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

Bailey J. Anderson[1], Manuela I. Brunner[2,3], Louise J. Slater[1], and Simon J. Dadson[1,4]

[1]School of Geography and the Environment, University of Oxford, Oxford, UK

5 [2]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

[3]Institute for Snow and Avalanche Research SLF, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Davos, Switzerland

[4]NERC Centre for Ecology and Hydrology, Wallingford, UK

Correspondence: Bailey J. Anderson (bailey.anderson@ouce.ox.ac.uk)

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. 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.

1. 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 alone. 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., 2018; Sankarasubramanian et al., 2001; Tang et al., 2020; Tsai, 2017), and more recently, at aggregated multi-annual scales (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 and Villarini, 2016a).

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 and Shaw, 2015; Slater and Villarini, 2016a). 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 and Stieglitz, 2011).

Understanding the sensitivity of each of these components of the flow regime is important considering their unique roles in determining resilience and adaptability to climatic change. For instance, low flows are highly relevant for riverine ecology, water quality, and water availability for out-of-channel water uses like irrigation, power generation, and municipal water supply (Cooper et al., 2018; Smakhtin, 2001). High flows frequently correspond to flood events, and understanding their distributions and probability is essential for flood frequency estimation and infrastructure planning, among other things (François et al., 2019). The typical approaches of estimating elasticity for a single point along the flow distribution are insufficient for the objective of characterizing flow response to precipitation change, since the elasticity of the central summary of the distribution is unlikely to capture hydrologic behaviour in either low or high flow percentiles.

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 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 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:

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

2. What catchment attributes best explain between-catchment variation in $\varepsilon_{c,P}$ shape?

2. Methods
2.1 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., 2022). Gridded monthly precipitation and temperature (4 km resolution) were extracted from the Oregon State PRISM project using the R package prism (Edmund and Bell, 2020). 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, 1963; Lu et al., 2007) 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., 2007). 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.

2.2 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, are capable of controlling for a large portion of omitted variable bias, allowing for a more causal interpretation of regression results (Croissant and Millo, 2018; Hsiao, 1995; 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).

$$ln\left(Q_{i,t}^q\right) = \alpha_{i,t} + \varepsilon_P^q \ln\left(P_{i,t}\right) + \varepsilon_E^q \ln\left(E_{i,t}\right) + \eta_{i,t}^q \qquad 1$$

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.

160 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 and 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, 165 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. 170 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 175 its results are included in Appendices A and B.

2.3 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 180 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 185 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, 190 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 195 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.

2.4 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.

3.   Results
3.1 Normalized elasticity curves

Figure 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.

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 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 3). For instance, in the summer, 78% of catchments exhibit a flat elasticity curve (Figure 3 summer: A; C). However, while skewed towards zero, the distribution of possible elasticity magnitude is widespread (Figure 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.

3.2 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.

4. 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 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 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 2 summer; Figure 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 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 and Shaw, 2015; Slater and Villarini, 2016b). 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 and 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 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 4 spring). The shape may also reflect, in part, the climatic drivers dominant over different regions.

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 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 4). Interestingly, at the annual timescale, curve type C (negative) catchments are in some ways similar to those with

385   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

390   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 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

395   type C curves may differ from positive $\varepsilon_{c,p}$s predominantly in that they exhibit highly flashy low flow behaviour (Figure 4).

4.1 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

400   River at Van Buren (gauge id: 07067000), and Reddies River at North Wilkesboro (gauge id: 02111500) are included in Figure 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 and Gannon, 2018), but has

405   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 1.

At the seasonal timescale, both of the Ozarks catchments (Figure 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, it is clear that the catchment with young Karstic geology (06918460) and comparatively less long-

410   term storage experiences a higher absolute magnitude of elasticity to precipitation (Figure 5 A) when compared to its counterpart. This is particularly clear in summer, where the curve shape is similar (Figure 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

415   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;

420   Figure 5, represented in green), has relatively similar characteristics to 07067000 (Figure 5; Table 1) and exhibits similar curve structure at the annual and seasonal timescales, although with a slightly flatter overall condition.

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 A and Figure B1) 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., 2016, 2015), how groundwater contributes to flood-generation (Berghuijs and 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 A) – a method which may be well suited to regional calculation of elasticity and estimation in ungauged basins.

5. 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.

