Constructing a geography of heavy-tailed flood distributions: insights from common streamflow dynamics.

Hsing-Jui Wang[1], Ralf Merz[2], and Stefano Basso[3]

[1]Helmholtz-Centre for Environmental Research UFZ
[2]Helmholtz Centre for Environmental Research (UFZ)
[3]Norwegian Institute for Water Research (NIVA), Oslo, Norway

October 5, 2023

Abstract

Heavy-tailed flood distributions depict the higher occurrence probability of extreme floods. Understanding the spatial distribution of heavy tail floods is essential for effective risk assessment. Conventional methods often encounter data limitations, leading to uncertainty across regions. To address this challenge, we utilize hydrograph recession exponents derived from common streamflow dynamics, which have proven to be a robust indicator of flood tail propensity across analyses with varying data lengths. Analyzing extensive datasets from Germany, the United Kingdom (UK), Norway, and the United States (US), we uncover distinct patterns: prevalent heavy tails in Germany and the UK, diverse behavior in the US, and predominantly nonheavy tails in Norway. The regional tail behavior has been observed in relation to the interplay between terrain and meteorological characteristics, and we further conducted quantitative analyses to assess the influence of hydroclimatic conditions using Köppen classifications. Notably, temporal variations in catchment storage are a crucial mechanism driving highly nonlinear catchment responses that favor heavy-tailed floods, often intensified by concurrent dry periods and high temperatures. Furthermore, this mechanism is influenced by various flood generation processes, which can be shaped by both hydroclimatic seasonality and catchment scale. These insights deepen our understanding of the interplay between climate, physiographical settings, and flood behavior, while highlighting the utility of hydrograph recession exponents in flood hazard assessment.

Constructing a geography of heavy-tailed flood distributions: insights from common streamflow dynamics

Hsing-Jui Wang[1], Ralf Merz[1,2], and Stefano Basso[3]

[1]Department of Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale), 06120, Germany,

[2]Institute of Geosciences and Geography, Martin-Luther University Halle-Wittenberg, Halle (Saale), 06120, Germany,

[3]Norwegian Institute for Water Research (NIVA), Oslo, 0579, Norway

Correspondence to: Hsing–Jui Wang (hsing-jui.wang@ufz.de)

Key Points

1. Regional propensity of flood tail behavior is shown across a diverse geographical range based on the analysis of common streamflow dynamics.
2. Temporal variability in catchment storage, driven by high evapotranspiration and dry soils, is a key mechanism for heavy-tailed floods.
3. Examining flood generation processes aids in unraveling the connections between flood tail behavior and seasons or catchment sizes.

Abstract

Heavy-tailed flood distributions depict the higher occurrence probability of extreme floods. Understanding the spatial distribution of heavy tail floods is essential for effective risk assessment. Conventional methods often encounter data limitations, leading to uncertainty across regions. To address this challenge, we utilize hydrograph recession exponents derived from common streamflow dynamics, which have proven to be a robust indicator of flood tail propensity across analyses with varying data lengths. Analyzing extensive datasets from Germany, the United Kingdom (UK), Norway, and the United States (US), we uncover distinct patterns: prevalent heavy tails in Germany and the UK, diverse behavior in the US, and predominantly nonheavy tails in Norway. The regional tail behavior has been observed in relation to the interplay between terrain and meteorological characteristics, and we further conducted quantitative analyses to assess the influence of hydroclimatic conditions using Köppen classifications. Notably, temporal variations in catchment storage are a crucial mechanism driving highly nonlinear catchment responses that favor heavy-tailed floods, often intensified by concurrent dry periods and high temperatures. Furthermore, this mechanism is influenced by various flood generation processes, which can be shaped by both hydroclimatic seasonality and catchment scale. These insights deepen our understanding of the interplay between climate, physiographical settings, and flood behavior, while highlighting the utility of hydrograph recession exponents in flood hazard assessment.

1 Introduction

Floods are devastating natural hazards that pose significant risks to infrastructure, property, and human life (McDermott, 2022; Bevere and Remondi, 2022). The unprecedented magnitude of extreme floods often characterizes these hazards, which is better depicted by the heavy-tailed behavior exhibited in flood frequency distributions (Smith et al., 2018; Merz et al., 2021; Merz et al., 2022). The concept of heavy-tailed behavior finds broad application in various fields to describe the likelihood of extreme event occurrences (Katz, 2002; Kondor et al., 2014; Malamud, 2004; Sartori and Schiavo, 2015; Wang et al., 2022). In particular, it is widely recognized as a prevalent feature in hydrologic extremes (Papalexiou and Koutsoyiannis, 2013; Smith et al., 2018).

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

In the review of Merz et al. (2022), it becomes evident that multiple hydro-physiographic characteristics interact within a complex system, collectively shaping flood tail behavior. Specifically, the interplay between characteristic flood generation (Bernardara et al., 2008; Thorarinsdottir et al., 2018), the presence of mixed flood types (Morrison and Smith, 2002; Villarini and Smith, 2010), the tail heaviness of rainfall distributions (Gaume, 2006), catchment aridity (Molnar et al., 2006; Merz and Blöschl, 2009; Guo et al., 2014), and catchment area (Pallard et al., 2009; Villarini and Smith, 2010) are proposed as contributing factors to the nonlinearity of catchment responses. This nonlinearity is increasingly recognized as a plausible driver of heavy-tailed flood behavior (Fiorentino et al., 2007; Struthers and Sivapalan, 2007; Gioia et al., 2008; Rogger et al., 2012; Basso et al., 2015; Merz et al., 2022; Basso et al., 2023; Wang et al., 2023).

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

$$\frac{dq}{dt} = -B \cdot q^{a}$$

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

In our prior research (Wang et al., 2023), we introduced hydrograph recession exponents as a newly proposed indicator for heavy-tailed flood behavior. This indicator allows for an inference of heavy-tailed flood distributions based on physical mechanisms (i.e., typical hydrological processes within common streamflow dynamics). Importantly, it has shown its capacity to provide robust estimates for both short and long data records. Unlike traditional methods for identifying heavy tails, such as the upper tail ratio (Villarini et al., 2011; Lu et al., 2017) and the shape parameter of a fitted Generalized Extreme Value (GEV) distribution (Morrison and Smith, 2002; Villarini and Smith, 2010), which are sensitive to sample sizes (Wietzke et al., 2020), recession exponents offer more consistent estimates of flood tail heaviness across various data lengths. This characteristic makes it a valuable tool for analyzing regions with diverse gauge data records.

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

We organize the structure of this paper as follows: Section 2 describes the study areas and the hydrological data based on an extensive dataset composed of four countries, Section 3 describes the methods of estimation and validation of hydrograph recession exponents in identifying heavy-tailed flood behavior in the dataset, the framework of the analyses of spatial patterns of inferred heavy-tailed flood behavior, the framework of the analyses of seasonal dynamics of inferred heavy-tailed flood behavior, and statistical tests. In Section 4, we present the validation results of our heavy-tailed flood behavior index, along with analyses of the relationships between flood tail behavior and geographical spatial characteristics, seasonal patterns, and catchment scales in these comparable countries. Physical interpretations of the results and remarks from the literature are discussed in Section 5. The main conclusions are summarized in Section 6.

2 Study areas and data

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

(US), known for its diverse range of physiographical settings, enables us to validate and consolidate the transferability of our findings.

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

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

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

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

We collected daily continuous streamflow records from 575 gauges across the four study regions to conduct our analyses. The corresponding drainage areas range from 4 to 40,504 km^2. Our analysis was performed on a seasonal basis, considering spring (March-May), summer (June-August), autumn (September-November), and winter (December-February) to account for the seasonality of hydrograph recessions (Tashie et al., 2020) and flood distributions (Durrans et al., 2003). Each analysis conducted on a specific river gauge during a season was treated as a case study.

We excluded gauges located downstream of large dams in all four regions (Lehner et al., 2011; Wang et al., 2022). Consistent with previous studies (e.g., Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Basso et al., 2021; Basso et al., 2023), we chose case studies in Germany, the UK, and the US characterized by limited snowfall, which minimizes the potential transfer of water across seasons due to strong snow accumulation and melting. Specifically, this condition is defined as having an average daily temperature below zero degrees Celsius during precipitation events for over 50% of a season (Basso et al., 2021). However, recognizing that recession exponents can inherently capture both linear and nonlinear catchment responses, we intentionally included case studies in Norway, which are characterized by a dominant runoff generation process driven by snow dynamics. This deliberate inclusion provides a counter-verification, allowing us to explore the capability of the recession exponent as a measure of flood tail behavior in regions primarily characterized by snowmelt-driven flood generation processes. In summary, this analysis encompasses regions dominated by both rainfall-driven and snowmelt-driven floods, providing an extensive examination of these factors. These procedures resulted in a total of 1997 case studies, distributed as follows: 540 in spring, 520 in summer, 543 in autumn, and 394 in winter (refer to Table A1 for detailed information of each region).

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

3 Methods

3.1 Inferring Heavy Tails of Flood Distributions from Common Streamflow Dynamics

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

Within the PHEV framework, we obtain consistent expressions for the probability distributions of various flow metrics, including daily streamflow, ordinary peak flows (local flow peaks resulting from streamflow-producing rainfall events), and floods (flow maxima within a specified timeframe) (Basso et al., 2016). The description of runoff generation and streamflow dynamics provided by this framework has been successfully tested across a diverse range of hydroclimatic and physiographic conditions through a number of studies (Arai et al., 2020; Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Mejía et al., 2014; Müller et al., 2014; Müller et al., 2021; Pumo et al., 2014; Santos et al., 2018; Schaefli et al., 2013).

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

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

3.2 Validation of Hydrograph Recession Exponents as An Index of Heavy-Tailed Flood Behavior

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

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

We conduct our approach using three distinct empirical data distributions: daily streamflow, ordinary peaks, and monthly maxima. These multiple analyses strengthen our validation process and enhance the evaluation of our results. It's worth noting that our chosen benchmark, the empirical power law, may be influenced by fitting uncertainty due to data scarcity in certain cases, particularly when analyzing maxima. To mitigate this, we consider monthly maxima (Fischer and Schumann, 2016; Malamud and Turcotte, 2006) instead of the seasonal maxima previously used in the literature (e.g., Basso et al., 2021) in order to expand the sample size. Parallel analyses for cases with larger sample sizes (i.e., daily streamflow and ordinary peaks) provide more robust validation and lend support to the interpretation of results for maxima.

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

3.3 Analyses of Spatial and Seasonal Patterns of Inferred Flood Tail Behavior

We construct a geographical representation of inferred heavy-tailed flood behavior by utilizing estimated recession exponents derived from common streamflow dynamics across study countries and for each season. This representation serves as an evaluation of the propensity of heavy-tailed flood behavior across various regions and seasons. We simplify the seasonal results by identifying the dominant tail behavior, which refers to the majority of seasons exhibiting either heavy-tailed or nonheavy-tailed behavior, as the representative inferred flood tail behavior in the analysis of spatial pattern (Section 4.2).

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

To analyze seasonal patterns, we initially investigate the coherence of inferred flood tail behavior across seasons, focusing on consistency between heavy- or nonheavy-tailed behavior. Catchments with valid recession exponents from only one season are omitted from this analysis. As a result, the selection comprises 98 out of 98 catchments in Germany, 81 out of 82 in the UK, 79 out of 82 in Norway, and 290 out of 313 in the US. We also employ the Wilcoxon signed-rank test (Wilcoxon, 1945), a non-parametric statistical hypothesis test, at a significance level of 0.05 in this section. This test assesses whether the median of recession exponents (within a climate group on a seasonal basis) is above two, below two, or shows no significant difference from two (Figure 7).

4 Results

4.1 Effectiveness of Identifying Heavy-tailed Flood Behavior Using Common Discharge Dynamics

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

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

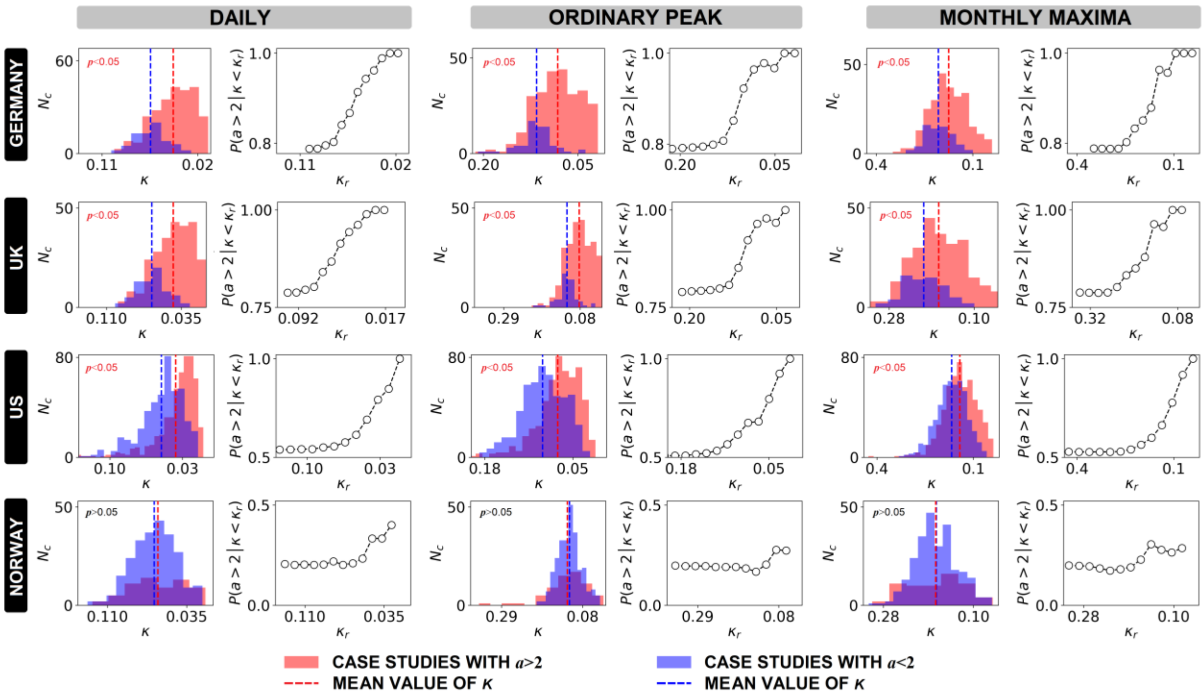


Figure 1. Effectiveness of identifying heavy-tailed flood behavior using hydrograph recession exponents. Case studies in each country are categorized into two groups: red histograms representing recession exponents a above two and blue histograms representing recession exponents a below two. In all three analyses (daily streamflow, ordinary peak flow, and monthly maxima), every case study is subjected to empirical power law fitting, resulting in a representative power law for the dataset, measured by the KS statistic κ (where $\kappa \in [0,\infty]$ and κ=0 signifies maximum reliability). The histograms portray the count of case studies $\boldsymbol{N_c}$ analyzed as a function of $\boldsymbol{\kappa}$ for two distinct groups. Dashed lines on the histogram plots indicate the means of the histograms. The means of two groups (a>2 and a<2) are subjected to Welch's t-test at a significance level of 0.05 to determine whether they are significantly different ($p < 0.05$) or not ($p > 0.05$). The line chart shows the accuracy of using the recession exponent to identify heavy-tailed behavior (denoted as $\boldsymbol{P(a > 2|\kappa < \kappa_r) = N_c(a > 2|\kappa < \kappa_r)/N_c(\kappa < \kappa_r)}$) as the $\boldsymbol{\kappa_r}$ threshold decreases (i.e., as the reliability of empirical power laws increases). The results for Germany are reproduced from Wang et al. (2023).

In Figure 2, we explore the correlation between the values of empirical power law exponents b and the values of recession exponents a for case studies confirmed to exhibit heavy-tailed behavior. This is achieved by utilizing the goodness-of-fit testing procedure of Clauset et al. (2009) to categorize case studies into ‘confirmed power-law-tailed case studies’ and ‘uncertain case studies.’ The former are depicted as black dots, while the latter are depicted as gray dots. The presence of a sizable number of uncertain case studies indicates the difficulty of establishing with certainty whether the underlying distribution of empirical data is or not a power law. This difficulty is often due to limited data availability, although the possibility that they indeed do not follow power laws cannot be excluded.

