org/10.1111/j.1467-9876.2005.00510.x

Rigby, R. A., & Stasinopoulos, D. M. (2017). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 507–554. https://doi.org/10.1111/j.1467-9876.2005.00510.x@10.1111/(ISSN)1467-9876.TOP_SERIES_C_RESEARCH

Robinson, M., Rodda, J. C., & Sutcliffe, J. V. (2013). Long-term environmental monitoring in the UK: Origins and achievements of the Plynlimon catchment study. *Transactions of the Institute of British Geographers*, *38*(3), 451–463. https://doi.org/10.1111/j.1475-5661.2012.00534.x

Rodell, M., & Famiglietti, J. S. (2002). The potential for satellite-based monitoring of groundwater storage changes using GRACE: The High Plains aquifer, Central US. *Journal of Hydrology*, *263*(1), 245–256. https://doi.org/10.1016/S0022-1694(02)00060-4

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), Article 7707. https://doi.org/10.1038/s41586-018-0123-1

Rogger, M., Agnoletti, M., Alaoui, A., Bathurst, J. C., Bodner, G., Borga, M., Chaplot, V., Gallart, F., Glatzel, G., Hall, J., Holden, J., Holko, L., Horn, R., Kiss, A., Kohnová, S., Leitinger, G., Lennartz, B., Parajka, J., Perdigão, R., … Blöschl, G. (2017). Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research. *Water Resources Research*, *53*(7), 5209–5219. https://doi.org/10.1002/2017WR020723

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R.,

Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., …

Zscheischler, J. (2019). Inferring causation from time series in Earth system sciences. *Nature*

Communications, *10*(1), 2553. https://doi.org/10.1038/s41467-019-10105-3

Sadri, S., Kam, J., & Sheffield, J. (2016). Nonstationarity of low flows and their timing in the

eastern United States. *Hydrology and Earth System Sciences*, *20*(2), 633–649.

https://doi.org/10.5194/hess-20-633-2016

Safeeq, M., Grant, G. E., Lewis, S. L., & Tague, Christina. L. (2013). Coupling snowpack

and groundwater dynamics to interpret historical streamflow trends in the western United

States. *Hydrological Processes*, *27*(5), 655–668. https://doi.org/10.1002/hyp.9628

Safeeq, M., & Hunsaker, C. T. (2016). Characterizing Runoff and Water Yield for Headwater

Catchments in the Southern Sierra Nevada. *JAWRA Journal of the American Water*

Resources Association, *52*(6), 1327–1346. https://doi.org/10.1111/1752-1688.12457

Saft, M., Peel, M. C., Western, A. W., & Zhang, L. (2016). Predicting shifts in rainfall-runoff

partitioning during multiyear drought: Roles of dry period and catchment characteristics.

Water Resources Research, *52*(12), 9290–9305. https://doi.org/10.1002/2016WR019525

Saft, M., Western, A. W., Zhang, L., Peel, M. C., & Potter, N. J. (2015). The influence of

multiyear drought on the annual rainfall-runoff relationship: An Australian perspective.

Water Resources Research, *51*(4), 2444–2463. https://doi.org/10.1002/2014WR015348

Salavati, B., Oudin, L., Furusho-Percot, C., & Ribstein, P. (2016). Modeling approaches to

detect land-use changes: Urbanization analyzed on a set of 43 US catchments. *Journal of*

Hydrology, *538*, 138–151. https://doi.org/10.1016/j.jhydrol.2016.04.010

Sankarasubramanian, A., Vogel, R. M., & Limbrunner, J. F. (2001). Climate elasticity of streamflow in the United States. *Water Resources Research*, *37*(6), 1771–1781. https://doi.org/10.1029/2000WR900330

Schaake, J. C. (1990). From climate to flow. In P. E. Waggoner, *Climate change and US water resources* (Vol. 8, pp. 177–206). https://api.semanticscholar.org/CorpusID:140694770

Schilling, K. E., Gassman, P. W., Kling, C. L., Campbell, T., Jha, M. K., Wolter, C. F., & Arnold, J. G. (2014). The potential for agricultural land use change to reduce flood risk in a large watershed. *Hydrological Processes*, *28*(8), 3314–3325. https://doi.org/10.1002/hyp.9865

Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters*, *85*(2), 461–464. https://doi.org/10.1103/PhysRevLett.85.461

Searcy, J. K. (1959). Flow-duration curves. In *Water Supply Paper* (1542-A). U.S. Govt. Print. Off.,. https://doi.org/10.3133/wsp1542A

Seibert, J., & McDonnell, J. J. (2010). Land-cover impacts on streamflow: A change-detection modelling approach that incorporates parameter uncertainty. *Hydrological Sciences Journal*, *55*(3), 316–332. https://doi.org/10.1080/02626661003683264

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., & Zhou, B. (2021). Weather and climate extreme events in a changing climate. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, Ö. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of*

the *Intergovernmental Panel on Climate Change* (pp. 1513–1766). Cambridge University Press. https://doi.org/10.1017/9781009157896.001

Serinaldi, F., & Kilsby, C. G. (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Advances in Water Resources*, *77*, 17–36.

Shadish, W., Cook, Thomas D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Boston, MA.

Shao, S., Zhang, H., Singh, V. P., Ding, H., Zhang, J., & Wu, Y. (2022). Nonstationary analysis of hydrological drought index in a coupled human-water system: Application of the GAMLSS with meteorological and anthropogenic covariates in the Wuding River basin, China. *Journal of Hydrology*, *608*, 127692. https://doi.org/10.1016/j.jhydrol.2022.127692

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If Precipitation Extremes Are Increasing, Why Aren't Floods? *Water Resources Research*, *54*(11), 8545–8551. https://doi.org/10.1029/2018WR023749

Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., & Smith, D. R. (2005). Impacts of impervious surface on watershed hydrology: A review. *Urban Water Journal*, *2*(4), 263–275. https://doi.org/10.1080/15730620500386529

Sims, N., Green, C., Newnham, G., England, J., Held, A., Wulder, M., Herold, M., Cox, S., Huete, A., Kumar, L., & others. (2017). Good practice guidance. SDG indicator 15.3. 1: Proportion of land that is degraded over total land area. *United Nations Convention to Combat Desertification (UNCCD), Bonn, Germany*.

Singh, R., Wagener, T., van Werkhoven, K., Mann, M. E., & Crane, R. (2011). A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing

climate – accounting for changing watershed behavior. *Hydrology and Earth System Sciences*, *15*(11), 3591–3603. https://doi.org/10.5194/hess-15-3591-2011

Sivapalan, M., Yaeger, M. A., Harman, C. J., Xu, X., & Troch, P. A. (2011). Functional model of water balance variability at the catchment scale: 1. Evidence of hydrologic similarity and space-time symmetry. *Water Resources Research*, *47*(2). https://doi.org/10.1029/2010WR009568

Sklash, M. G., & Farvolden, R. N. (1979). The role of groundwater in storm runoff. *Journal of Hydrology*, *43*(1), 45–65. https://doi.org/10.1016/0022-1694(79)90164-1

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C., & Wilby, R. L. (2020). Nonstationary weather and water extremes: A review of methods for their detection, attribution, and management. *Hydrology and Earth System Sciences Discussions*, 1–54. https://doi.org/10.5194/hess-2020-576

Slater, L. J., & Villarini, G. (2016a). Recent trends in U.S. flood risk. *Geophysical Research Letters*, *43*(24), 12,428-12,436. https://doi.org/10.1002/2016GL071199

Slater, L. J., & Villarini, G. (2016b). Recent trends in U.S. flood risk. *Geophysical Research Letters*, *43*(24), 12,428-12,436. https://doi.org/10.1002/2016GL071199

Slater, L. J., & Villarini, G. (2016c). Recent trends in U.S. flood risk. *Geophysical Research Letters*, *43*(24), 12,428-12,436. https://doi.org/10.1002/2016GL071199

Slater, L. J., & Villarini, G. (2017). Evaluating the Drivers of Seasonal Streamflow in the U.S. Midwest. *Water*, *9*(9), 695. https://doi.org/10.3390/w9090695

Slater, L. J., Villarini, G., Archfield, S., Faulkner, D., Lamb, R., Khouakhi, A., & Yin, J. (2021). Global Changes in 20-Year, 50-Year, and 100-Year River Floods. *Geophysical Research Letters*, *48*(6), e2020GL091824. https://doi.org/10.1029/2020GL091824