485 **6. Figures**

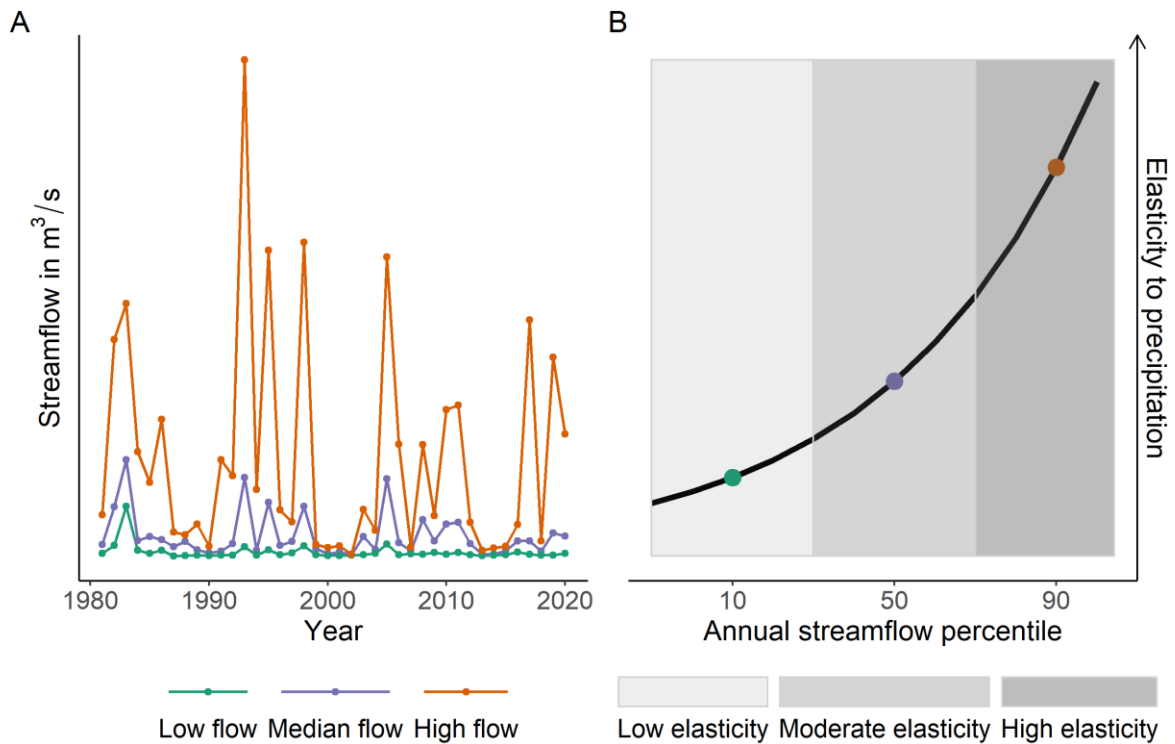



Figure 1. Conceptual diagram demonstrating how to read an elasticity curve. Where plot panel A. shows hypothetical high, low, and median annual streamflow (10th, 50th and 90th percentiles of the flow distribution in each year) and plot panel B. shows the hypothesised relative elasticity of each of these 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.