To highlight the correlation, we binned the confirmed power-law-tailed case studies and used red markers showing the median values of a and b (squares), the interquartile intervals of b (vertical bars), and the binning ranges of a. In each country, the composition of each bin encompasses one-seventh of the total number of case studies, except for Norway, where this fraction is adjusted to one-fifth due to the limited number of confirmed power-law-tailed cases. We calculated Spearman correlations $\boldsymbol{r_s}$ (Spearman, 1904) to test the correlation between a and b, which is valid for both linear and nonlinear associations between random variables. We found that a and b are significantly correlated at a significance level of 0.05 in Germany, the US, and the UK. In these three countries, a larger number of uncertain case studies emerge in the analysis of flow maxima compared to the analysis of daily streamflow and ordinary peak flow (respectively for daily streamflow, ordinary peaks, and flow maxima: 265, 270, and 352 out of 386 case studies in Germany; 258, 280, and 306 out of 325 case studies in the UK; and 589, 624, and 836 out of 980 case studies in the US). Since the same case studies have already been confirmed to exhibit power-law-tailed distributions in their daily streamflow and ordinary peak flow data, the increase of uncertain case studies in the analysis of flow maxima suggests that the greater level of uncertainty is due to limited data availability rather than indicating a rise in the number of non-power-law-tailed case studies.

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

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

Conversely, regions with lower recession exponents (below two), reflecting relatively linear responses in catchments, are more likely to signify nonheavy-tailed flood behavior.

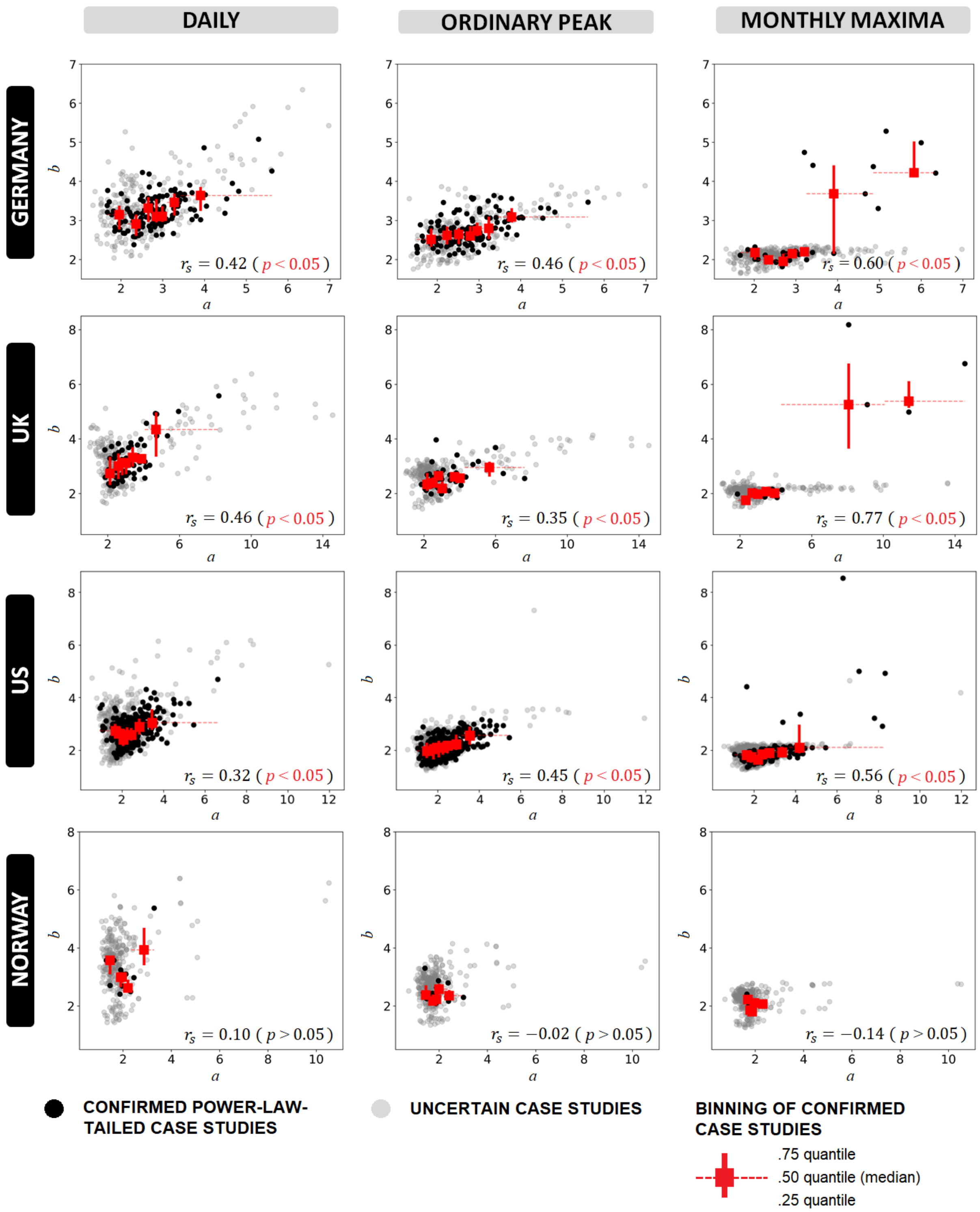


Figure 2. Empirical power law exponent *b* as a function of the hydrograph recession exponent *a* (physically-based index of heavy-tailed flood behavior). Case studies are classified

into groups of confirmed power-law-tailed (black dots) and uncertain (gray dots) case studies on the basis of the goodness-of-fit test (Clauset et al., 2009). The former group shows statistical confirmation that the data's distribution tail can be accurately characterized by a power law, indicating heavy-tailed behavior. Conversely, the latter group indicates our inability to statistically affirm whether the data follows a power law distribution or not. For the confirmed power-law-tailed case studies, the correlation between the empirical power law exponent *b* and the hydrograph recession exponent *a* is underlined by red markers. This correlation is quantified using the Spearman correlation coefficient $\boldsymbol{r_s}$ at a significance level of 0.05. The squares represent the median values of *a* and *b*, vertical bars indicate the interquartile intervals of *b*, and horizontal dashed bars indicate the binning ranges of *a*. In each country, the composition of each bin encompasses one-seventh of the total number of case studies, except for Norway, where this fraction is adjusted to one-fifth due to the constraint posed by the total number of confirmed power-law-tailed case studies. The count of the confirmed power-law-tailed case studies in the analyses of daily streamflows, ordinary peak flows, and monthly flow maxima are as follows: 121, 116, and 34 out of 386 case studies for Germany, respectively; 67, 45, and 19 out of 325 case studies for the UK, respectively; 391, 356, and 144 out of 980 case studies for the US, respectively; and 15, 17, and 6 out of 306 case studies for Norway, respectively. The results for Germany are reproduced from Wang et al. (2023).

4.2 Spatial Patterns of Inferred Flood Tail Behavior

Figure 3 displays the spatial distribution of dominant flood tail behavior across seasons, based on the recession exponent values, for Germany, the UK, Norway, and the US, respectively. This dominant behavior represents either heavy or nonheavy tails, depending on what is observed in the majority of seasons. Additionally, Figure 4 and Table 1 provide quantitative analyses of the propensity of flood tail behavior across different regions.

In Germany (Figure 3a), approximately 81% of catchments are identified as sites with dominant heavy-tailed flood behavior (red dots), indicating a prevalence of such behavior. This result agrees with the findings of Mushtaq et al. (2022), which reported that a distribution with a relatively heavier tail (i.e., the log-normal) best represent ordinary peak flows in the majority of German basins considered in their study. The inferred heavy-tailed sites are spread across Germany. They dominate in the eastern part, while there are mixed patterns of heavy- and nonheavy-tailed behavior in the western part. The climate conditions are primarily humid continental (Dfb) and temperate oceanic (Cfb). Humid continental climate is prominent in the east, while temperate oceanic climate generally covers the west.

In the UK (Figure 3b), four climate types are present, with temperate oceanic climate (Cfb) being the dominant one. The terrain of this country in comparison to the other three countries is relatively homogeneous, with no high mountains. According to our findings, heavy-tailed flood behavior is prevalent in the UK, with a prevalence of 77%, especially in the eastern and southern coastal regions. Huntingford et al. (2014) reported a case in which a rapid succession of vigorous Atlantic low-pressure systems crossed much of the UK, resulting in repeated heavy rainfall events. Southeast England was identified as a distinct region characterized by exceptionally high flows, exacerbated by increasingly saturated catchments. These catchment characteristics and hydrological responses align with our findings, which indicate the pronounced heavy tails in such a region.

In Norway (Figure 3c), however, nonheavy-tailed flood behavior dominates. Approximately 89% of sites are inferred to have nonheavy-tailed flood behavior. Norway encompasses nine climate types but is primarily covered by Subarctic climate (Dfc), characterized by low temperatures and reduced evapotranspiration. Hydrological processes are significantly influenced by snow dynamics, which generally determine linear catchment responses as a result of snow accumulation and melting processes (Santos et al., 2018).

In contrast to the aforementioned countries with relatively consistent climate and dominant flood behavior, the US (Figure 3d) display a diverse range of climate types and a balanced propensity toward heavy- and nonheavy-tailed flood behavior. The eastern regions dominated by humid subtropical climate (Cfa), hot-summer humid continental climate (Dfa), and temperate oceanic climate (Cfb) from south to north. The interior western states feature a cold semi-arid climate (BSk), while mixed patterns are observed in the western mountainous and coastal areas. An overall relatively even distribution of inferred heavy-tailed (52%) and nonheavy-tailed (48%) flood behavior prevails in this diverse climate country.

Figure 3e provides an example of how the spatial distribution of flood behavior is influenced by regional physioclimatic features. In particular, catchments on the east side of the mountains exhibit pronounced heavy-tailed flood behavior, which aligns with the findings of Smith et al. (2018). This is likely due to the interaction between cold air from the inland polar jet stream and warm ocean currents leads to the formation of Nor'easters, which are synoptic-scale extratropical cyclones in the western North Atlantic Ocean along the US northeast coast. These weather systems often resulted in heavy rain or rain-on-snow events. Conversely, on the west side of mountains, catchments tend to exhibit nonheavy-tailed behavior, potentially due to the leeward rain shadow effect.

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

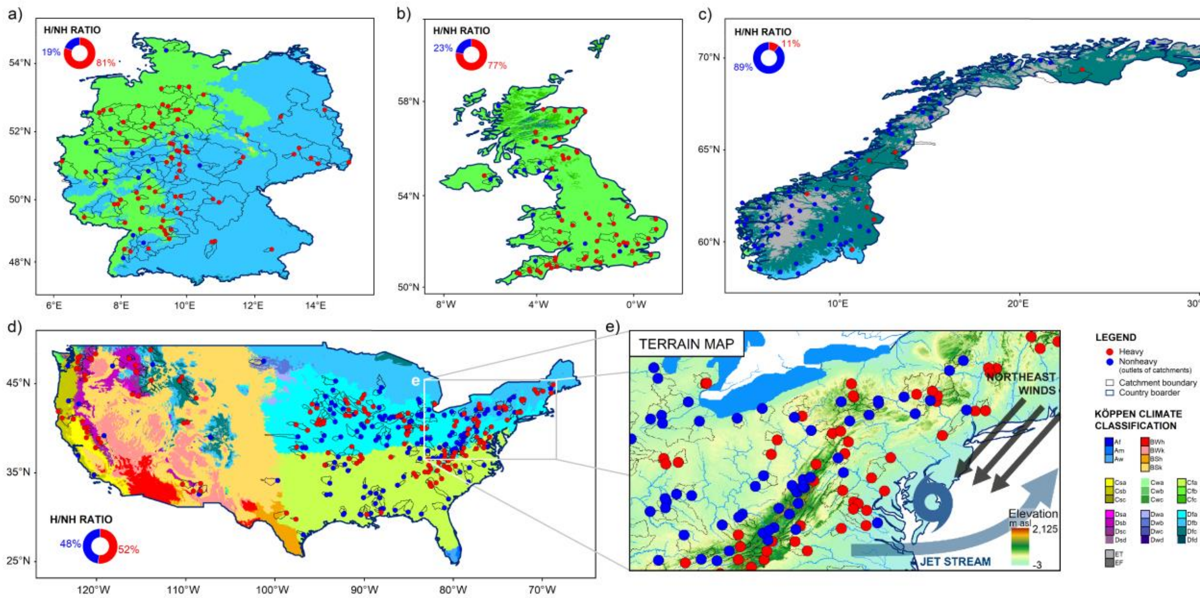


Figure 3. Spatial distribution of dominant flood behavior. The dominant pattern determines the representative flood tail behavior of catchments across all study countries, whether it is heavy or non-heavy, which is defined by the major pattern recognized across seasons. Tail behavior is inferred by hydrograph recession exponents. The ratio of heavy- to nonheavy-tailed catchments is indicated as H/NH ratio. Köppen climate classification based on Beck et al. (2018), with present climate types outlined by bold dark frames in the legend. (a) Germany, a total of 98 gauges represent catchments ranging from 110 to 23,843 km^2, with a median area of 1,195 km^2. (b) The UK, a total of 82 gauges represent catchments ranging from 15 to 9,948 km^2, with a median area of 283 km^2. (c) Norway, a total of 82 gauges represent catchments ranging from 4 to 40,504 km^2, with a median area of 234 km^2. (Note: some catchment boundaries are absent in the dataset for catchments in the UK and Norway.) (d) The US, a total of 313 gauges represent catchments ranging from 66 to 9,935 km^2, with a median area of 1769 km^2. (e) A zoomed-in map illustrates the discernible patterns of flood tail behavior resulting from specific flood generation processes influenced by the interplay between regional terrain and meteorological features. (Note: the cylinder map projection is employed in these maps.)

To obtain quantitative results we examine the predominant flood tail behavior (inferred by recession exponents) of catchments across various climate regions and sort these regions based on the proportion of heavy-tailed catchments from high to low, as illustrated in Figure 4. By categorizing climate type regions based on the proportion of heavy-tailed catchments, we establish three groups according to their propensity of flood tail behavior: Heavy-tailed group, indicating regions with over 66.6% of catchments dominated by heavy tails; Neutral group, encompassing regions with 33.3% to 66.6% of catchments dominated by heavy tails, represents a relatively even propensity for both heavy and nonheavy tails in the catchments within these regions; and Nonheavy-tailed group, representing regions with less than 33.3% of catchments dominated by heavy tails, denotes the propensity for nonheavy tails. According to the Köppen climate type classification, the overarching hydroclimatic characteristics can be delineated by three hierarchical features: 1. the main group, which encompasses five areas—Tropical, Arid, Temperate, Continental, and Polar; 2. precipitation characteristics; and 3. temperature characteristics. The findings are synthesized in Figure 4 and Table 1, where the groups of flood tail behavior propensity are juxtaposed with the distinctive traits of each climate region.

Five climate regions are identified as having a higher propensity for heavy tails: mediterranean climate (Csa), hot semi-arid climates (BSh), humid continental climate (Dsb), temperate oceanic climate (Cfb), and cool-summer mediterranean climate (Csb). These regions are characterized by warm to hot temperatures, often accompanied by occasional dry periods (except for Cfb). Based on the definition of Köppen climate classification the occurrence of dry periods is a result of significantly uneven rainfall throughout the year, with at least three times as much rainfall in the wettest month as in the driest month. In semi-arid climates (BSh), there is generally lower annual rainfall (summarized in Table 1). Higher temperatures increase the potentail evapotranspiration, often enhancing atmospheric moisture content and facilitating convective rainfall. Moreover, the dynamics of evapotranspiration in hillslopes influence the nonlinearity of runoff processes in catchments (Tashie et al., 2019). Dry periods can lead to lower catchment soil moisture, facilitating nonlinear runoff generation (Merz and Blöschl, 2009; Viglione et al., 2009). The findings presented here indicate that heavy-tailed flood behavior tends to emerge due to the substantial nonlinearity observed in catchment hydrological processes, which is facilitated by temporally uneven rainfall and higher evapotranspiration variation throughout the year.