Slinski, K. M., Hogue, T. S., Porter, A. T., & McCray, J. E. (2016). Recent bark beetle outbreaks have little impact on streamflow in the Western United States. *Environmental Research Letters*, *11*(7), 074010. https://doi.org/10.1088/1748-9326/11/7/074010

Smakhtin, V. U. (2001). Low flow hydrology: A review. *Journal of Hydrology*, *240*(3), 147–186. https://doi.org/10.1016/S0022-1694(00)00340-1

Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *Journal of Statistical Software*, *23*(7). https://doi.org/10.18637/jss.v023.i07

Stasinopoulos, D., Rigby, R., Heller, G., Voudouris, V., & De Bastiani, F. (2017). Flexible regression and smoothing: Using GAMLSS in R. In *Flexible Regression and Smoothing: Using GAMLSS in R*. https://doi.org/10.1201/b21973

Staudinger, M., Stoelzle, M., Seeger, S., Seibert, J., Weiler, M., & Stahl, K. (2017). Catchment water storage variation with elevation. *Hydrological Processes*, *31*(11), 2000–2015. https://doi.org/10.1002/hyp.11158

Steinschneider, S., Yang, Y.-C. E., & Brown, C. (2013). Panel regression techniques for identifying impacts of anthropogenic landscape change on hydrologic response. *Water Resources Research*, *49*(12), 7874–7886. https://doi.org/10.1002/2013WR013818

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward Earlier Streamflow Timing across Western North America. *Journal of Climate*, *18*(8), 1136–1155. https://doi.org/10.1175/JCLI3321.1

Stoelzle, M., Schuetz, T., Weiler, M., Stahl, K., & Tallaksen, L. M. (2020). Beyond binary baseflow separation: A delayed-flow index for multiple streamflow contributions. *Hydrology and Earth System Sciences*, *24*(2), 849–867. https://doi.org/10.5194/hess-24-849-2020

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting Causality in Complex Ecosystems. *Science*, *338*(6106), 496–500. https://doi.org/10.1126/science.1227079

Sun, X., Li, Z., & Tian, Q. (2020). Assessment of hydrological drought based on nonstationary runoff data. *Hydrology Research*, *51*(5), 894–910. https://doi.org/10.2166/nh.2020.029

Swank, W. T., Vose, J. M., & Elliott, K. J. (2001). Long-term hydrologic and water quality responses following commercial clearcutting of mixed hardwoods on a southern Appalachian catchment. *Forest Ecology and Management*, *143*(1), 163–178. https://doi.org/10.1016/S0378-1127(00)00515-6

Tague, C., & Grant, G. E. (2009). Groundwater dynamics mediate low-flow response to global warming in snow-dominated alpine regions. *Water Resources Research*, *45*(7). https://doi.org/10.1029/2008WR007179

Tague, C., Grant, G., Farrell, M., Choate, J., & Jefferson, A. (2008). Deep groundwater mediates streamflow response to climate warming in the Oregon Cascades. *Climatic Change*, *86*(1), 189–210. https://doi.org/10.1007/s10584-007-9294-8

Tamaddun, K., Kalra, A., & Ahmad, S. (2016). Identification of Streamflow Changes across the Continental United States Using Variable Record Lengths. *Hydrology*, *3*(2), 24. https://doi.org/10.3390/hydrology3020024

Tang, Y., Tang, Q., Wang, Z., Chiew, F. H. S., Zhang, X., & Xiao, H. (2019). Different Precipitation Elasticity of Runoff for Precipitation Increase and Decrease at Watershed Scale. *Journal of Geophysical Research: Atmospheres*, *124*(22), 11932–11943. https://doi.org/10.1029/2018JD030129

Tang, Y., Tang, Q., & Zhang, L. (2020). Derivation of Interannual Climate Elasticity of Streamflow. *Water Resources Research*, *56*(11), e2020WR027703. https://doi.org/10.1029/2020WR027703

Tsai, Y. (2017). The multivariate climatic and anthropogenic elasticity of streamflow in the Eastern United States. *Journal of Hydrology: Regional Studies*, *9*, 199–215. https://doi.org/10.1016/j.ejrh.2016.12.078

Turc, L. (1954). The water balance of soils. Relation between precipitation, evaporation and flow. *Ann. Agron*, *5*, 491–569.