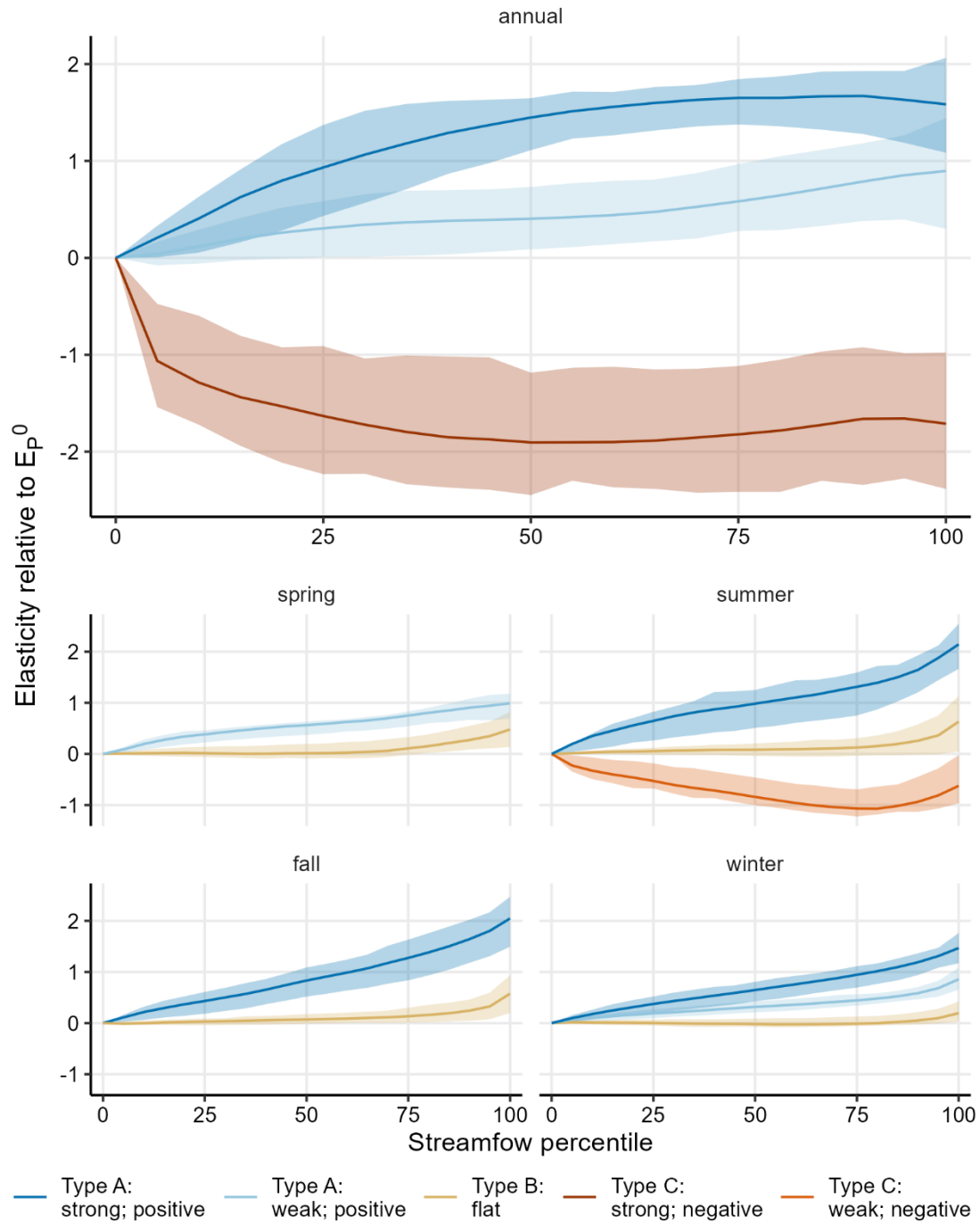



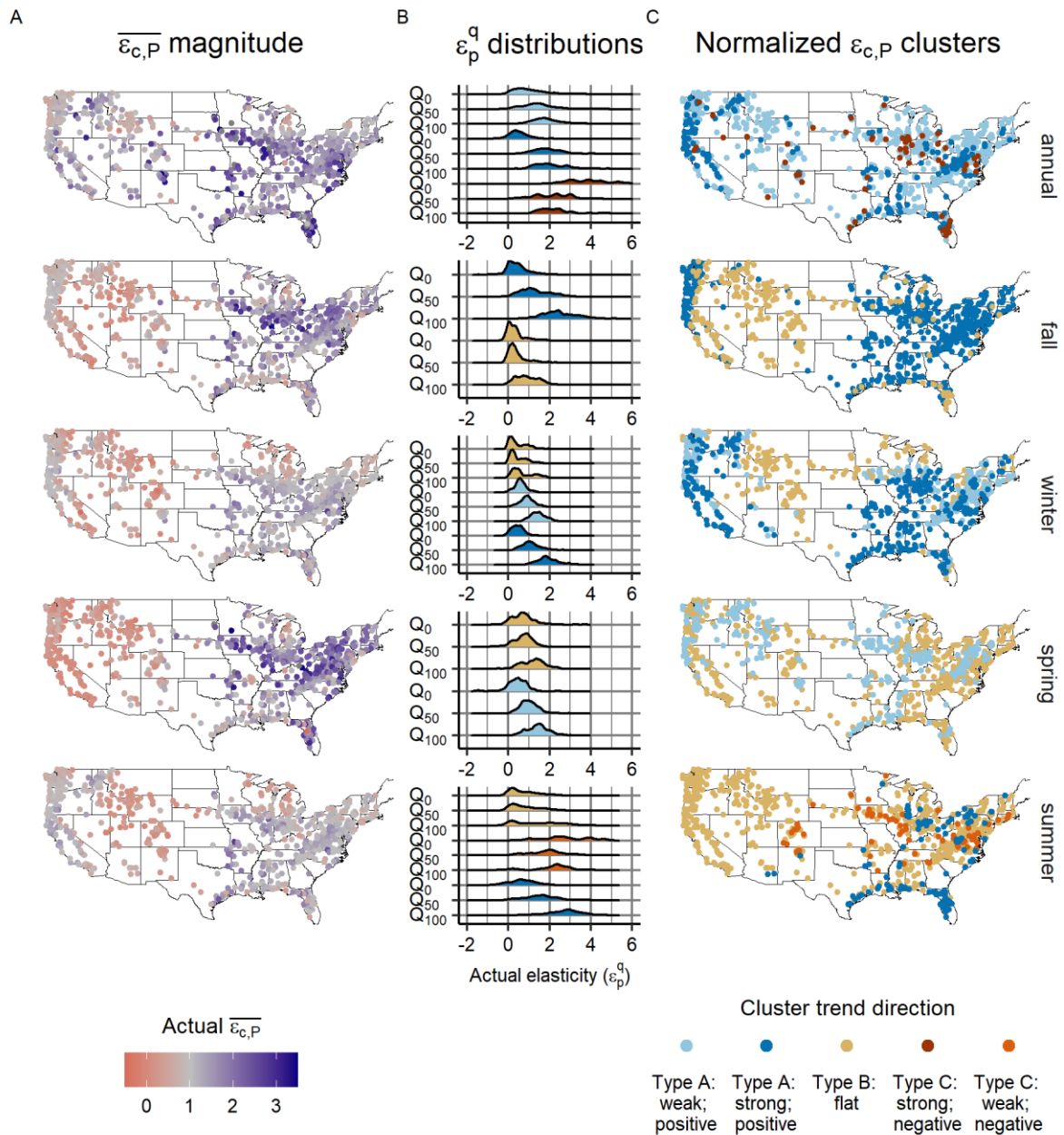
Figure 2. Normalized elasticity curves, shows the curves resulting from the single catchment linear models (lm) where the line is the mean of the distribution of elasticity point estimates (for a cluster of sites) and the bands are the inter-quartile range. Note that spring and fall have 2 clusters while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