We also find that certain regions show a relatively neutral propensity regarding flood tail behavior (either heavy- or nonheavy-tailed) and aggregate them into the second group of Figure 4 and Table 1. These regions encompass cold semi-arid climates (BSk), humid continental climate (Dfb), humid subtropical climate (Cfa), and humid continental climate (Dfa). While cold semi-arid climates (BSk) experience dryness, they are characterized by very limit precipitation. In the other three regions (Dfb, Cfa, and Dfa), heavy tails may still occur due to higher evapotranspiration, which is driven by high temperatures. However, the relatively even distribution of rainfall throughout the year in these regions may reduce the propensity for heavy tails, resulting in a smoother occurrence of heavy-tailed flood behavior. In summary, the regions in this group still have a certain probability of exhibiting heavy-tailed flood behavior. However, the absence of either a drier state of the catchment (caused by uneven rainfall) or higher temperatures (that ensure sufficient atmospheric moisture for rainfall and strengthened nonlinearity) could constrain the occurrence of such behavior.

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

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

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

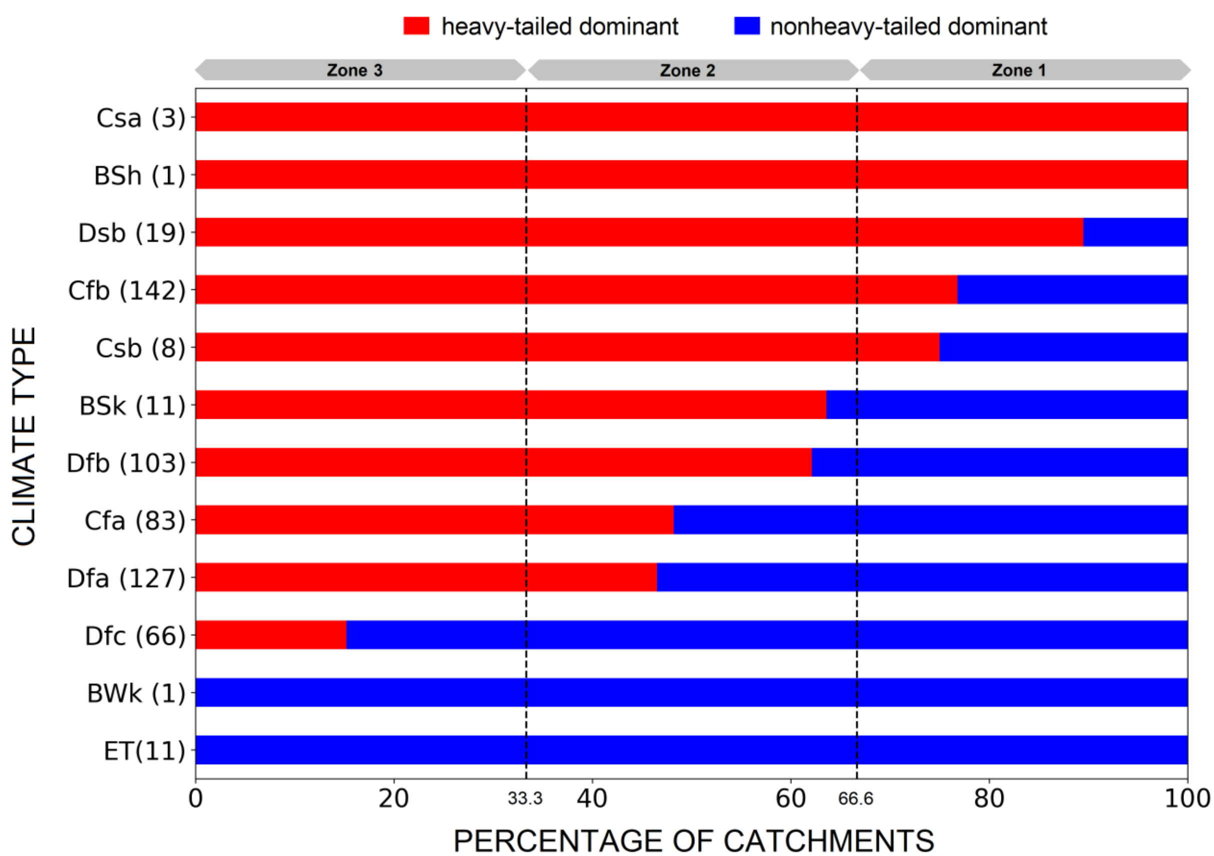


Figure 4. Propensity of inferred flood tail behavior in diverse climate regions. Catchments are categorized by climate types and grouped by dominant (across seasons) heavy-tailed case percentages. Three groups are defined by heavy-tailed case proportions: Zone 1 (>66.6%) represents heavy tails, Zone 2 (33.3-66.6%) is neutral, and Zone 3 (<33.3%) represents nonheavy

tails. The number of catchments in each climate region is indicated in parentheses after the climate type.

Table 1. Comparison of inferred flood tail behavior propensity with climate characteristics.

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

4.3 Seasonal patterns of Inferred Flood Tail Behavior

We analyze the seasonality of flood tail behavior, an aspect of this phenomenon which has been previously suggested but remains poorly understood (Durrans et al., 2003; Basso et al., 2015; Macdonald et al., 2022). Figure 5 illustrates the spatial distribution of catchments with consistent tail behavior across seasons (i.e., with either heavy or nonheavy tails across all seasons; black triangles) and those with varying tail behavior across seasons (green dots). Catchments exhibiting inconsistent behavior spread across the whole US and Germany, whereas they are mostly concentrated in the southern parts of the UK and in the central mountainous regions of Norway. The percentages of catchments exhibiting inconsistent flood tail behaviors are respectively 33%, 33%, 17%, and 34% in the US, Germany, the UK, and Norway. The results indicate that although the majority of catchments tend to exhibit stable heavy-/nonheavy-tailed behavior, still around one-third reveal changing patterns across seasons. Notably, there is a particularly high percentage of consistent patterns (83%) in the UK, likely due to the relatively uniform climate and terrain conditions across the country characterized by continuous rainfall throughout a year (as shown in Figure 3b).

We further investigate the dynamics of heavy- and nonheavy-tailed case studies across seasons in Figure 6. Heavy-tailed case studies increase from spring to autumn (approximately corresponding to the growing season in the northern hemisphere) and decrease from autumn to spring (approximately corresponding to the dormant season in the northern hemisphere), as seen in the aggregated patterns across all regions (panel a). This pattern can be attributed to the increasing temperature in the growing season, during which increasing evapotranspiration consumes water storage in the shallow subsurface, escalating the nonlinearity of catchment responses (Tashie et al., 2019). The seasonality of evapotranspiration effects on catchment nonlinearity is supported by the findings of Tarasova et al. (2018), who observed clear seasonal dynamics of catchment average runoff coefficients. These coefficients tend to be higher in wet winters and lower in dry summers. It has been shown that significant variation in runoff coefficients is linked to high nonlinearity of hydrological responses, facilitating heavier-tailed floods. This phenomenon is often observed in dry catchments (Merz and Blöschl, 2009). Other studies confirmed that the nonlinearity of catchment responses favors the emergence of heavy-tailed flood behavior (Gioia et al., 2008; Rogger et al., 2012; Basso et al. 2015), and is often expressed by quicker recession during high flow periods and greater stability during low flow periods. Conversely, during the dormant season, nonlinearity decreases due to reduced competition from evapotranspiration and replenished water storage. We underscore that the significant variability in evapotranspiration amplifies the fluctuation of catchment storage conditions, causing soil moisture levels to oscillate between drier and wetter states. This alternation leads to the occurrence of both very small and very large events, which are characteristic of heavy-tailed flood behavior.

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

We delve into the seasonal characteristics of this behavior further by combining the regional patterns based on climate classification. In Figure 7, the square dots represent the median of each box, marked as red, blue, or black to indicate the significance of its value above 2 (heavy), below 2 (nonheavy), or not significantly different from 2, respectively. The last one (black squares) may imply an equal occurrence of heavy- and nonheavy-tailed cases or a lack of samples to draw conclusions. Based on the patterns of significance across seasons, regions with seasonality (defined as having different tail behavior propensity across seasons according to the significant values of median) are grouped in the white area, while those considered stable in heavy tails are in the red area, and those stable in nonheavy tails are in the blue area. For regions where statistical significance cannot be concluded for all seasons, we group them based on the absolute values of their medians.

We find that the grouping based on their distinct patterns of seasonality (Figure 7) closely aligns with the grouping based on the analysis of dominant patterns throughout the year (Figure 4 and Table 1). Regions (red area in Figure 7 corresponded to the heavy-tailed group in Table 1) characterized by uneven rainfall distribution throughout the year, leading to pronounced fluctuations between drier and wetter soil states, combined with higher evapotranspiration rates (indicated by warm to hot temperatures), tend to exhibit a dominance of heavy-tailed behavior in their hydrological responses across all seasons. In areas (white area in Figure 7 corresponded to the neutral group in Table 1) where rainfall is more evenly distributed annually, the emergence of heavy-tailed behavior is often linked to increased evapotranspiration during the growing seasons, particularly in spring and summer, and is less prominent during dormant seasons. This mechanism, which depends on evapotranspiration dynamics, substantiates the seasonality of flood tail behavior. Regions (blue area in Figure 7 corresponded to the nonheavy-tailed group in Table 1) where runoff generation is primarily influenced by snow dynamics tend to display linear hydrological responses. This is due to the fact that most runoff in these areas results from snowmelt during the growing seasons, driven by energy availability. These findings support the proposed mechanism of heavy-tailed flood behavior concluded in the spatial analyses and further demonstrate the pivotal effect played by the variation of evapotranspiration and catchment storage on the emergence of heavy-tailed flood behavior.

In summary, while heavy-/nonheavy-tailed behavior is generally consistent across seasons, there is a certain probability for cases to exhibit seasonality. This seasonality of inferred heavy-tailed behavior shows a dynamic pattern of increasing during the growing season and decreasing during the dormant season. Regions with pronounced temperature variations across seasons, particularly with hot summers, display such dynamics and highlight the role of evapotranspiration in catchments in driving this seasonality.

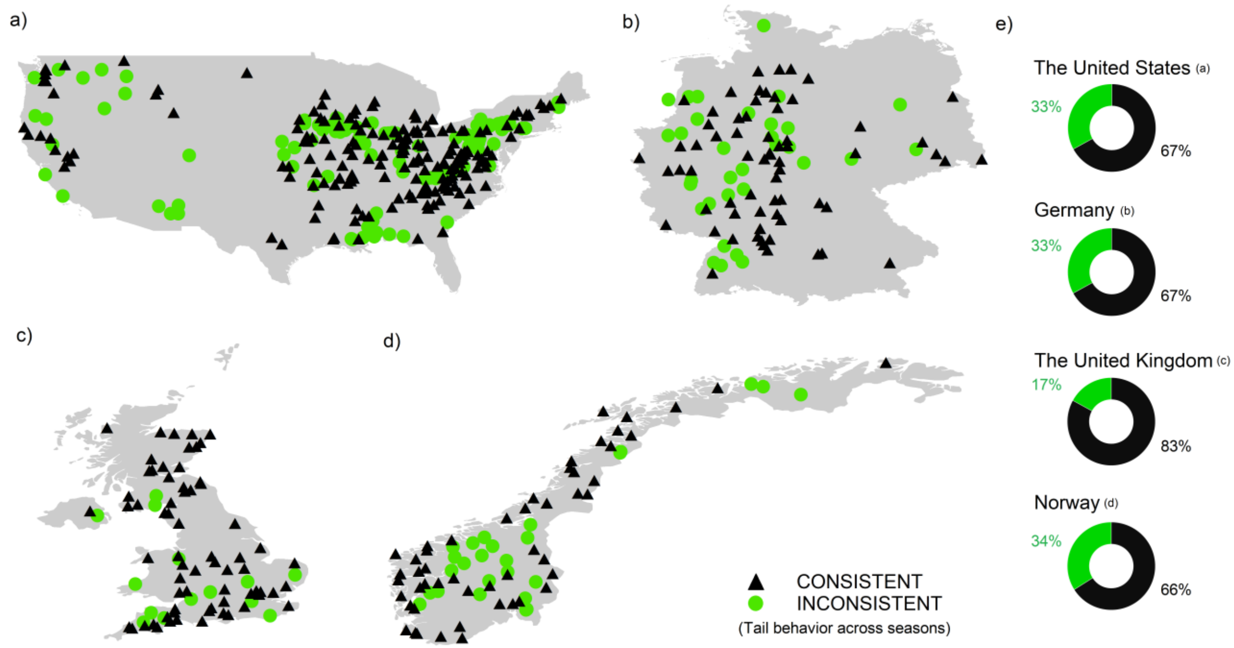


Figure 5. Consistency of inferred flood tail behavior across seasons. (a) 290 catchments in the US. (b) 98 catchments in Germany. (c) 81 catchments in the UK. (d) 79 catchments in Norway. (e) Percentage of consistent and inconsistent catchments in each country.

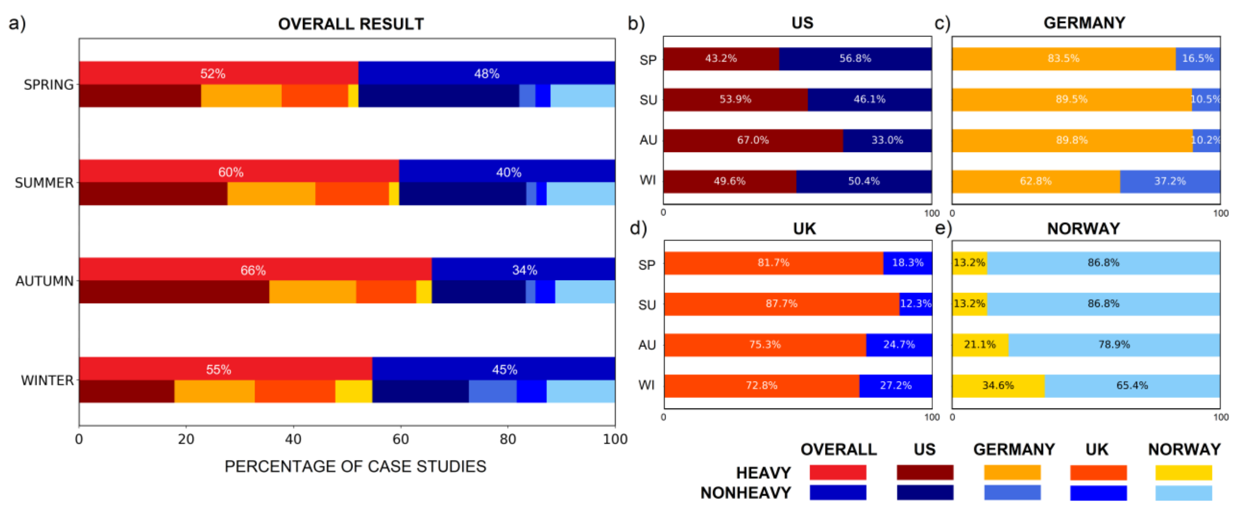
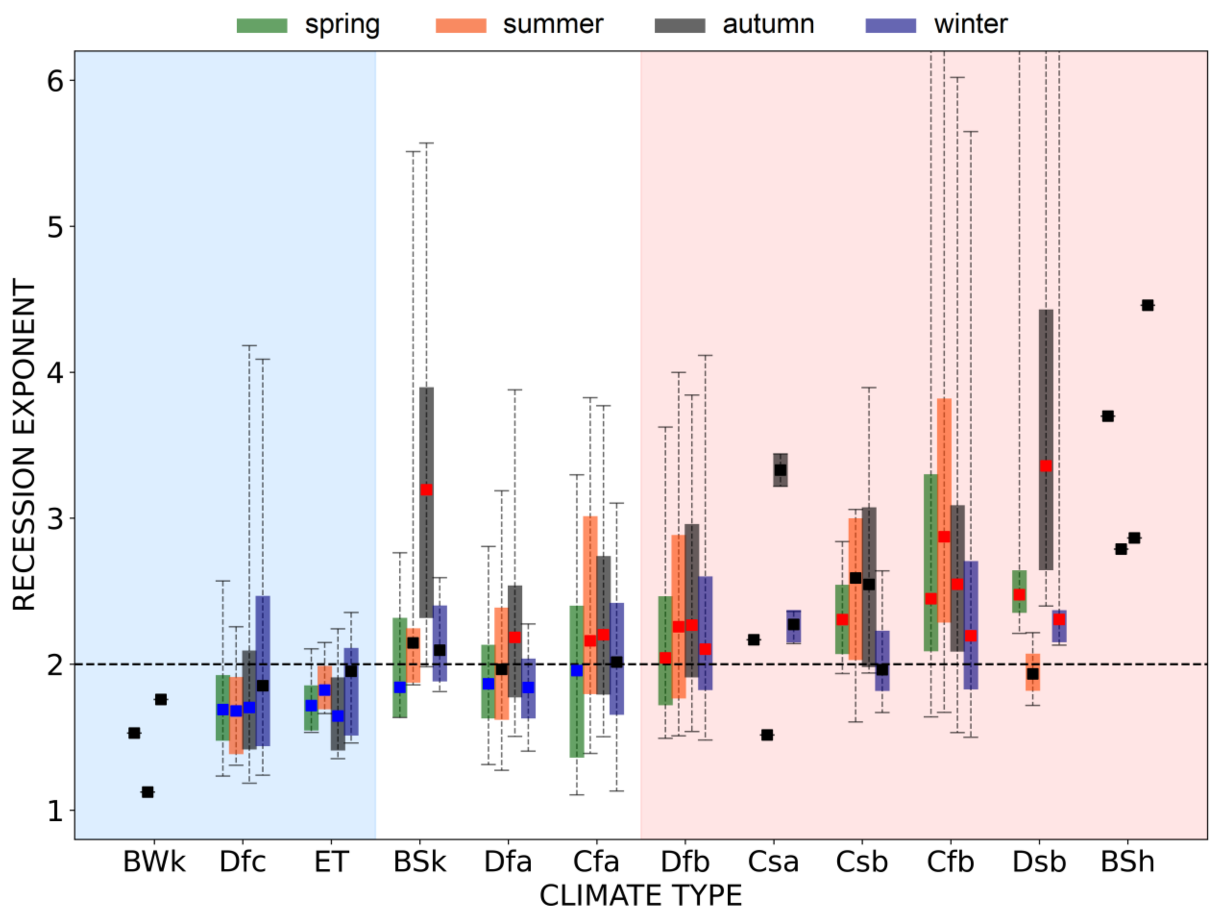