United States Geological Survey and United States Department of Agriculture. (2020). *Watershed boundary dataset for the nation.* [dataset]. https://prd-tnm.s3.amazonaws.com/index.html?prefix=StagedProducts/Hydrography/WBD/National/GDB/

Vano, J. A., & Lettenmaier, D. P. (2014). A sensitivity-based approach to evaluating future changes in Colorado River discharge. *Climatic Change*, *122*(4), 621–634. https://doi.org/10.1007/s10584-013-1023-x

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., López-Moreno, J. I., Beguería, S., Noguera, I., Harrigan, S., & Vidal, J.-P. (2019). Climate, irrigation, and land-cover change explain streamflow trends in countries bordering the Northeast Atlantic. *Geophysical Research Letters*, *0*(ja). https://doi.org/10.1029/2019GL084084

Villarini, G., Howell Taylor, S., Wobus, C., Hecht, J., White, K., Baker, B., Gilroy, K., Olsen, J., & Raff, D. (2018). *Floods and nonstationarity—A review*.

Villarini, G., Serinaldi, F., Smith, J. A., & Krajewski, W. F. (2009). On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, *45*(8). https://doi.org/10.1029/2008WR007645

Villarini, G., Smith, J. A., Serinaldi, F., Bales, J., Bates, P. D., & Krajewski, W. F. (2009). Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources*, *32*(8), 1255–1266. https://doi.org/10.1016/j.advwatres.2009.05.003

Vörösmarty, C. J., Federer, C. A., & Schloss, A. L. (1998). Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling. *Journal of Hydrology*, *207*(3), 147–169. https://doi.org/10.1016/S0022-1694(98)00109-7

Wang, C., Yang, J., Myint, S. W., Wang, Z.-H., & Tong, B. (2016). Empirical modeling and spatio-temporal patterns of urban evapotranspiration for the Phoenix metropolitan area, Arizona. *GIScience & Remote Sensing*, *53*(6), 778–792. https://doi.org/10.1080/15481603.2016.1243399

Wang, D. (2012). Evaluating interannual water storage changes at watersheds in Illinois based on long-term soil moisture and groundwater level data. *Water Resources Research*, *48*(3). https://doi.org/10.1029/2011WR010759

Wang, D., & Hejazi, M. (2011). Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States. *Water Resources Research*, *47*(10). https://doi.org/10.1029/2010WR010283

Wang-Erlandsson, L., Fetzer, I., Keys, P. W., van der Ent, R. J., Savenije, H. H. G., & Gordon, L. J. (2018). Remote land use impacts on river flows through atmospheric teleconnections. *Hydrology and Earth System Sciences*, *22*(8), 4311–4328. https://doi.org/10.5194/hess-22-4311-2018

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Wey, H.-W., Lo, M.-H., Lee, S.-Y., Yu, J.-Y., & Hsu, H.-H. (2015). Potential impacts of wintertime soil moisture anomalies from agricultural irrigation at low latitudes on regional and global climates. *Geophysical Research Letters*, *42*(20), 8605–8614. https://doi.org/10.1002/2015GL065883

Williamson, E. J., Aitken, Z., Lawrie, J., Dharmage, S. C., Burgess, J. A., & Forbes, A. B. (2014). Introduction to causal diagrams for confounder selection. *Respirology*, *19*(3), 303–311. https://doi.org/10.1111/resp.12238

Winter, T. C. (2000). *Ground Water and Surface Water: A Single Resource*. DIANE Publishing.

Wittenberg, H. (1999). Baseflow recession and recharge as nonlinear storage processes. *Hydrological Processes*, *13*(5), 715–726. https://doi.org/10.1002/(SICI)1099-1085(19990415)13:5<715::AID-HYP775>3.0.CO;2-N

Woodhouse, C. A., Pederson, G. T., Morino, K., McAfee, S. A., & McCabe, G. J. (2016). Increasing influence of air temperature on upper Colorado River streamflow. *Geophysical Research Letters*, *43*(5), 2174–2181. https://doi.org/10.1002/2015GL067613