490

Figure 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.

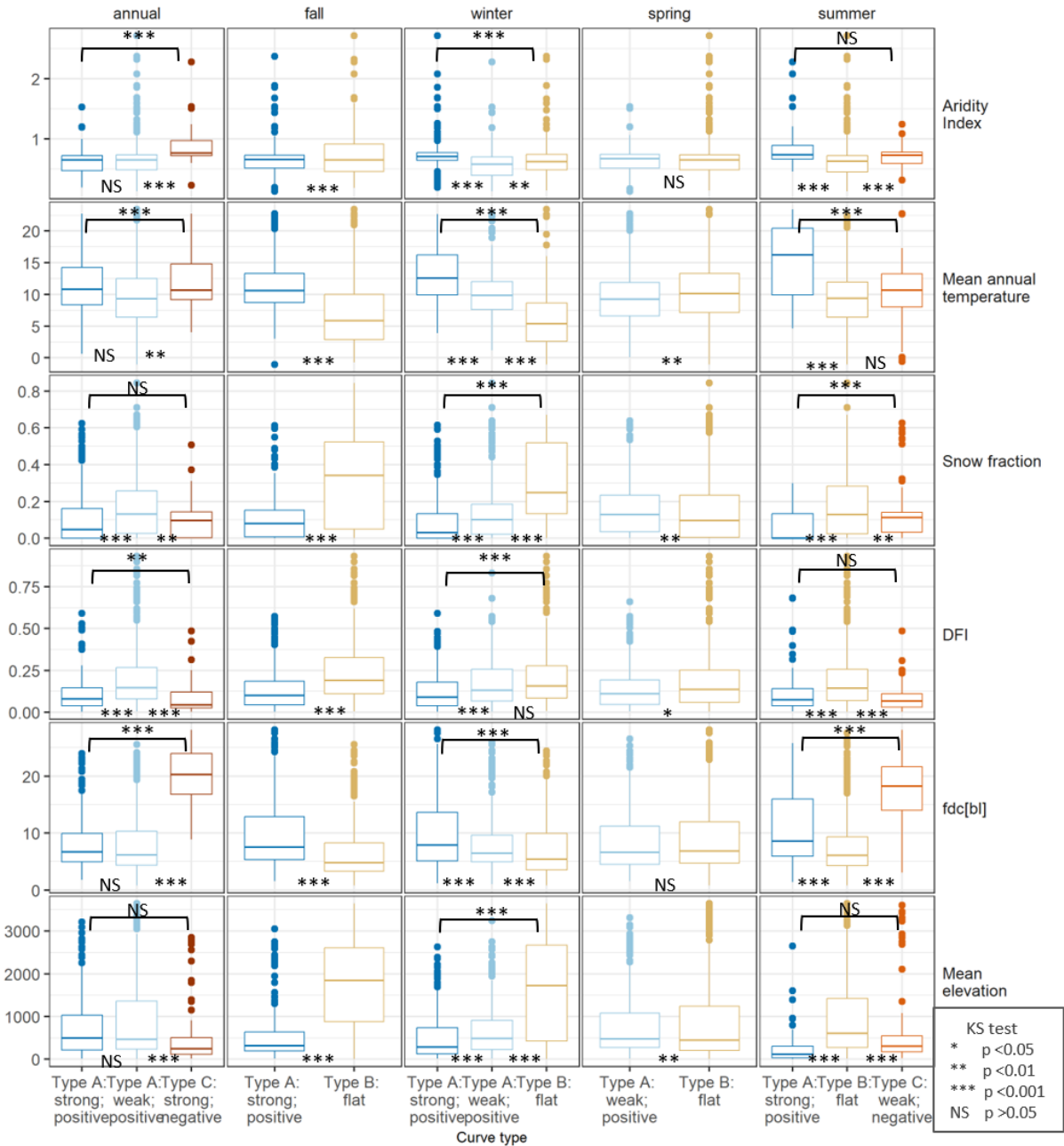



Figure 4. Boxplots showing the distributions of static catchment attributes split by time period and cluster membership. Significance is shown between each box for neighbouring distribution plots and the significance of the difference between the first and last distribution in each time period is plotted 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.