Figure 6. Seasonal variations in the percentage of inferred flood tail behavior between heavy and nonheavy case studies. (a) The aggregated results encompass all study regions, while the second line provides a breakdown by country. In total, there are 1,997 case studies composed by

581 540 in spring, 520 in summer, 543 in autumn, and 394 in winter. (b)-(e) Results for each study
582 country (see Table A1 for detailed case numbers across seasons in each country).



583

584 **Figure 7. Seasonal variations in recession exponents (inferring flood tail behavior) across**
585 **diverse climate regions.** Case studies grouped by climate regions based on seasons. Medians of
586 recession exponents in each group are compared with a value of two using Wilcoxon signed-rank
587 test (significance level: 0.05). Red squares indicate significantly heavy-tailed (recession exponents
588 > 2) groups, blue squares indicate significantly nonheavy-tailed (recession exponents < 2) groups,
589 and black squares denote insignificance. Climate regions are categorized as follows: the red area
590 denotes regions with prominent heavy tails across seasons, the blue area denotes regions with
591 prominent nonheavy tails across seasons, and the white area denotes regions with significant
592 seasonality in flood tail behavior.

593 4.4 Factors associated with catchment scales and their role in flood tail behavior

594 It remains unclear how flood tail behavior varies across catchment scales and what the underlying
595 drivers and mechanisms are (Merz et al., 2022). We employ catchment nonlinearity, represented
596 by recession exponents, to explore the influence of catchment scales on flood tail behavior, as
597 depicted in Figure 8. We utilize the categorization of regions characterized by distinct controls on

flood tail behavior, primarily influenced by characteristic runoff generation processes (as three groups identified in Figure 7), to elucidate the underlying mechanisms. Case studies are categorized into bins based on catchment areas, with the median values represented by squares, interquartile intervals depicted by vertical bars, and catchment area ranges indicated by horizontal dashed bars. Panels a, b, c, and d present results for all regions, regions exhibiting significant heavy tails across seasons, regions with a neutral propensity and seasonal variations, and regions displaying pronounced nonheavy tails across seasons, respectively. Each panel comprises a total of 30 bins, with approximately 67, 33, 24, and 10 case studies in panels a, b, c, and d, respectively (with minor variations due to rounding).

From the perspective of all case studies (Figure 8a), the pattern appears somewhat unclear. Apart from the case studies involving extremely small and large catchment areas, there seems to be a decrease in nonlinearity as catchment areas increase. Nevertheless, the relationship is rather weak and lacks clarity. These findings align with previous discussions on this matter (e.g., Merz and Blöschl, 2009; Villarini and Smith, 2010; Smith et al., 2018), which have suggested a relatively weak inverse correlation between catchment area and the occurrence of heavy-tailed flood behavior.

However, we can evidently clarify this relationship by considering the distinct runoff generation processes recognized in different regions. Panel b illustrates that catchment area plays no significant role in catchment nonlinearity in regions characterized by prominent heavy tails. Whereas a clear inverse relationship between nonlinearity and catchment area is shown in panel c, representing regions characterized by a neutral propensity for heavy and nonheavy tails. In contrast, a proportional relationship between nonlinearity and catchment area is identified in panel d, representing regions characterized by prominent nonheavy tails.

As shown by the previous sections, nonlinearity in neutral regions is primarily driven by high evapotranspiration facilitated by high temperatures. When the catchment area increases, it has a higher chance of encompassing diverse terrain types, including areas with higher altitudes, such as mountainous regions. Increased altitude tends to result in lower temperatures and evapotranspiration rates, negating the evapotranspiration variation and its impact on catchment nonlinearity, which is the main driver of nonlinearity in this region and thus substantiates an inverse relationship (Figure 8c). In regions with prominent heavy tails (Figure 8b), nonlinearity is generated from the interplay of uneven rainfall and evapotranspiration dynamics, and the enlargement of catchments does not substantively change this relationship. For regions with prominent nonheavy tails (Figure 8d), the underlying mechanisms are similar to the neutral regions but work in the opposite direction due to the differently dominant mechanism. Recall that the runoff process in this region is generally dominated by snow dynamics. The region is mainly located in high mountain or high latitude areas. As catchments expand, more diverse terrain is encompassed, potentially introducing a mixture of flood generation processes due to the incorporation of lowland or coastal areas. Particularly, more rain-on-snow events or rainfall-driven events may be encompassed in a same catchment together with snowmelt-driven events (Vormoor et al., 2016). Therefore, an increase in nonlinearity is facilitated due to the mixture of flood types, favoring the emergence of heavier tails in flood distributions (Tarasova et al., 2020). It should be noted that the tail patterns, based on Figure 8d, are still more likely to be nonheavy tails (i.e., recession exponents below two), even though nonlinearity indeed appears to show an increasing tendency along with catchment area.

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

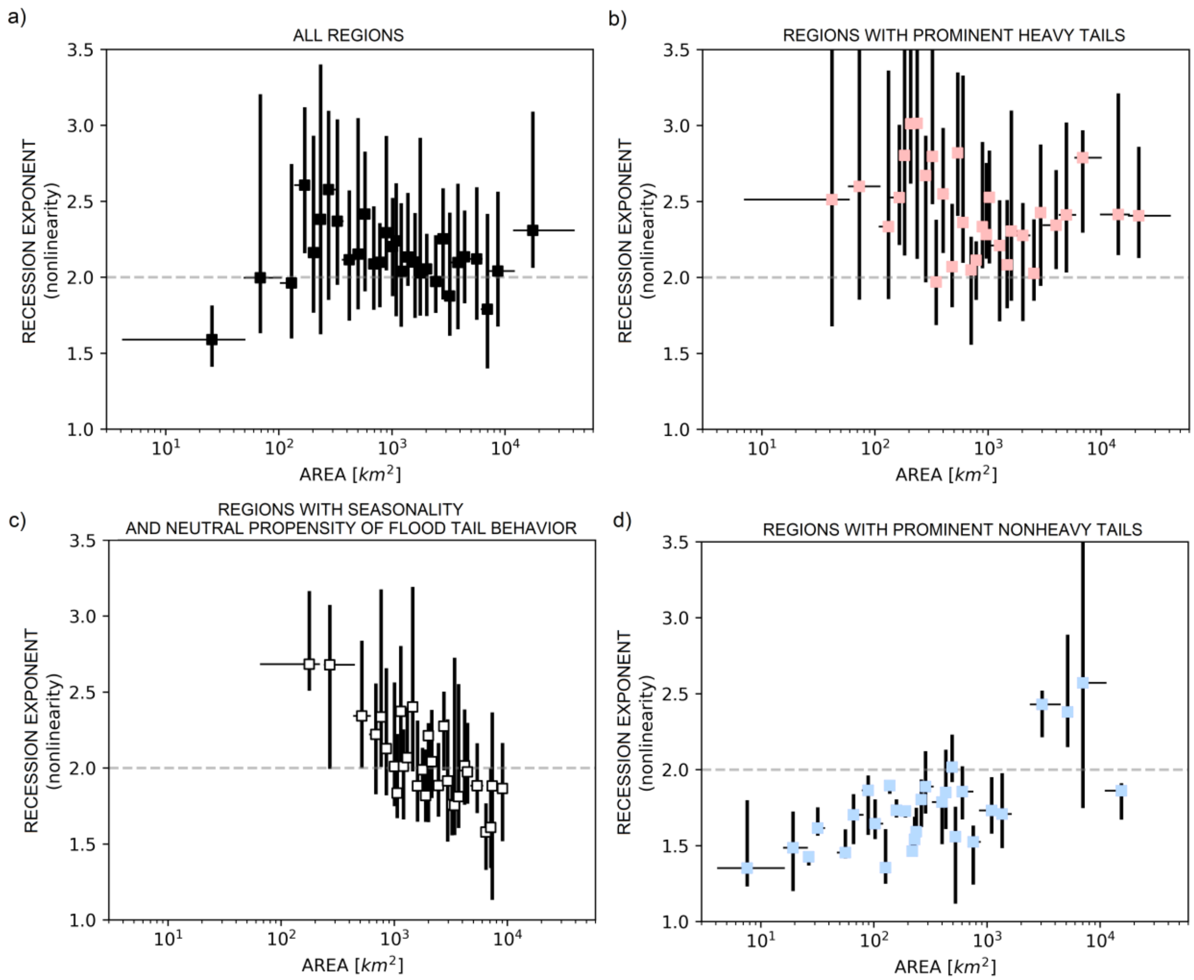


Figure 8. Catchment nonlinearity as a function of catchment area. The recession exponents, representing catchment nonlinearity, have been evenly grouped into bins based on catchment areas. The squares denote the median values, vertical bars represent the interquartile intervals of the recession exponents, and horizontal dashed bars indicate the catchment area ranges for each bin. (a) All regions (encompassing case studies, *n*=1997). (b)-(d) show case studies separately according to categorization recognized in Figure 7. (b) Regions with prominent heavy tails (*n*=978). (c) Regions with seasonality and neutral propensity of flood tail behavior (*n*=733). (d) Regions with prominent nonheavy tails (*n*=286). In each panel, there are a total of 30 bins, each containing

approximately 67, 33, 24, and 10 case studies in panels a, b, c, and d, respectively (with slight variations due to rounding).

5 Discussion

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

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

We further identify key drivers of heavy-tailed flood behavior by conducting large scale physioclimatic analyses. Specifically, our findings reveal that regions with a pronounced propensity for heavy tails exhibit distinct characteristics: the presence of a dry period and higher temperatures (as shown in Figure 4 and Table 1). This aligns with previous studies based on the mathematical analysis which associates heavier-tailed flood behavior with a lower frequency of streamflow-triggering rainfall events. Such lower frequency often results from erratic rainfall patterns and higher rates of evapotranspiration, leading to drier catchment conditions (Botter, 2010; Basso et al., 2016). In line with this theory, our large scale analysis provides evidence by showing a prevalent propensity for heavy tails in regions characterized by uneven rainfall patterns throughout the year (i.e., more erratic rainfall), contributing to the presence of dry periods, along with higher potential evapotranspiration rates, as indicated by higher temperatures.

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

The seasonality of flood tail behavior has been suggested in previous studies but remains less understood (Basso et al., 2015; Smith et al., 2018; Macdonald et al., 2022). It's noteworthy that more than one-third of catchments appear to exhibit inconsistent flood tail behavior across seasons (Figure 5). In these catchments, some seasons show a tendency toward nonheavy tails, while others tend to display heavy tails. Identifying these catchments and understanding the factors driving them to exhibit heavy tails is vital for hazard assessment. This understanding allows us to pinpoint catchments where extreme floods could potentially occur, even if methods solely based on annual maximum floods might estimate the flood tail as nonheavy based on annual maxima, when heavy tails can still occur within a single season. We have identified that regions characterized by stronger evapotranspiration dynamics across seasons favor this seasonality of flood tail behavior, as it leads to larger variations in water storage during higher evapotranspiration seasons, such as the growing seasons (highlighted in white in Figure 7). This finding aligns with previous studies that have observed similar seasonal dynamics in the nonlinearity of hydrological responses (Tashie et al., 2019; Tarasova et al., 2018).

In this study, we also found that the relationship between flood tail behavior and the expansion of catchment scales can be explained by changes in catchment nonlinearity, which are influenced by distinct flood generation processes. Previous studies have presented diverse perspectives on the relationship between flood tail behavior and catchment scales. While some studies have suggested that smaller catchments tend to exhibit heavier tails (e.g., Meigh et al., 1997; Pallard et al., 2009), others have noted a similar trend but with only a weak correlation (Merz and Blöschl, 2009; Villarini and Smith, 2010). Meanwhile, some studies have found no significant relationship between these two variables (Morrison and Smith, 2002; Smith et al., 2018). These studies have explored this topic without reaching a consensus, and many conclusions lack sufficient evidence and a clear understanding. In contrast, our findings (Figure 8) distinctly differentiate between various patterns by considering region classifications based on distinct dominant flood generation processes, thereby providing a mechanistic understanding. As a catchment expands, it encompasses more diverse terrain, which in turn facilitates a wider range of altitudes and flood types. In regions where tail behavior is primarily influenced by evapotranspiration dynamics (Figure 8c), the presence of diverse altitudes tends to moderate the effect of higher temperatures, reducing the influence of high evapotranspiration on the emergence of heavy tails. In regions where tail behavior is primarily controlled by snowmelt (Figure 8d) (mainly composed of catchments in Norway in this study), it has been shown that larger catchments are more likely to encompass a mix of flood types, including snowmelt-driven and rainfall-driven floods (Vormoor et al., 2016). Merz et al. (2022) suggested that heavier-tailed behavior in rainfall-driven floods tends to dominate in such mixed conditions. Our findings support this hypothesis by demonstrating an increase in tail heaviness as catchment area enlarges. In regions where heavy tails are pronounced due to the strong nonlinearity resulting from the interplay of uneven rainfall and high evapotranspiration, there is no significant relationship between catchment nonlinearity and catchment area (Figure 8b). This lack of relationship may be because the expansion of the catchment area does not appear to significantly enhance or reduce this interplay.

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

"Hypothesis 2 (of the review paper): The Characteristic Flood Generation Process Shapes the Upper Flood Tail Catchments." While previous studies have hinted at the possibility that regions where flood generation is dominated by rainfall-driven floods tend to exhibit heavier-tailed flood behavior compared to regions dominated by snowmelt (Bernardara et al., 2008; Thorarinsdottir et al., 2018), more explicit process explanations are desired. In line with this hypothesis, we present further evidence showing that the specific nonlinearity inherent in each flood generation process is the primary driver of flood tail behavior. Specifically, we show that in snowmelt-dominated regions, such as the case studies in Norway, hydrological responses closely resemble linear behavior and thus floods tend to exhibit pronounced nonheavy-tailed behavior. Conversely, heavy-tailed floods are more prominent in regions like the UK, where hydrological responses display nonlinearity (as indicated by recession exponents above two). In these areas, flood generation processes are primarily driven by rainfall events. Furthermore, our study reveals that flood generation processes are significantly influenced by the interplay between regional terrain and meteorological features. These factors, in turn, impact the nonlinearity of hydrological responses and can determine the occurrence of heavy or nonheavy tails in flood distributions (Figure 3e).