Wrede, S., Fenicia, F., Martínez-Carreras, N., Juilleret, J., Hissler, C., Krein, A., Savenije, H. H. G., Uhlenbrook, S., Kavetski, D., & Pfister, L. (2015). Towards more systematic perceptual model development: A case study using 3 Luxembourgish catchments. *Hydrological Processes*, *29*(12), 2731–2750. https://doi.org/10.1002/hyp.10393

Wu, C., Yeh, P. J.-F., Hu, B. X., & Huang, G. (2018). Controlling factors of errors in the predicted annual and monthly evaporation from the Budyko framework. *Advances in Water Resources*, *121*, 432–445. https://doi.org/10.1016/j.advwatres.2018.09.013

Yang, H., & Yang, D. (2011). Derivation of climate elasticity of runoff to assess the effects of climate change on annual runoff. *Water Resources Research*, *47*(7). https://doi.org/10.1029/2010WR009287

Yang, W., Yang, H., Yang, D., & Hou, A. (2021). Causal effects of dams and land cover changes on flood changes in mainland China. *Hydrology and Earth System Sciences*, *25*(5), 2705–2720. https://doi.org/10.5194/hess-25-2705-2021

Yin, J., Slater, L. J., Khouakhi, A., Yu, L., Liu, P., Li, F., Pokhrel, Y., & Gentine, P. (2023). GTWS-MLrec: Global terrestrial water storage reconstruction by machine learning from 1940 to present. *Earth System Science Data Discussions*, 1–29. https://doi.org/10.5194/essd-2023-315

Zhang, H., & Delworth, T. L. (2018). Robustness of anthropogenically forced decadal precipitation changes projected for the 21st century. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-03611-3

Zhang, L., Zhao, F. F., & Brown, A. E. (2012). Predicting effects of plantation expansion on streamflow regime for catchments in Australia. *Hydrology and Earth System Sciences*, *16*(7), 2109–2121. https://doi.org/10.5194/hess-16-2109-2012

Zhang, M., Liu, N., Harper, R., Li, Q., Liu, K., Wei, X., Ning, D., Hou, Y., & Liu, S. (2017). A global review on hydrological responses to forest change across multiple spatial scales: Importance of scale, climate, forest type and hydrological regime. *Journal of Hydrology*, *546*, 44–59. https://doi.org/10.1016/j.jhydrol.2016.12.040

Zhang, W., Villarini, G., Vecchi, G. A., & Smith, J. A. (2018). Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. *Nature*, *563*(7731), 384–388. https://doi.org/10.1038/s41586-018-0676-z

Zhang, X., Dong, Q., Zhang, Q., & Yu, Y. (2020). A unified framework of water balance models for monthly, annual, and mean annual timescales. *Journal of Hydrology*, *589*, 125186. https://doi.org/10.1016/j.jhydrol.2020.125186

Zhang, Y., Viglione, A., & Blöschl, G. (2022). Temporal Scaling of Streamflow Elasticity to Precipitation: A Global Analysis. *Water Resources Research*, *58*(1), e2021WR030601. https://doi.org/10.1029/2021WR030601

Zhang, Y., Zheng, H., Zhang, X., Leung, L. R., Liu, C., Zheng, C., Guo, Y., Chiew, F. H. S., Post, D., Kong, D., Beck, H. E., Li, C., & Blöschl, G. (2023). Future global streamflow declines are probably more severe than previously estimated. *Nature Water*, 1–11. https://doi.org/10.1038/s44221-023-00030-7

Zhang, Y.-K., & Schilling, K. E. (2006). Increasing streamflow and baseflow in Mississippi River since the 1940s: Effect of land use change. *Journal of Hydrology*, *324*(1), 412–422. https://doi.org/10.1016/j.jhydrol.2005.09.033

Zhou, G., Wei, X., Chen, X., Zhou, P., Liu, X., Xiao, Y., Sun, G., Scott, D. F., Zhou, S., Han, L., & Su, Y. (2015). Global pattern for the effect of climate and land cover on water yield. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms6918

Zimmer, M. A., & Gannon, J. P. (2018). Run-off processes from mountains to foothills: The role of soil stratigraphy and structure in influencing run-off characteristics across high to low relief landscapes. *Hydrological Processes*, *32*(11), 1546–1560. https://doi.org/10.1002/hyp.11488

ADDIN ZOTERO_BIBL {"uncited":[],"omitted":[],"custom":[]} CSL_BIBLIOGRAPHY