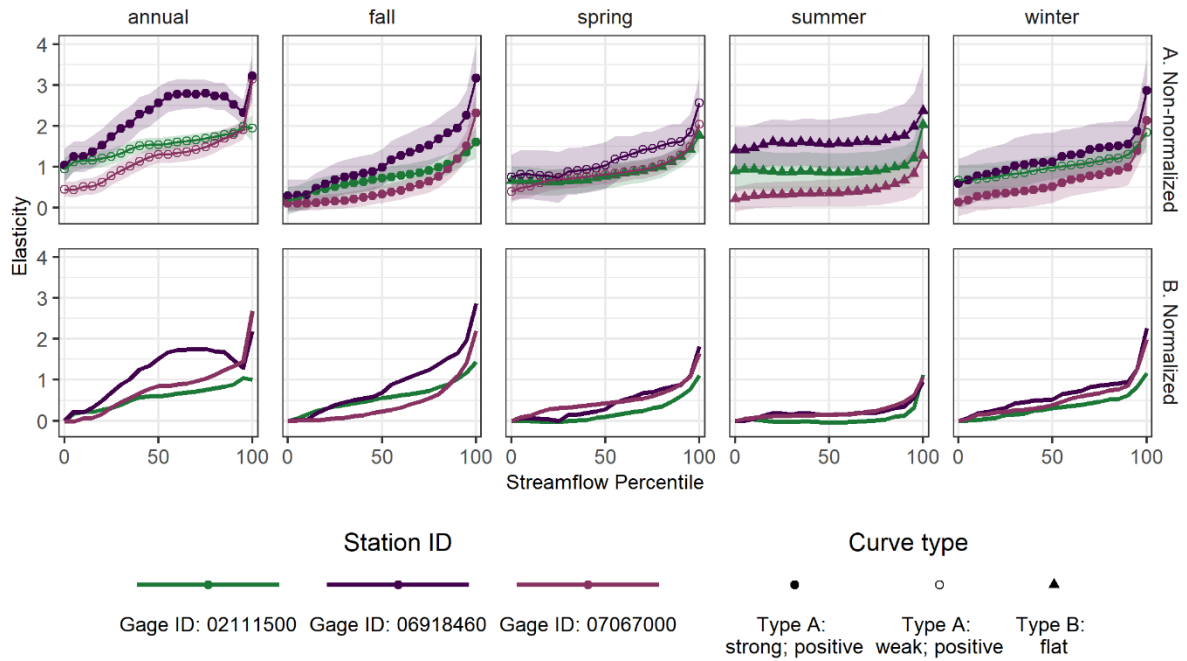



Figure 5. Examples of elasticity curves from three example catchments. Turnback Creek above Greenfield (gauge id: 06918460), Current River at Van Buren (gauge id: 07067000), and Reddies River at North Wilkesboro (gauge id: 02111500). Panel A shows the non-normalized curves to demonstrate actual elasticity, 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.

495

Table 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

7. Appendix A: Panel regression model results
A1. Panel model design

In order to further validate the elasticity estimates, we constructed a fixed-effects panel regression model (Equation A1) 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 A1) is used. The panel model is represented by:

$$ln\left(Q_{i,t}^q\right) = \alpha_{i,t} + \beta_1 \, ln\left(P_{i,t}\right) + \beta_2 \, ln\left(E_{i,t}\right) + \varepsilon_P^{g,q} \, ln\left(P_{i,t}\right) g_i + \varepsilon_E^{g,q} \, ln\left(E_{i,t}\right) g_i + \eta_{i,t}^q \qquad A1$$

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 A1) 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. Appendix B: Figures and tables in appendices

Table B1. 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	$$\text{Aridity} = \left(\frac{\overline{PET}}{\overline{P}}\right) * 100$$	The aridity index as a percentage of mean potential evaporation ($\overline{PET}$) divided by mean precipitation ($\overline{P}$).
Runoff Coefficient	$$RC = \frac{\overline{Q}/D}{\overline{P}} * 100$$	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$_b$	$$fdc_b = \frac{\ln(Q_{33}) - \ln(Q_{66})}{(0.66 - 0.33)}$$	Slope of the annual flow duration curve calculated with daily flow between the 33rd and 66th flow exceedance probabilities
fdc$_{bu}$	$$fdc_{bu} = \frac{\ln(Q_0) - \ln(Q_{32})}{0.32}$$	Slope of the annual flow duration curve calculated with daily flow between the 0th and 32nd flow exceedance probabilities
fdc$_{bl}$	$$fdc_{bl} = \frac{\ln(Q_{67}) - \ln(Q_{100})}{(1 - 0.67)}$$	Slope of the annual flow duration curve calculated with daily flow between the 67th and 100th flow exceedance probabilities