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

"Hypothesis 4: Non-Linear Response to Precipitation Causes Heavy Flood Tails." Studies have consistently highlighted the significance of nonlinearity in hydrological processes within catchments as a key determinant in the emergence of heavy-tailed flood behavior (e.g., Struthers and Sivapalan, 2007; Rogger et al., 2012; Basso et al., 2015). In our research, we contribute by introducing a quantitative approach that employs hydrograph recession exponents as a measure of nonlinearity in flood tail analyses and validate its effectiveness in identifying heavy-tailed flood behavior in a large scale analysis. While nonlinearity has long been acknowledged as a contributing factor, our works uniquely utilizes this driver as a reliable index by establishing a specific recession exponent threshold that robustly discriminates heavy-tailed distributions, characterized by power-law tails, from nonheavy ones, offering a valuable tool to the field.

Furthermore, our large scale analysis identifies rainfall unevenness and high temperatures as crucial drivers behind the observed nonlinearity in flood responses. Specifically, they intensify catchment soil dryness and amplify water balance storage variations, thereby facilitating both very small and very large runoff events, translating into heavy-tailed flood behavior.

"Hypothesis 5: Drier Catchments Have Heavier Flood Tails Due To Interaction of Water Balance Processes." In alignment with previous studies that suggest the water balance processes in drier catchments contribute to the emergence of heavy-tailed flood behavior (e.g., Molnar et al., 2006; Merz and Blöschl, 2009; Guo et al., 2014), we emphasize the critical interplay between uneven rainfall and evapotranspiration dynamics in facilitating these processes and shaping such the behavior. Specifically, our findings show that heavy-tailed flood behavior is more likely to occur in catchments characterized by lower rainfall and higher evapotranspiration in one season (e.g., summer), contrasted with more rainfall and lower evapotranspiration in another season (e.g., winter). When one of these conditions is lacking, heavy-tailed behavior may be less pronounced. For example, regions classified as BSh and BSk, both of which exhibit semi-arid characteristics based on their rainfall patterns, exhibit differences in the prevalence of heavy-tailed cases. BSk regions, despite their semi-arid status, exhibit fewer pronounced heavy-tailed cases due to colder temperatures (Table 1) and only show a higher rate of heavy-tailed cases during the summer (Figure 7). This interplay highlights the importance of considering the seasonality of flood tail behavior, particularly in regions that do not experience significant dry periods based on their rainfall patterns. In such regions, heavy tails are still likely to occur in seasons with higher evapotranspiration rates (indicated by the white area in Figure 7).

"Hypothesis 6: Smaller Catchments Have Heavier Flood Tails Due To Less Pronounced Spatial Aggregation Effects." A commonly debated question among hydrologists is whether the roles identified in large catchments are applicable to smaller ones, and vice versa. This issue has also arisen in discussions regarding flood tail heaviness, but evidence on the matter has been scattered. While smaller catchments have been suggested to exhibit heavier tails (Meigh et al., 1997; Pallard et al., 2009), previous research has revealed weak (Merz and Blöschl, 2009; Villarini and Smith, 2010) to no (Morrison and Smith, 2002; Smith et al., 2018) correlations between catchment size and tail heaviness. Our findings (Figure 8) help clarify the relationship between catchment nonlinearity (used as an indicator of tail heaviness) and catchment sizes. We observe distinct patterns among regions characterized by strong, neutral, and weak conditions of heavy tail behavior. These findings underscore the importance of considering the dominant flood generation processes in each region and elucidate how catchment size interacts with flood tail behavior by influencing these dominant processes—either amplifying, reducing, or having no significant effect.

6 Conclusions

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

1. **Capability of Recession Exponents for Detecting Heavy-Tailed Flood Behavior:** The capability of this index to discern between case studies which display heavy-tailed flood

distributions and those exhibit nonheavy-tailed behavior is validated by using empirical data from catchments across Germany, Norway, the UK, and the US. This extensive analysis provides a well-rounded evaluation due to the inclusion of regions with divergent conditions, such as rainfall-driven floods (Germany, the UK, and the US) versus snowmelt-driven floods (Norway), as well as regions characterized by single/dominant hydroclimates (Germany, the UK, and Norway) versus those with mixed hydroclimates (the US).

2. **Regional Propensity for Heavy-Tailed Flood Behavior:** Germany and the UK are characterized by a propensity for heavy-tailed flood behavior, which is prevalent in these regions. Conversely, a tendency for nonheavy-tailed flood behavior is predominant in Norway under current hydroclimatic conditions, as indicated by the degree of catchment nonlinearity in each region. The US exhibits a mixture of heavy- and nonheavy-tailed behavior. This is likely the results of overarching climatic characteristics, which also shape river network morphology, interacting with diverse regional physioclimatic settings. We emphasize that the relatively more uniform climates in Germany, the UK, and Norway contribute to a dominant presence of heavy or nonheavy-tailed behaviorsin these countries, while the US experiences more complex regional patterns due to more diverse hydroclimatic conditions.
3. **Factors Influencing Heavy-Tailed Flood Behavior:** The presence of simultaneous dry periods (defined by uneven rainfall throughout the year) and higher temperatures emerge as the pivotal conditions favoring heavy-tailed flood behavior. Drier catchments alter the runoff generation process, resulting in higher nonlinearity of catchment responses, while higher temperatures elevate evapotranspiration rates, enhancing nonlinearity but also maintaining atmospheric moisture preventing precipitation limitations. The absence of either condition diminishes the prevalence of heavy-tailed flood behavior. More generalized climate categorizations like Arid, Temperate, and Continental exhibit minimal influence on our results.
4. **Seasonality of Flood Tail Behavior:** We contribute to a better understanding of the seasonality of flood tail behavior. Around two-thirds of catchments exhibit consistent behavior across seasons, with the remaining one-third demonstrating seasonality. Heavy-tailed flood behavior is more likely during the growing season (spring to autumn) and diminishes during the dormant season (autumn to winter). These findings hint at the role of temperature-driven evapotranspiration dynamics for the emergence of heavy-tailed flood behavior, which are particularly important in regions which do not experience simultaneous dry conditions and high temperatures.
5. **Influences of Catchment Area on Flood Tail Behavior:** We elucidate that the impacts of catchment size on flood tail behavior are primarily contingent on the dominant flood generation processes within each region. Specifically, the expansion of catchment area tends to have three distinct effects: (1) It diminishes tail heaviness in regions with moderate nonlinearity, characterized by strong evapotranspiration dynamics and relatively even rainfall throughout the year. This reduction is attributed to the smoothing of evapotranspiration variations. (2) Conversely, in regions with low nonlinearity, characterized by snowfall dynamics, increasing catchment area intensifies tail heaviness. This effect results from the inclusion of diverse flood types, particularly rainfall-driven floods. (3) In regions with high nonlinearity, characterized by simultaneous strong evapotranspiration dynamics and uneven rainfall throughout the year, catchment size expansion appears to have no significant impact on tail heaviness. This lack of effect is likely due to the absence of significant influence on rainfall patterns, which are critical in determining the presence of drier soil conditions.

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

Appendix A Information on Study Regions

Table A1. Daily Hydrological Data Information

Region	Germany	UK	Norway	US
Gauge Number	98	82	82	313
Catchment Size [*km²*]	110 – 23,843 (median: 1,195)	15 – 9,948 (median: 283)	4 – 40,504 (median: 234)	66 – 9,935 (median: 1,769)
Streamflow Record Length [*year*]	35 – 63 (median: 58)	50 – 138 (median: 59)	50 – 148 (median: 96)	24 – 55 (median: 55)
Streamflow Record Duration	1951 – 2013	1883 – 2021	1871 – 2019	1948 – 2002
Number of Case Study (spring / summer / autumn/ winter)	386 (97 / 96 / 98 / 95)	325 (82 / 81 / 81 / 81)	306 (76 / 76 / 76 / 78)	980 (285 / 267 / 288 / 140)

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Project number 421396820 "Propensity of rivers to extreme floods: climate-landscape controls and early detection (PREDICTED)" and Research Group FOR 2416 "Space-Time Dynamics of Extreme Floods (SPATE)". We also acknowledge the financial support provided by the Helmholtz Centre for Environmental Research - UFZ. The manuscript and supporting information contain all the necessary details to replicate the results.

Data Availability Statement

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

References

Arai, R., Toyoda, Y., & Kazama, S. (2020). Runoff recession features in an analytical probabilistic streamflow model. Journal of Hydrology, 597, 125745. https://doi.org/10.1016/j.jhydrol.2020.125745

Basso, S., Botter, G., Merz, R., & Miniussi, A. (2021). PHEV! The PHysically-based Extreme Value distribution of river flows. Environmental Research Letters, 16(12). https://doi.org/10.1088/1748-9326/ac3d59

Basso, S., Schirmer, M., & Botter, G. (2015). On the emergence of heavy-tailed streamflow distributions. Advances in Water Resources, 82, 98–105. https://doi.org/10.1016/j.advwatres.2015.04.013

Basso, S., Schirmer, M., & Botter, G. (2016). A physically based analytical model of flood frequency curves. Geophysical Research Letters, 43(17), 9070–9076. https://doi.org/10.1002/2016GL069915

Basso, S., Merz, R., Tarasova, L., & Miniussi, A. (2023). Extreme flooding controlled by stream network organization and flow regime. Nature Geoscience, 16(April), 339–343. https://doi.org/10.1038/s41561-023-01155-w

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data, 5, 180214. https://doi.org/10.1038/sdata.2018.214

Bernardara, P., Schertzer, D., Sauquet, E., Tchiguirinskaia, I., & Lang, M. (2008). The flood probability distribution tail: How heavy is it? Stochastic Environmental Research and Risk Assessment, 22(1), 107–122. https://doi.org/10.1007/s00477-006-0101-2

Bevere, L., & Remondi, F. (2022). Natural catastrophes in 2021: the floodgates are open. Swiss Re Institute sigma research.

Biswal, B. (2021). Decorrelation is not dissociation: There is no means to entirely decouple the Brutsaert-Nieber parameters in streamflow recession analysis. Advances in Water Resources, 147, 103822. https://doi.org/https://doi.org/10.1016/j.advwatres.2020.103822

Biswal, B., & Marani, M. (2010). Geomorphological origin of recession curves. Geophysical Research Letters, 37(24), 1–5. https://doi.org/10.1029/2010GL045415

Botter, G. (2010). Stochastic recession rates and the probabilistic structure of stream flows. Water Resources Research, 46(12). https://doi.org/10.1029/2010WR009217

Botter, G., Basso, S., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). Natural streamflow regime alterations: Damming of the Piave river basin (Italy). Water Resources Research, 46(6), 1–14. https://doi.org/10.1029/2009WR008523

Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007). Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. climate regimes. Water Resources Research, 43(11), 1–10. https://doi.org/10.1029/2007WR006162

Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. Water Resources Research, 43(2), 1–14. https://doi.org/10.1029/2006WR005043

Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2009). Nonlinear storage-discharge relations and catchment streamflow regimes. Water Resources Research, 45(10), 1–16. https://doi.org/10.1029/2008WR007658

Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resources Research, 13(3), 637–643. https://doi.org/10.1029/WR013i003p00637

Ceola, S., Botter, G., Bertuzzo, E., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). Comparative study of ecohydrological streamflow probability distributions. Water Resources Research, 46(9), 1–12. https://doi.org/10.1029/2010WR009102

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. SIAM Review, 51(4), 661–703. https://doi.org/10.1137/070710111

Cunderlik, J. M., & Burn, D. H. (2002). Utilisation d'une information sur le régime des crues dans une analyse fréquentielle régionale des crues. Hydrological Sciences Journal, 47(1), 77–92. https://doi.org/10.1080/02626660209492909

Derrick, B., Toher, D., & White, P. (2016). Why Welchs test is Type I error robust. The Quantitative Methods for Psychology, 12(1), 30–38. https://doi.org/doi:10.20982/tqmp.12.1.p030

Doulatyari, B., Betterle, A., Basso, S., Biswal, B., Schirmer, M., & Botter, G. (2015). Predicting streamflow distributions and flow duration curves from landscape and climate. Advances in Water Resources, 83, 285–298. https://doi.org/10.1016/j.advwatres.2015.06.013

Durrans, S. R., Eiffe, M. A., Thomas, W. O., & Goranflo, H. M. (2003). Joint Seasonal /Annual Flood Frequency Analysis. Journal of Hydrologic Engineering, 8(4), 181–189. https://doi.org/10.1061/(asce)1084-0699(2003)8:4(181)

El Adlouni, S., Bobée, B., & Ouarda, T. B. M. J. (2008). On the tails of extreme event distributions in hydrology. Journal of Hydrology, 355(1–4), 16–33. https://doi.org/10.1016/j.jhydrol.2008.02.011

Fiorentino, M., Manfreda, S., & Iacobellis, V. (2007). Peak runoff contributing area as hydrological signature of the probability distribution of floods. Advances in Water Resources, 30(10), 2123–2134. https://doi.org/10.1016/j.advwatres.2006.11.017

Fischer, S., & Schumann, A. (2016). Robust flood statistics: comparison of peak over threshold approaches based on monthly maxima and TL-moments. Hydrological Sciences Journal, 61(3), 457–470. https://doi.org/10.1080/02626667.2015.1054391

Gaume, E. (2006). On the asymptotic behavior of flood peak distributions. Hydrology and Earth System Sciences, 10(2), 233–243. https://doi.org/10.5194/hess-10-233-2006

Gioia, A., Iacobellis, V., Manfreda, S., & Fiorentino, M. (2008). Runoff thresholds in derived flood frequency distributions. Hydrology and Earth System Sciences, 12(6), 1295–1307. https://doi.org/10.5194/hess-12-1295-2008

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

Huntingford, C., Marsh, T., Scaife, A. A., Kendon, E. J., Hannaford, J., Kay, A. L., et al. (2014). Potential influences on the United Kingdom’s floods of winter 2013/14. Nature Climate Change, 4(9), 769–777. https://doi.org/10.1038/nclimate2314

Jachens, E. R., Rupp, D. E., Roques, C., & Selker, J. S. (2020). Recession analysis revisited: Impacts of climate on parameter estimation. Hydrology and Earth System Sciences, 24(3), 1159–1170. https://doi.org/10.5194/hess-24-1159-2020

Katz, R. (2002). Statistics of Extremes in Climatology and Hydrology. Advances in Water Resources, 25, 1287–1304.

Kondor, D., Pósfai, M., Csabai, I., & Vattay, G. (2014). Do the rich get richer? An empirical analysis of the Bitcoin transaction network. PLoS ONE, 9(2). https://doi.org/10.1371/journal.pone.0086197

Laio, F., Porporato, A., Fernandez-Illescas, C. P., & Rodriguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and responce to water stress IV. Discussion of real cases. Advances in Water Resources, 24(7), 745–762. https://doi.org/10.1016/S0309-1708(01)00007-0

Lehner, B., Liermann, C. R., Revenga, C., Vörömsmarty, C., Fekete, B., Crouzet, P., et al. (2011). High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment, 9(9), 494–502. https://doi.org/10.1890/100125

Lins, H. F. (2008). Challenges to hydrological observations. WMO Bulletin, 57(January), 55–58.

Lu, P., Smith, J. A., & Lin, N. (2017). Spatial characterization of flood magnitudes over the drainage network of the Delaware river basin. Journal of Hydrometeorology, 18(4), 957–976. https://doi.org/10.1175/JHM-D-16-0071.1

Macdonald, E., Merz, B., Guse, B., Wietzke, L., Ullrich, S., Kemter, M., et al. (2022). Event and Catchment Controls of Heavy Tail Behavior of Floods. Water Resources Research, 58(6), 1–25. https://doi.org/10.1029/2021wr031260

Malamud, B. D. (2004). Tails of natural hazards. Physics World, 17(8), 31–35. https://doi.org/10.1088/2058-7058/17/8/35

Malamud, B. D., & Turcotte, D. L. (2006). The applicability of power-law frequency statistics to floods. Journal of Hydrology, 322(1–4), 168–180. https://doi.org/10.1016/j.jhydrol.2005.02.032

McCuen, R. H., & Smith, E. (2008). Origin of Flood Skew. Journal of Hydrologic Engineering, 13(9), 771–775. https://doi.org/10.1061/(asce)1084-0699(2008)13:9(771)

McDermott, T. K. J. (2022). Global exposure to flood risk and poverty. Nature Communications, 13(1), 6–8. https://doi.org/10.1038/s41467-022-30725-6

Meigh, J. R., Farquharson, F. A. K., & Sutcliffe, J. V. (1997). A worldwide comparison of regional flood estimation methods and climate. Hydrological Sciences Journal, 42(2), 225–244. https://doi.org/10.1080/02626669709492022

Mejía, A., Daly, E., Rossel, F., Javanovic, T., & Gironás, J. (2014). A stochastic model of streamflow for urbanized basins. Water Resources Research, 50, 1984–2001. https://doi.org/10.1002/2013WR014834

Merz, B., Basso, S., Fischer, S., Lun, D., Blöschl, G., Merz, R., et al. (2022). Understanding heavy tails of flood peak distributions. Water Resources Research, 1–37. https://doi.org/10.1029/2021wr030506

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. Nature Reviews Earth and Environment, 2(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3

Merz, R., & Blöschl, G. (2009). Process controls on the statistical flood moments - a data based analysis. Hydrological Processes, 23(5), 675–696. https://doi.org/10.1002/hyp

Molnar, P., Anderson, R. S., Kier, G., & Rose, J. (2006). Relationships among probability distributions of stream discharges in floods, climate, bed load transport, and river incision. Journal of Geophysical Research: Earth Surface, 111(2), 1–10. https://doi.org/10.1029/2005JF000310

Morrison, J. E., & Smith, J. A. (2002). Stochastic modeling of flood peaks using the generalized extreme value distribution. Water Resources Research, 38(12), 41-1-41–12. https://doi.org/10.1029/2001wr000502

Müller, M. F., Dralle, D. N., & Thompson, S. E. (2014). Analytical model for flow duration curves in seasonally dry climates. Water Resources Research, 50, 5510–5531. https://doi.org/10.1002/2014WR015301

Müller, M. F., Roche, K. R., & Dralle, D. N. (2021). Catchment processes can amplify the effect of increasing rainfall variability. Environmental Research Letters, 16(8). https://doi.org/10.1088/1748-9326/ac153e

Mushtaq, S., Miniussi, A., Merz, R., & Basso, S. (2022). Reliable estimation of high floods: A method to select the most suitable ordinary distribution in the Metastatistical extreme value framework. Advances in Water Resources, 161(September 2021), 104127. https://doi.org/10.1016/j.advwatres.2022.104127

Pallard, B., Castellarin, A., & Montanari, A. (2009). A look at the links between drainage density and flood statistics. Hydrol. Earth Syst. Sci (Vol. 13). Retrieved from www.hydrol-earth-syst-sci.net/13/1019/2009/

Papalexiou, S. M., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? An assessment of daily rainfall distribution tails. Hydrology and Earth System Sciences, 17(2), 851–862. https://doi.org/10.5194/hess-17-851-2013

Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. American Naturalist, 164(5), 625–632. https://doi.org/10.1086/424970

Pumo, D., Viola, F., La Loggia, G., & Noto, L. V. (2014). Annual flow duration curves assessment in ephemeral small basins. Journal of Hydrology, 519(PA), 258–270. https://doi.org/10.1016/j.jhydrol.2014.07.024

Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., & Merz, R. (2012). Step changes in the flood frequency curve : Process controls. Water Resources Research, 48, 1–15. https://doi.org/10.1029/2011WR011187

Santos, A. C., Portela, M. M., Rinaldo, A., & Schaefli, B. (2018). Analytical flow duration curves for summer streamflow in Switzerland. Hydrology and Earth System Sciences, 22(4), 2377–2389. https://doi.org/10.5194/hess-22-2377-2018

Sartori, M., & Schiavo, S. (2015). Connected we stand: A network perspective on trade and global food security. Food Policy, 57, 114–127. https://doi.org/https://doi.org/10.1016/j.foodpol.2015.10.004

Schaefli, B., Rinaldo, A., & Botter, G. (2013). Analytic probability distributions for snow-dominated streamflow. Water Resources Research, 49(5), 2701–2713. https://doi.org/10.1002/wrcr.20234

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

Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L., & Bates, P. (2018). Strange Floods: The Upper Tail of Flood Peaks in the United States. Water Resources Research, 54(9), 6510–6542. https://doi.org/10.1029/2018WR022539

Spearman, C. (1904). The proof and measurement of association between two things. American Journal of Psychology, 15(1), 72–101. https://doi.org/10.2307/1412159

Struthers, I., & Sivapalan, M. (2007). A conceptual investigation of process controls upon flood frequency: Role of thresholds. Hydrology and Earth System Sciences, 11(4), 1405–1416. https://doi.org/10.5194/hess-11-1405-2007

Tarasova, L., Basso, S., & Merz, R. (2020). Transformation of Generation Processes From Small Runoff Events to Large Floods. Geophysical Research Letters, 47(22). https://doi.org/10.1029/2020GL090547

Tarasova, L., Basso, S., Zink, M., & Merz, R. (2018). Exploring Controls on Rainfall-Runoff Events: 1. Time Series-Based Event Separation and Temporal Dynamics of Event Runoff Response in Germany. Water Resources Research, 54(10), 7711–7732. https://doi.org/10.1029/2018WR022587

Tarasova, L., Lun, D., Merz, R., Blöschl, G., Basso, S., Bertola, M., et al. (2023). Shifts in flood generation processes exacerbate regional flood anomalies in Europe. Communications Earth and Environment, 4(1), 1–12. https://doi.org/10.1038/s43247-023-00714-8

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

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

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

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

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

Villarini, G., Smith, J. A., Baeck, M. L., Marchok, T., & Vecchi, G. A. (2011). Characterization of rainfall distribution and flooding associated with U.S. landfalling tropical cyclones: Analyses of Hurricanes Frances, Ivan, and Jeanne (2004). Journal of Geophysical Research Atmospheres, 116(23). https://doi.org/10.1029/2011JD016175

Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D., & Wong, W. K. (2016). Evidence for changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway. Journal of Hydrology, 538, 33–48. https://doi.org/10.1016/j.jhydrol.2016.03.066

Wang, H.-J., Merz, R., Yang, S., & Basso, S. (2023). Inferring heavy tails of flood distributions from common discharge dynamics. EGUsphere, 2023(May), 1–24. https://doi.org/https://doi.org/10.5194/egusphere-2023-660

Wang, H., Merz, R., Yang, S., Tarasova, L., & Basso, S. (2022). Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. Advances in Water Resources Journal, 171(104359). https://doi.org/10.1016/j.advwatres.2022.104359

Ward, A. S., Wondzell, S. M., Schmadel, N. M., & Herzog, S. P. (2020). Climate Change Causes River Network Contraction and Disconnection in the H.J. Andrews Experimental Forest, Oregon, USA. Frontiers in Water, 2(April), 1–10. https://doi.org/10.3389/frwa.2020.00007

Welch, B. L. (1947). The generalization of “Student’s” problem when several different population variances are involved. Biometrika, 34(1–2), 28–35. https://doi.org/doi:10.1093/biomet/34.1-2.28. MR 0019277. PMID 20287819

Wietzke, L. M., Merz, B., Gerlitz, L., Kreibich, H., Guse, B., Castellarin, A., & Vorogushyn, S. (2020). Comparative analysis of scalar upper tail indicators. Hydrological Sciences Journal, 65(10), 1625–1639. https://doi.org/10.1080/02626667.2020.1769104

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics Bulletin, 1(6), 80–83. https://doi.org/10.2307/3001968

Wu, Q., Ke, L., Wang, J., Pavelsky, T. M., Allen, G. H., Sheng, Y., et al. (2023). Satellites reveal hotspots of global river extent change. Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-37061-3

Constructing a geography of heavy-tailed flood distributions: insights from common streamflow dynamics

Hsing-Jui Wang[1], Ralf Merz[1,2], and Stefano Basso[3]

[1]Department of Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale), 06120, Germany,

[2]Institute of Geosciences and Geography, Martin-Luther University Halle-Wittenberg, Halle (Saale), 06120, Germany,

[3]Norwegian Institute for Water Research (NIVA), Oslo, 0579, Norway

Correspondence to: Hsing–Jui Wang (hsing-jui.wang@ufz.de)

Key Points

1. Regional propensity of flood tail behavior is shown across a diverse geographical range based on the analysis of common streamflow dynamics.
2. Temporal variability in catchment storage, driven by high evapotranspiration and dry soils, is a key mechanism for heavy-tailed floods.
3. Examining flood generation processes aids in unraveling the connections between flood tail behavior and seasons or catchment sizes.

Abstract

Heavy-tailed flood distributions depict the higher occurrence probability of extreme floods. Understanding the spatial distribution of heavy tail floods is essential for effective risk assessment. Conventional methods often encounter data limitations, leading to uncertainty across regions. To address this challenge, we utilize hydrograph recession exponents derived from common streamflow dynamics, which have proven to be a robust indicator of flood tail propensity across analyses with varying data lengths. Analyzing extensive datasets from Germany, the United Kingdom (UK), Norway, and the United States (US), we uncover distinct patterns: prevalent heavy tails in Germany and the UK, diverse behavior in the US, and predominantly nonheavy tails in Norway. The regional tail behavior has been observed in relation to the interplay between terrain and meteorological characteristics, and we further conducted quantitative analyses to assess the influence of hydroclimatic conditions using Köppen classifications. Notably, temporal variations in catchment storage are a crucial mechanism driving highly nonlinear catchment responses that favor heavy-tailed floods, often intensified by concurrent dry periods and high temperatures. Furthermore, this mechanism is influenced by various flood generation processes, which can be shaped by both hydroclimatic seasonality and catchment scale. These insights deepen our understanding of the interplay between climate, physiographical settings, and flood behavior, while highlighting the utility of hydrograph recession exponents in flood hazard assessment.

1 Introduction

Floods are devastating natural hazards that pose significant risks to infrastructure, property, and human life (McDermott, 2022; Bevere and Remondi, 2022). The unprecedented magnitude of extreme floods often characterizes these hazards, which is better depicted by the heavy-tailed behavior exhibited in flood frequency distributions (Smith et al., 2018; Merz et al., 2021; Merz et al., 2022). The concept of heavy-tailed behavior finds broad application in various fields to describe the likelihood of extreme event occurrences (Katz, 2002; Kondor et al., 2014; Malamud, 2004; Sartori and Schiavo, 2015; Wang et al., 2022). In particular, it is widely recognized as a prevalent feature in hydrologic extremes (Papalexiou and Koutsoyiannis, 2013; Smith et al., 2018).

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

In the review of Merz et al. (2022), it becomes evident that multiple hydro-physiographic characteristics interact within a complex system, collectively shaping flood tail behavior. Specifically, the interplay between characteristic flood generation (Bernardara et al., 2008; Thorarinsdottir et al., 2018), the presence of mixed flood types (Morrison and Smith, 2002; Villarini and Smith, 2010), the tail heaviness of rainfall distributions (Gaume, 2006), catchment aridity (Molnar et al., 2006; Merz and Blöschl, 2009; Guo et al., 2014), and catchment area (Pallard et al., 2009; Villarini and Smith, 2010) are proposed as contributing factors to the nonlinearity of catchment responses. This nonlinearity is increasingly recognized as a plausible driver of heavy-tailed flood behavior (Fiorentino et al., 2007; Struthers and Sivapalan, 2007; Gioia et al., 2008; Rogger et al., 2012; Basso et al., 2015; Merz et al., 2022; Basso et al., 2023; Wang et al., 2023).

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

$$\frac{dq}{dt} = -B \cdot q^{a}$$

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

In our prior research (Wang et al., 2023), we introduced hydrograph recession exponents as a newly proposed indicator for heavy-tailed flood behavior. This indicator allows for an inference of heavy-tailed flood distributions based on physical mechanisms (i.e., typical hydrological processes within common streamflow dynamics). Importantly, it has shown its capacity to provide robust estimates for both short and long data records. Unlike traditional methods for identifying heavy tails, such as the upper tail ratio (Villarini et al., 2011; Lu et al., 2017) and the shape parameter of a fitted Generalized Extreme Value (GEV) distribution (Morrison and Smith, 2002; Villarini and Smith, 2010), which are sensitive to sample sizes (Wietzke et al., 2020), recession exponents offer more consistent estimates of flood tail heaviness across various data lengths. This characteristic makes it a valuable tool for analyzing regions with diverse gauge data records.

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

We organize the structure of this paper as follows: Section 2 describes the study areas and the hydrological data based on an extensive dataset composed of four countries, Section 3 describes the methods of estimation and validation of hydrograph recession exponents in identifying heavy-tailed flood behavior in the dataset, the framework of the analyses of spatial patterns of inferred heavy-tailed flood behavior, the framework of the analyses of seasonal dynamics of inferred heavy-tailed flood behavior, and statistical tests. In Section 4, we present the validation results of our heavy-tailed flood behavior index, along with analyses of the relationships between flood tail behavior and geographical spatial characteristics, seasonal patterns, and catchment scales in these comparable countries. Physical interpretations of the results and remarks from the literature are discussed in Section 5. The main conclusions are summarized in Section 6.

2 Study areas and data

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

(US), known for its diverse range of physiographical settings, enables us to validate and consolidate the transferability of our findings.

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

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

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

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

We collected daily continuous streamflow records from 575 gauges across the four study regions to conduct our analyses. The corresponding drainage areas range from 4 to 40,504 km^2. Our analysis was performed on a seasonal basis, considering spring (March-May), summer (June-August), autumn (September-November), and winter (December-February) to account for the seasonality of hydrograph recessions (Tashie et al., 2020) and flood distributions (Durrans et al., 2003). Each analysis conducted on a specific river gauge during a season was treated as a case study.

We excluded gauges located downstream of large dams in all four regions (Lehner et al., 2011; Wang et al., 2022). Consistent with previous studies (e.g., Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Basso et al., 2021; Basso et al., 2023), we chose case studies in Germany, the UK, and the US characterized by limited snowfall, which minimizes the potential transfer of water across seasons due to strong snow accumulation and melting. Specifically, this condition is defined as having an average daily temperature below zero degrees Celsius during precipitation events for over 50% of a season (Basso et al., 2021). However, recognizing that recession exponents can inherently capture both linear and nonlinear catchment responses, we intentionally included case studies in Norway, which are characterized by a dominant runoff generation process driven by snow dynamics. This deliberate inclusion provides a counter-verification, allowing us to explore the capability of the recession exponent as a measure of flood tail behavior in regions primarily characterized by snowmelt-driven flood generation processes. In summary, this analysis encompasses regions dominated by both rainfall-driven and snowmelt-driven floods, providing an extensive examination of these factors. These procedures resulted in a total of 1997 case studies, distributed as follows: 540 in spring, 520 in summer, 543 in autumn, and 394 in winter (refer to Table A1 for detailed information of each region).

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

3 Methods

3.1 Inferring Heavy Tails of Flood Distributions from Common Streamflow Dynamics

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

Within the PHEV framework, we obtain consistent expressions for the probability distributions of various flow metrics, including daily streamflow, ordinary peak flows (local flow peaks resulting from streamflow-producing rainfall events), and floods (flow maxima within a specified timeframe) (Basso et al., 2016). The description of runoff generation and streamflow dynamics provided by this framework has been successfully tested across a diverse range of hydroclimatic and physiographic conditions through a number of studies (Arai et al., 2020; Botter et al., 2007a; Botter et al., 2010; Ceola et al., 2010; Doulatyari et al., 2015; Mejía et al., 2014; Müller et al., 2014; Müller et al., 2021; Pumo et al., 2014; Santos et al., 2018; Schaefli et al., 2013).

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

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

3.2 Validation of Hydrograph Recession Exponents as An Index of Heavy-Tailed Flood Behavior

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

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

We conduct our approach using three distinct empirical data distributions: daily streamflow, ordinary peaks, and monthly maxima. These multiple analyses strengthen our validation process and enhance the evaluation of our results. It's worth noting that our chosen benchmark, the empirical power law, may be influenced by fitting uncertainty due to data scarcity in certain cases, particularly when analyzing maxima. To mitigate this, we consider monthly maxima (Fischer and Schumann, 2016; Malamud and Turcotte, 2006) instead of the seasonal maxima previously used in the literature (e.g., Basso et al., 2021) in order to expand the sample size. Parallel analyses for cases with larger sample sizes (i.e., daily streamflow and ordinary peaks) provide more robust validation and lend support to the interpretation of results for maxima.