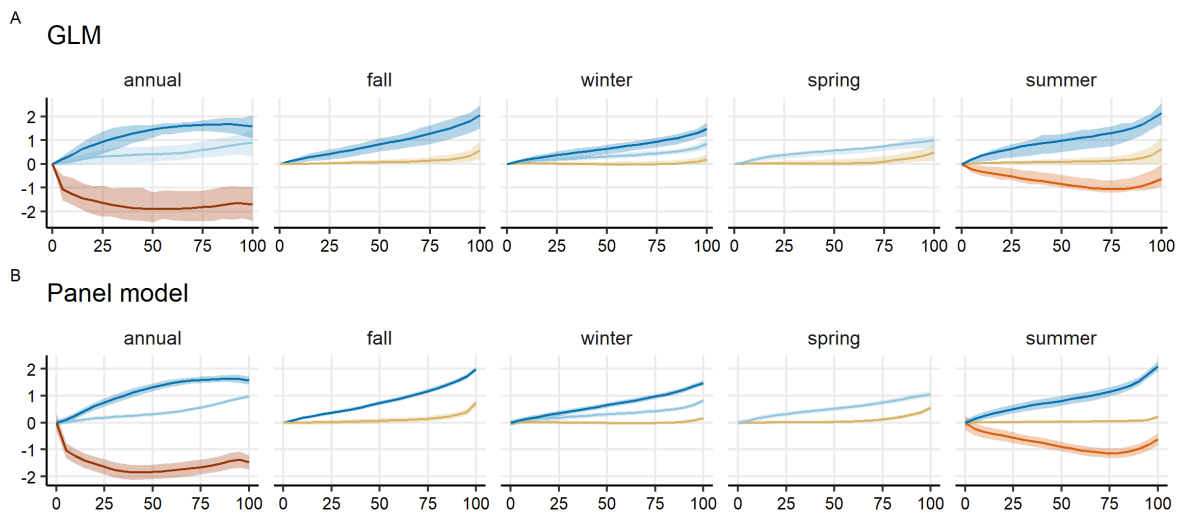
Annual temperature		Mean annual temperature
Mean catchment elevation		In meters (Falcone, 2017)
Latitude		Latitude at gage site (Falcone, 2017)
Drainage area		In Km^2 (Falcone, 2017)
Average catchment slope		In degrees (Falcone, 2017)
Coefficient of variation	$CV = \dfrac{sd(Q)}{\overline{Q}}$	CV of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

545

Figure A1. Elasticity curves as estimated using the single-site regression models (Panel A) and the aggregated panel regression models described in appendix A (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 2 here to facilitate comparison.

550

Code and data availability. Datasets are publicly 450 available as of December 2, 2022 at: total upstream dam storage, average annual runoff, drainage area, mean catchment elevation, latitude and average catchment slope are available through GAGES II at (https://water.usgs.gov/GIS/dsdl/basinchar_and in Falcone (2011); watershed boundaries are also available through GAGES II at

555   (https://water.usgs.gov/GIS/dsdl/gagesII_9322_point_shapefile.zip) and in Falcone (2017); climate data is available from PRISM at (https://www.prism.oregonstate.edu/recent/) and can be downloaded using the prism R package (Edmund and Bell, 2020). Streamflow data can be downloaded from the National Water Information System (NWIS) using the R package dataRetrieval (Cicco et al., 2022). R code for the complete analysis is available as of December 2, 2022 at: https://doi.org/10.5281/zenodo.7391227 (Anderson, 2022)

560   **Author contributions.** Conceptualizaton, Methodology: BJA, MIB, and LJS; Data curation, Formal analysis, Investigation, Project administration, Software, Visualization, Writing – original draft preparation: BJA; Supervision, Validation: LJS, SJD, MIB; Writing – review & editing: LJS, SJD, MIB, BJA

Competing interests. At least one of the co-authors is on the editorial board of HESS. The authors have no other competing interests to declare.

565   **Acknowledgements.** LJS is supported by UKRI (MR/V022008/1). BJA is supported by the Clarendon scholarship and Hertford College, Oxford. We additionally thank the attendees of the 2022 STAHY workshop, the hydrology research group at the University of Freiburg, the Water lab at the University of Oxford, as well as Ross Woods and Linda Speight for their valuable input. In particular, we thank Richard Vogel, Michael Stoelzle, Markus Weiler, and Kerstin Stahl whose comments led to methodological and conceptual 570   improvements in the work.

9.   References

Allaire, M. C., Vogel, R. M., and Kroll, C. N.: The hydromorphology of an urbanizing watershed using

      multivariate elasticity, Adv. Water Resour., 86, 147–154,

      https://doi.org/10.1016/j.advwatres.2015.09.022, 2015.

575 Anderson, B. J., Slater, L. J., Dadson, S. J., Blum, A. G., and Prosdocimi, I.: Statistical Attribution of the

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

      Approaches, Water Resour. Res., 58, e2021WR030742, https://doi.org/10.1029/2021WR030742, 2022.