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

3.3 Analyses of Spatial and Seasonal Patterns of Inferred Flood Tail Behavior

We construct a geographical representation of inferred heavy-tailed flood behavior by utilizing estimated recession exponents derived from common streamflow dynamics across study countries and for each season. This representation serves as an evaluation of the propensity of heavy-tailed flood behavior across various regions and seasons. We simplify the seasonal results by identifying the dominant tail behavior, which refers to the majority of seasons exhibiting either heavy-tailed or nonheavy-tailed behavior, as the representative inferred flood tail behavior in the analysis of spatial pattern (Section 4.2).

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

To analyze seasonal patterns, we initially investigate the coherence of inferred flood tail behavior across seasons, focusing on consistency between heavy- or nonheavy-tailed behavior. Catchments with valid recession exponents from only one season are omitted from this analysis. As a result, the selection comprises 98 out of 98 catchments in Germany, 81 out of 82 in the UK, 79 out of 82 in Norway, and 290 out of 313 in the US. We also employ the Wilcoxon signed-rank test (Wilcoxon, 1945), a non-parametric statistical hypothesis test, at a significance level of 0.05 in this section. This test assesses whether the median of recession exponents (within a climate group on a seasonal basis) is above two, below two, or shows no significant difference from two (Figure 7).

4 Results

4.1 Effectiveness of Identifying Heavy-tailed Flood Behavior Using Common Discharge Dynamics

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

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

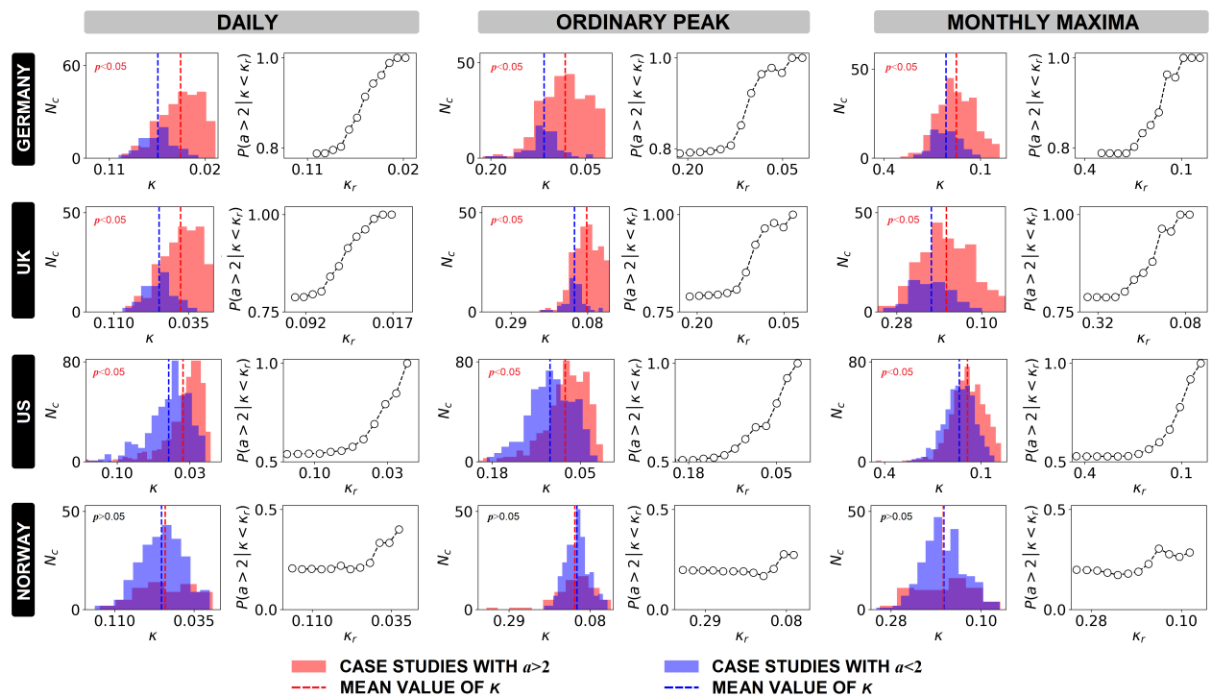


Figure 1. Effectiveness of identifying heavy-tailed flood behavior using hydrograph recession exponents. Case studies in each country are categorized into two groups: red histograms representing recession exponents a above two and blue histograms representing recession exponents a below two. In all three analyses (daily streamflow, ordinary peak flow, and monthly maxima), every case study is subjected to empirical power law fitting, resulting in a representative power law for the dataset, measured by the KS statistic κ (where $\kappa \in [0,\infty]$ and $\kappa=0$ signifies maximum reliability). The histograms portray the count of case studies $\boldsymbol{N_c}$ analyzed as a function of $\boldsymbol{\kappa}$ for two distinct groups. Dashed lines on the histogram plots indicate the means of the histograms. The means of two groups (a>2 and a<2) are subjected to Welch's t-test at a significance level of 0.05 to determine whether they are significantly different ($p < 0.05$) or not ($p > 0.05$). The line chart shows the accuracy of using the recession exponent to identify heavy-tailed behavior (denoted as $\boldsymbol{P(a > 2|\kappa < \kappa_r) = N_c(a > 2|\kappa < \kappa_r)/N_c(\kappa < \kappa_r)}$) as the $\boldsymbol{\kappa_r}$ threshold decreases (i.e., as the reliability of empirical power laws increases). The results for Germany are reproduced from Wang et al. (2023).

In Figure 2, we explore the correlation between the values of empirical power law exponents b and the values of recession exponents a for case studies confirmed to exhibit heavy-tailed behavior. This is achieved by utilizing the goodness-of-fit testing procedure of Clauset et al. (2009) to categorize case studies into ‘confirmed power-law-tailed case studies’ and ‘uncertain case studies.’ The former are depicted as black dots, while the latter are depicted as gray dots. The presence of a sizable number of uncertain case studies indicates the difficulty of establishing with certainty whether the underlying distribution of empirical data is or not a power law. This difficulty is often due to limited data availability, although the possibility that they indeed do not follow power laws cannot be excluded.

To highlight the correlation, we binned the confirmed power-law-tailed case studies and used red markers showing the median values of *a* and *b* (squares), the interquartile intervals of *b* (vertical bars), and the binning ranges of *a*. In each country, the composition of each bin encompasses one-seventh of the total number of case studies, except for Norway, where this fraction is adjusted to one-fifth due to the limited number of confirmed power-law-tailed cases. We calculated Spearman correlations $\boldsymbol{r_s}$ (Spearman, 1904) to test the correlation between *a* and *b*, which is valid for both linear and nonlinear associations between random variables. We found that *a* and *b* are significantly correlated at a significance level of 0.05 in Germany, the US, and the UK. In these three countries, a larger number of uncertain case studies emerge in the analysis of flow maxima compared to the analysis of daily streamflow and ordinary peak flow (respectively for daily streamflow, ordinary peaks, and flow maxima: 265, 270, and 352 out of 386 case studies in Germany; 258, 280, and 306 out of 325 case studies in the UK; and 589, 624, and 836 out of 980 case studies in the US). Since the same case studies have already been confirmed to exhibit power-law-tailed distributions in their daily streamflow and ordinary peak flow data, the increase of uncertain case studies in the analysis of flow maxima suggests that the greater level of uncertainty is due to limited data availability rather than indicating a rise in the number of non-power-law-tailed case studies.

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

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

Conversely, regions with lower recession exponents (below two), reflecting relatively linear responses in catchments, are more likely to signify nonheavy-tailed flood behavior.

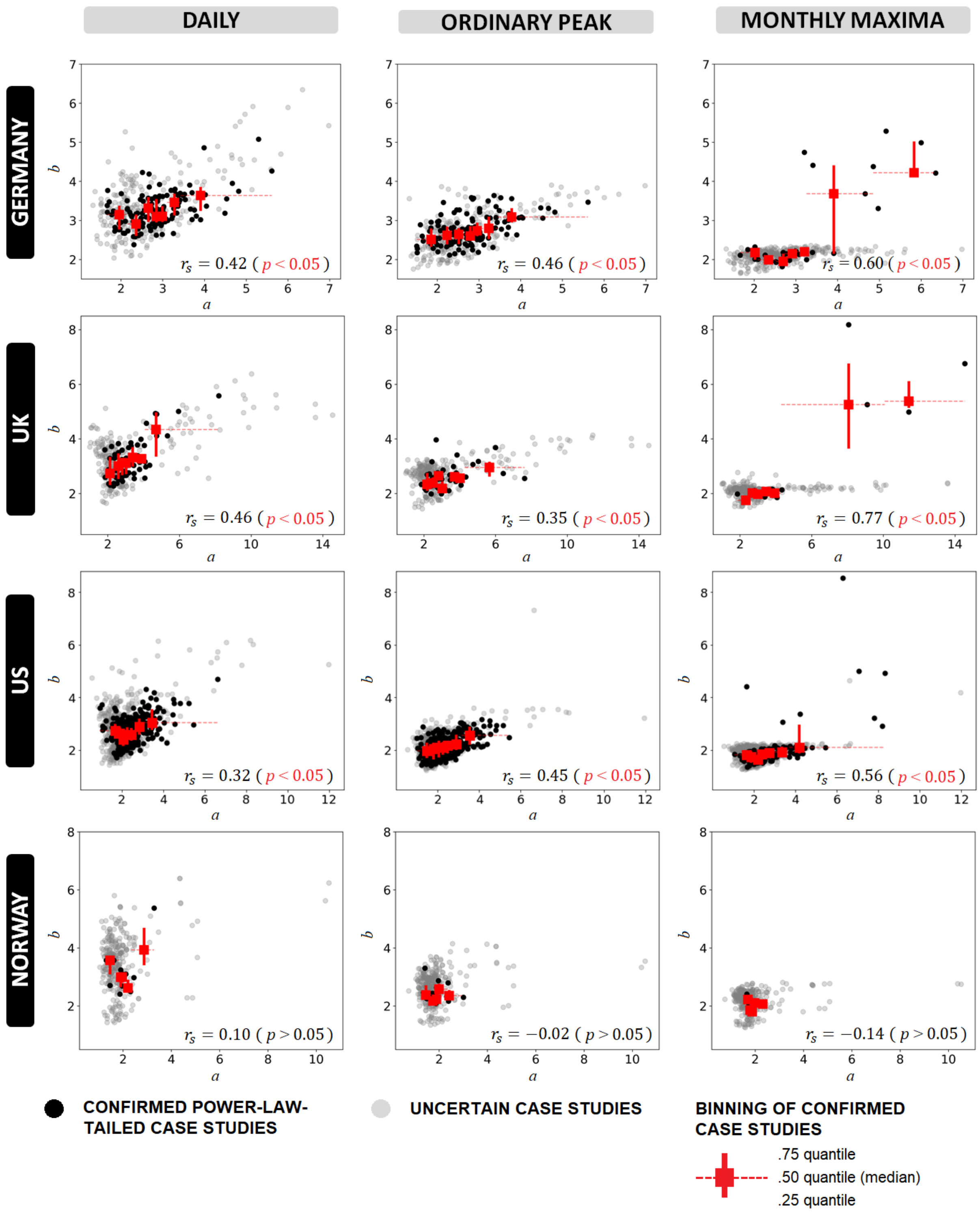


Figure 2. Empirical power law exponent *b* as a function of the hydrograph recession exponent *a* (physically-based index of heavy-tailed flood behavior). Case studies are classified

into groups of confirmed power-law-tailed (black dots) and uncertain (gray dots) case studies on the basis of the goodness-of-fit test (Clauset et al., 2009). The former group shows statistical confirmation that the data's distribution tail can be accurately characterized by a power law, indicating heavy-tailed behavior. Conversely, the latter group indicates our inability to statistically affirm whether the data follows a power law distribution or not. For the confirmed power-law-tailed case studies, the correlation between the empirical power law exponent *b* and the hydrograph recession exponent *a* is underlined by red markers. This correlation is quantified using the Spearman correlation coefficient $\boldsymbol{r_s}$ at a significance level of 0.05. The squares represent the median values of *a* and *b*, vertical bars indicate the interquartile intervals of *b*, and horizontal dashed bars indicate the binning ranges of *a*. In each country, the composition of each bin encompasses one-seventh of the total number of case studies, except for Norway, where this fraction is adjusted to one-fifth due to the constraint posed by the total number of confirmed power-law-tailed case studies. The count of the confirmed power-law-tailed case studies in the analyses of daily streamflows, ordinary peak flows, and monthly flow maxima are as follows: 121, 116, and 34 out of 386 case studies for Germany, respectively; 67, 45, and 19 out of 325 case studies for the UK, respectively; 391, 356, and 144 out of 980 case studies for the US, respectively; and 15, 17, and 6 out of 306 case studies for Norway, respectively. The results for Germany are reproduced from Wang et al. (2023).

4.2 Spatial Patterns of Inferred Flood Tail Behavior

Figure 3 displays the spatial distribution of dominant flood tail behavior across seasons, based on the recession exponent values, for Germany, the UK, Norway, and the US, respectively. This dominant behavior represents either heavy or nonheavy tails, depending on what is observed in the majority of seasons. Additionally, Figure 4 and Table 1 provide quantitative analyses of the propensity of flood tail behavior across different regions.

In Germany (Figure 3a), approximately 81% of catchments are identified as sites with dominant heavy-tailed flood behavior (red dots), indicating a prevalence of such behavior. This result agrees with the findings of Mushtaq et al. (2022), which reported that a distribution with a relatively heavier tail (i.e., the log-normal) best represent ordinary peak flows in the majority of German basins considered in their study. The inferred heavy-tailed sites are spread across Germany. They dominate in the eastern part, while there are mixed patterns of heavy- and nonheavy-tailed behavior in the western part. The climate conditions are primarily humid continental (Dfb) and temperate oceanic (Cfb). Humid continental climate is prominent in the east, while temperate oceanic climate generally covers the west.

In the UK (Figure 3b), four climate types are present, with temperate oceanic climate (Cfb) being the dominant one. The terrain of this country in comparison to the other three countries is relatively homogeneous, with no high mountains. According to our findings, heavy-tailed flood behavior is prevalent in the UK, with a prevalence of 77%, especially in the eastern and southern coastal regions. Huntingford et al. (2014) reported a case in which a rapid succession of vigorous Atlantic low-pressure systems crossed much of the UK, resulting in repeated heavy rainfall events. Southeast England was identified as a distinct region characterized by exceptionally high flows, exacerbated by increasingly saturated catchments. These catchment characteristics and hydrological responses align with our findings, which indicate the pronounced heavy tails in such a region.

In Norway (Figure 3c), however, nonheavy-tailed flood behavior dominates. Approximately 89% of sites are inferred to have nonheavy-tailed flood behavior. Norway encompasses nine climate types but is primarily covered by Subarctic climate (Dfc), characterized by low temperatures and reduced evapotranspiration. Hydrological processes are significantly influenced by snow dynamics, which generally determine linear catchment responses as a result of snow accumulation and melting processes (Santos et al., 2018).

In contrast to the aforementioned countries with relatively consistent climate and dominant flood behavior, the US (Figure 3d) display a diverse range of climate types and a balanced propensity toward heavy- and nonheavy-tailed flood behavior. The eastern regions dominated by humid subtropical climate (Cfa), hot-summer humid continental climate (Dfa), and temperate oceanic climate (Cfb) from south to north. The interior western states feature a cold semi-arid climate (BSk), while mixed patterns are observed in the western mountainous and coastal areas. An overall relatively even distribution of inferred heavy-tailed (52%) and nonheavy-tailed (48%) flood behavior prevails in this diverse climate country.

Figure 3e provides an example of how the spatial distribution of flood behavior is influenced by regional physioclimatic features. In particular, catchments on the east side of the mountains exhibit pronounced heavy-tailed flood behavior, which aligns with the findings of Smith et al. (2018). This is likely due to the interaction between cold air from the inland polar jet stream and warm ocean currents leads to the formation of Nor'easters, which are synoptic-scale extratropical cyclones in the western North Atlantic Ocean along the US northeast coast. These weather systems often resulted in heavy rain or rain-on-snow events. Conversely, on the west side of mountains, catchments tend to exhibit nonheavy-tailed behavior, potentially due to the leeward rain shadow effect.