Bassiouni, M., Vogel, R. M., and Archfield, S. A.: Panel regressions to estimate low-flow response to rainfall

      variability in ungaged basins, Water Resour. Res., 52, 9470–9494,

580   https://doi.org/10.1002/2016WR018718, 2016.

Berghuijs, W. R. and Slater, L. J.: Groundwater shapes North American river floods, Environ. Res. Lett., 18,

      034043, https://doi.org/10.1088/1748-9326/acbecc, 2023.

Berghuijs, W. R., Larsen, J. R., Emmerik, T. H. M. van, and Woods, R. A.: A Global Assessment of Runoff

Sensitivity to Changes in Precipitation, Potential Evaporation, and Other Factors, Water Resour. Res.,

585   53, 8475–8486, https://doi.org/10.1002/2017WR021593, 2017.

Blum, A. G., Ferraro, P. J., Archfield, S. A., and Ryberg, K. R.: Causal Effect of Impervious Cover on Annual

Flood Magnitude for the United States, Geophys. Res. Lett., 47, e2019GL086480,

https://doi.org/10.1029/2019GL086480, 2020.

Brunner, M. I., Swain, D. L., Gilleland, E., and Wood, A. W.: Increasing importance of temperature as a

590   contributor to the spatial extent of streamflow drought, Environ. Res. Lett., 16, 024038,

https://doi.org/10.1088/1748-9326/abd2f0, 2021.

Chiew, F.: Estimation of rainfall elasticity of streamflow in Australia, https://doi.org/10.1623/hysj.51.4.613,

2006.

Chiew, F., Peel, M., McMahon, T., and Siriwardena, L.: Precipitation elasticity of streamflow in catchments

595   across the world, undefined, 2006.

Cooper, M. G., Schaperow, J. R., Cooley, S. W., Alam, S., Smith, L. C., and Lettenmaier, D. P.: Climate

Elasticity of Low Flows in the Maritime Western U.S. Mountains, Water Resour. Res., 54, 5602–5619,

https://doi.org/10.1029/2018WR022816, 2018.

Croissant, Y. and Millo, G. (Eds.): Endogeneity, in: Panel Data Econometrics with R, John Wiley & Sons, Ltd,

600   Chichester, UK, 139–159, https://doi.org/10.1002/9781119504641.ch6, 2018.

DeCicco, L., Hirsch, R., Lorenz, D., Watkins, D., and Johnson, M.: dataRetrieval: Retrieval Functions for

USGS and EPA Hydrologic and Water Quality Data, 2022.

Edmund, H. and Bell, K.: prism: Access Data from the Oregon State Prism Climate Project, 2020.

Falcone, J. A.: U.S. Geological Survey GAGES-II time series data from consistent sources of land use, water

605   use, agriculture, timber activities, dam removals, and other historical anthropogenic influences,

https://doi.org/10.5066/F7HQ3XS4, 2017.

François, B., Schlef, K. E., Wi, S., and Brown, C. M.: Design considerations for riverine floods in a changing

climate – A review, J. Hydrol., 574, 557–573, https://doi.org/10.1016/j.jhydrol.2019.04.068, 2019.

Gnann, S. J., McMillan, H. K., Woods, R. A., and Howden, N. J. K.: Including Regional Knowledge Improves

610   Baseflow Signature Predictions in Large Sample Hydrology, Water Resour. Res., 57,

e2020WR028354, https://doi.org/10.1029/2020WR028354, 2021.

Hamon, W. R.: Computation of direct runoff amounts from storm rainfall, Int. Assoc. Sci. Hydrol. Publ., 63, 52–62, 1963.

Harman, C. J., Troch, P. A., and Sivapalan, M.: 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 Resour. Res., 47, https://doi.org/10.1029/2010WR009656, 2011.

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

Hsiao, C.: Panel Data Analysis — Advantages and Challenges, Panel Data Anal., 63, 1995.

Ivancic, T. J. and Shaw, S. B.: Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge, Clim. Change, 133, 681–693, https://doi.org/10.1007/s10584-015-1476-1, 2015.

Koehn, C. R., Petrie, M. D., Bradford, J. B., Litvak, M. E., and Strachan, S.: Seasonal Precipitation and Soil Moisture Relationships Across Forests and Woodlands in the Southwestern United States, J. Geophys. Res. Biogeosciences, 126, e2020JG005986, https://doi.org/10.1029/2020JG005986, 2021.

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

Li, D., Wrzesien, M. L., Durand, M., Adam, J., and Lettenmaier, D. P.: How much runoff originates as snow in the western United States, and how will that change in the future?, Geophys. Res. Lett., 44, 6163–6172, https://doi.org/10.1002/2017GL073551, 2017.

Lu, J., Sun, G., McNulty, S. G., and Amatya, D. M.: A Comparison of Six Potential Evapotranspiration Methods for Regional Use in the Southeastern United States1, JAWRA J. Am. Water Resour. Assoc., 41, 621–633, https://doi.org/10.1111/j.1752-1688.2005.tb03759.x, 2007.

Milly, P. C. D., Kam, J., and Dunne, K. A.: On the Sensitivity of Annual Streamflow to Air Temperature, Water Resour. Res., 54, 2624–2641, https://doi.org/10.1002/2017WR021970, 2018.

Murtagh, F. and Contreras, P.: Algorithms for hierarchical clustering: an overview, WIREs Data Min. Knowl. Discov., 2, 86–97, https://doi.org/10.1002/widm.53, 2012.

Nichols, A.: Causal Inference with Observational Data, Stata J., 7, 507–541, https://doi.org/10.1177/1536867X0800700403, 2007.

Patil, S. and Stieglitz, M.: Hydrologic similarity among catchments under variable flow conditions, Hydrol. Earth Syst. Sci., 15, 989–997, https://doi.org/10.5194/hess-15-989-2011, 2011.

Price, K.: Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review:, Prog. Phys. Geogr., https://doi.org/10.1177/0309133311402714, 2011.

Prudhomme, C., Crooks, S., Kay, A. L., and Reynard, N.: Climate change and river flooding: part 1 classifying the sensitivity of British catchments, Clim. Change, 119, 933–948, 2013.

Saft, M., Western, A. W., Zhang, L., Peel, M. C., and Potter, N. J.: The influence of multiyear drought on the annual rainfall-runoff relationship: An Australian perspective, Water Resour. Res., 51, 2444–2463, https://doi.org/10.1002/2014WR015348, 2015.

Saft, M., Peel, M. C., Western, A. W., and Zhang, L.: Predicting shifts in rainfall-runoff partitioning during multiyear drought: Roles of dry period and catchment characteristics, Water Resour. Res., 52, 9290–9305, https://doi.org/10.1002/2016WR019525, 2016.

Sankarasubramanian, A., Vogel, R. M., and Limbrunner, J. F.: Climate elasticity of streamflow in the United States, Water Resour. Res., 37, 1771–1781, https://doi.org/10.1029/2000WR900330, 2001.

Schaake, J. C.: From climate to flow, in: Climate change and US water resources, vol. 8, 177–206, 1990.

Searcy, J. K.: Flow-duration curves, Water Supply Paper, U.S. Govt. Print. Off., https://doi.org/10.3133/wsp1542A, 1959.

Slater, L. J. and Villarini, G.: Recent trends in U.S. flood risk, Geophys. Res. Lett., 43, 12,428-12,436, https://doi.org/10.1002/2016GL071199, 2016a.

Slater, L. J. and Villarini, G.: Recent trends in U.S. flood risk, Geophys. Res. Lett., 43, 12,428-12,436, https://doi.org/10.1002/2016GL071199, 2016b.

Smakhtin, V. U.: Low flow hydrology: a review, J. Hydrol., 240, 147–186, https://doi.org/10.1016/S0022-1694(00)00340-1, 2001.

Stoelzle, M., Schuetz, T., Weiler, M., Stahl, K., and Tallaksen, L. M.: Beyond binary baseflow separation: a delayed-flow index for multiple streamflow contributions, Hydrol. Earth Syst. Sci., 24, 849–867, https://doi.org/10.5194/hess-24-849-2020, 2020.

Tang, Y., Tang, Q., Wang, Z., Chiew, F. H. S., Zhang, X., and Xiao, H.: Different Precipitation Elasticity of Runoff for Precipitation Increase and Decrease at Watershed Scale, J. Geophys. Res. Atmospheres, 124, 11932–11943, https://doi.org/10.1029/2018JD030129, 2019.

Tang, Y., Tang, Q., and Zhang, L.: Derivation of Interannual Climate Elasticity of Streamflow, Water Resour. Res., 56, e2020WR027703, https://doi.org/10.1029/2020WR027703, 2020.

Tsai, Y.: The multivariate climatic and anthropogenic elasticity of streamflow in the Eastern United States, J. Hydrol. Reg. Stud., 9, 199–215, https://doi.org/10.1016/j.ejrh.2016.12.078, 2017.

675   Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, J. Am. Stat. Assoc., 58, 236–244, https://doi.org/10.1080/01621459.1963.10500845, 1963.

Zhang, Y., Viglione, A., and Blöschl, G.: Temporal Scaling of Streamflow Elasticity to Precipitation: A Global Analysis, Water Resour. Res., 58, e2021WR030601, https://doi.org/10.1029/2021WR030601, 2022.

Zimmer, M. A. and Gannon, J. P.: Run-off processes from mountains to foothills: The role of soil stratigraphy

680   and structure in influencing run-off characteristics across high to low relief landscapes, Hydrol. Process., 32, 1546–1560, https://doi.org/10.1002/hyp.11488, 2018.