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

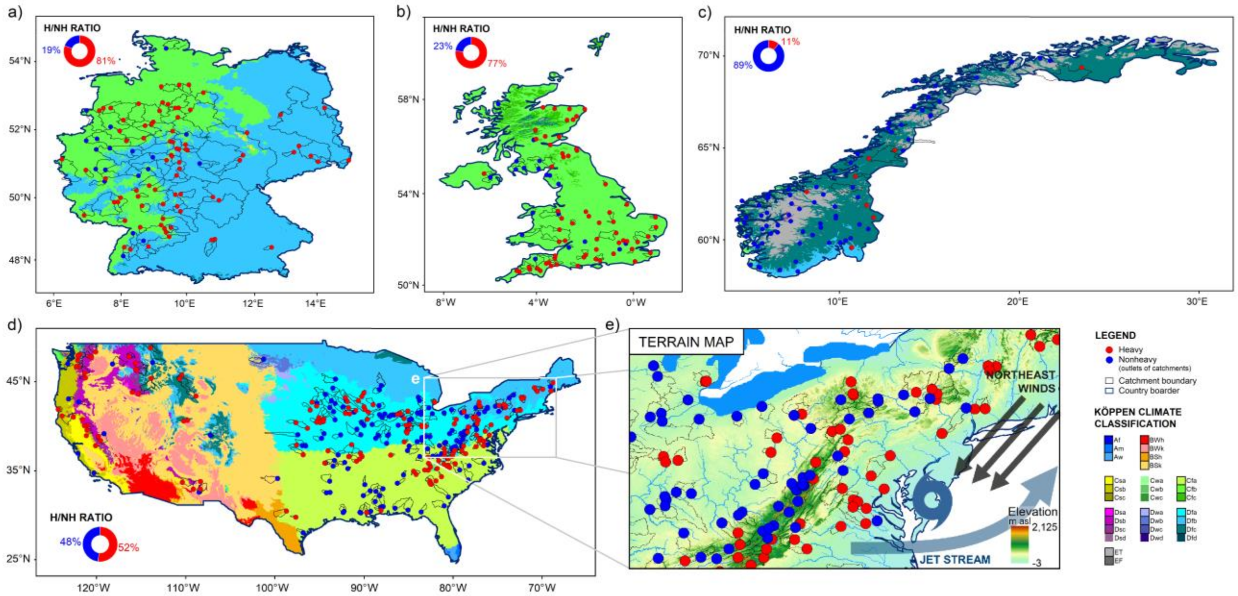


Figure 3. Spatial distribution of dominant flood behavior. The dominant pattern determines the representative flood tail behavior of catchments across all study countries, whether it is heavy or non-heavy, which is defined by the major pattern recognized across seasons. Tail behavior is inferred by hydrograph recession exponents. The ratio of heavy- to nonheavy-tailed catchments is indicated as H/NH ratio. Köppen climate classification based on Beck et al. (2018), with present climate types outlined by bold dark frames in the legend. (a) Germany, a total of 98 gauges represent catchments ranging from 110 to 23,843 km^2, with a median area of 1,195 km^2. (b) The UK, a total of 82 gauges represent catchments ranging from 15 to 9,948 km^2, with a median area of 283 km^2. (c) Norway, a total of 82 gauges represent catchments ranging from 4 to 40,504 km^2, with a median area of 234 km^2. (Note: some catchment boundaries are absent in the dataset for catchments in the UK and Norway.) (d) The US, a total of 313 gauges represent catchments ranging from 66 to 9,935 km^2, with a median area of 1769 km^2. (e) A zoomed-in map illustrates the discernible patterns of flood tail behavior resulting from specific flood generation processes influenced by the interplay between regional terrain and meteorological features. (Note: the cylinder map projection is employed in these maps.)

To obtain quantitative results we examine the predominant flood tail behavior (inferred by recession exponents) of catchments across various climate regions and sort these regions based on the proportion of heavy-tailed catchments from high to low, as illustrated in Figure 4. By categorizing climate type regions based on the proportion of heavy-tailed catchments, we establish three groups according to their propensity of flood tail behavior: Heavy-tailed group, indicating regions with over 66.6% of catchments dominated by heavy tails; Neutral group, encompassing regions with 33.3% to 66.6% of catchments dominated by heavy tails, represents a relatively even propensity for both heavy and nonheavy tails in the catchments within these regions; and Nonheavy-tailed group, representing regions with less than 33.3% of catchments dominated by heavy tails, denotes the propensity for nonheavy tails. According to the Köppen climate type classification, the overarching hydroclimatic characteristics can be delineated by three hierarchical features: 1. the main group, which encompasses five areas—Tropical, Arid, Temperate, Continental, and Polar; 2. precipitation characteristics; and 3. temperature characteristics. The findings are synthesized in Figure 4 and Table 1, where the groups of flood tail behavior propensity are juxtaposed with the distinctive traits of each climate region.

Five climate regions are identified as having a higher propensity for heavy tails: mediterranean climate (Csa), hot semi-arid climates (BSh), humid continental climate (Dsb), temperate oceanic climate (Cfb), and cool-summer mediterranean climate (Csb). These regions are characterized by warm to hot temperatures, often accompanied by occasional dry periods (except for Cfb). Based on the definition of Köppen climate classification the occurrence of dry periods is a result of significantly uneven rainfall throughout the year, with at least three times as much rainfall in the wettest month as in the driest month. In semi-arid climates (BSh), there is generally lower annual rainfall (summarized in Table 1). Higher temperatures increase the potentail evapotranspiration, often enhancing atmospheric moisture content and facilitating convective rainfall. Moreover, the dynamics of evapotranspiration in hillslopes influence the nonlinearity of runoff processes in catchments (Tashie et al., 2019). Dry periods can lead to lower catchment soil moisture, facilitating nonlinear runoff generation (Merz and Blöschl, 2009; Viglione et al., 2009). The findings presented here indicate that heavy-tailed flood behavior tends to emerge due to the substantial nonlinearity observed in catchment hydrological processes, which is facilitated by temporally uneven rainfall and higher evapotranspiration variation throughout the year.

We also find that certain regions show a relatively neutral propensity regarding flood tail behavior (either heavy- or nonheavy-tailed) and aggregate them into the second group of Figure 4 and Table 1. These regions encompass cold semi-arid climates (BSk), humid continental climate (Dfb), humid subtropical climate (Cfa), and humid continental climate (Dfa). While cold semi-arid climates (BSk) experience dryness, they are characterized by very limit precipitation. In the other three regions (Dfb, Cfa, and Dfa), heavy tails may still occur due to higher evapotranspiration, which is driven by high temperatures. However, the relatively even distribution of rainfall throughout the year in these regions may reduce the propensity for heavy tails, resulting in a smoother occurrence of heavy-tailed flood behavior. In summary, the regions in this group still have a certain probability of exhibiting heavy-tailed flood behavior. However, the absence of either a drier state of the catchment (caused by uneven rainfall) or higher temperatures (that ensure sufficient atmospheric moisture for rainfall and strengthened nonlinearity) could constrain the occurrence of such behavior.

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

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

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

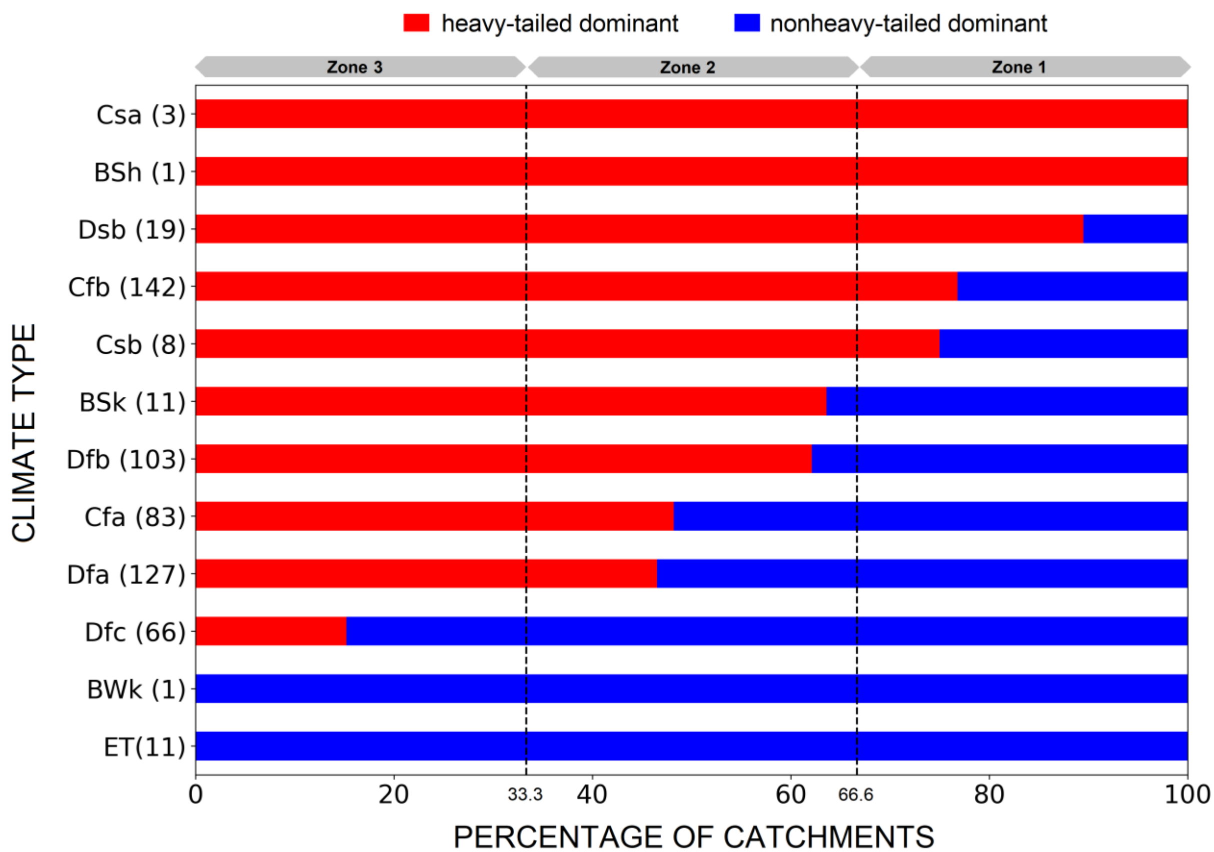


Figure 4. Propensity of inferred flood tail behavior in diverse climate regions. Catchments are categorized by climate types and grouped by dominant (across seasons) heavy-tailed case percentages. Three groups are defined by heavy-tailed case proportions: Zone 1 (>66.6%) represents heavy tails, Zone 2 (33.3-66.6%) is neutral, and Zone 3 (<33.3%) represents nonheavy

tails. The number of catchments in each climate region is indicated in parentheses after the climate type.

Table 1. Comparison of inferred flood tail behavior propensity with climate characteristics.

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

4.3 Seasonal patterns of Inferred Flood Tail Behavior

We analyze the seasonality of flood tail behavior, an aspect of this phenomenon which has been previously suggested but remains poorly understood (Durrans et al., 2003; Basso et al., 2015; Macdonald et al., 2022). Figure 5 illustrates the spatial distribution of catchments with consistent tail behavior across seasons (i.e., with either heavy or nonheavy tails across all seasons; black triangles) and those with varying tail behavior across seasons (green dots). Catchments exhibiting inconsistent behavior spread across the whole US and Germany, whereas they are mostly concentrated in the southern parts of the UK and in the central mountainous regions of Norway. The percentages of catchments exhibiting inconsistent flood tail behaviors are respectively 33%, 33%, 17%, and 34% in the US, Germany, the UK, and Norway. The results indicate that although the majority of catchments tend to exhibit stable heavy-/nonheavy-tailed behavior, still around one-third reveal changing patterns across seasons. Notably, there is a particularly high percentage of consistent patterns (83%) in the UK, likely due to the relatively uniform climate and terrain conditions across the country characterized by continuous rainfall throughout a year (as shown in Figure 3b).

We further investigate the dynamics of heavy- and nonheavy-tailed case studies across seasons in Figure 6. Heavy-tailed case studies increase from spring to autumn (approximately corresponding to the growing season in the northern hemisphere) and decrease from autumn to spring (approximately corresponding to the dormant season in the northern hemisphere), as seen in the aggregated patterns across all regions (panel a). This pattern can be attributed to the increasing temperature in the growing season, during which increasing evapotranspiration consumes water storage in the shallow subsurface, escalating the nonlinearity of catchment responses (Tashie et al., 2019). The seasonality of evapotranspiration effects on catchment nonlinearity is supported by the findings of Tarasova et al. (2018), who observed clear seasonal dynamics of catchment average runoff coefficients. These coefficients tend to be higher in wet winters and lower in dry summers. It has been shown that significant variation in runoff coefficients is linked to high nonlinearity of hydrological responses, facilitating heavier-tailed floods. This phenomenon is often observed in dry catchments (Merz and Blöschl, 2009). Other studies confirmed that the nonlinearity of catchment responses favors the emergence of heavy-tailed flood behavior (Gioia et al., 2008; Rogger et al., 2012; Basso et al. 2015), and is often expressed by quicker recession during high flow periods and greater stability during low flow periods. Conversely, during the dormant season, nonlinearity decreases due to reduced competition from evapotranspiration and replenished water storage. We underscore that the significant variability in evapotranspiration amplifies the fluctuation of catchment storage conditions, causing soil moisture levels to oscillate between drier and wetter states. This alternation leads to the occurrence of both very small and very large events, which are characteristic of heavy-tailed flood behavior.

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

We delve into the seasonal characteristics of this behavior further by combining the regional patterns based on climate classification. In Figure 7, the square dots represent the median of each box, marked as red, blue, or black to indicate the significance of its value above 2 (heavy), below 2 (nonheavy), or not significantly different from 2, respectively. The last one (black squares) may imply an equal occurrence of heavy- and nonheavy-tailed cases or a lack of samples to draw conclusions. Based on the patterns of significance across seasons, regions with seasonality (defined as having different tail behavior propensity across seasons according to the significant values of median) are grouped in the white area, while those considered stable in heavy tails are in the red area, and those stable in nonheavy tails are in the blue area. For regions where statistical significance cannot be concluded for all seasons, we group them based on the absolute values of their medians.

We find that the grouping based on their distinct patterns of seasonality (Figure 7) closely aligns with the grouping based on the analysis of dominant patterns throughout the year (Figure 4 and Table 1). Regions (red area in Figure 7 corresponded to the heavy-tailed group in Table 1) characterized by uneven rainfall distribution throughout the year, leading to pronounced fluctuations between drier and wetter soil states, combined with higher evapotranspiration rates (indicated by warm to hot temperatures), tend to exhibit a dominance of heavy-tailed behavior in their hydrological responses across all seasons. In areas (white area in Figure 7 corresponded to the neutral group in Table 1) where rainfall is more evenly distributed annually, the emergence of heavy-tailed behavior is often linked to increased evapotranspiration during the growing seasons, particularly in spring and summer, and is less prominent during dormant seasons. This mechanism, which depends on evapotranspiration dynamics, substantiates the seasonality of flood tail behavior. Regions (blue area in Figure 7 corresponded to the nonheavy-tailed group in Table 1) where runoff generation is primarily influenced by snow dynamics tend to display linear hydrological responses. This is due to the fact that most runoff in these areas results from snowmelt during the growing seasons, driven by energy availability. These findings support the proposed mechanism of heavy-tailed flood behavior concluded in the spatial analyses and further demonstrate the pivotal effect played by the variation of evapotranspiration and catchment storage on the emergence of heavy-tailed flood behavior.

In summary, while heavy-/nonheavy-tailed behavior is generally consistent across seasons, there is a certain probability for cases to exhibit seasonality. This seasonality of inferred heavy-tailed behavior shows a dynamic pattern of increasing during the growing season and decreasing during the dormant season. Regions with pronounced temperature variations across seasons, particularly with hot summers, display such dynamics and highlight the role of evapotranspiration in catchments in driving this seasonality.

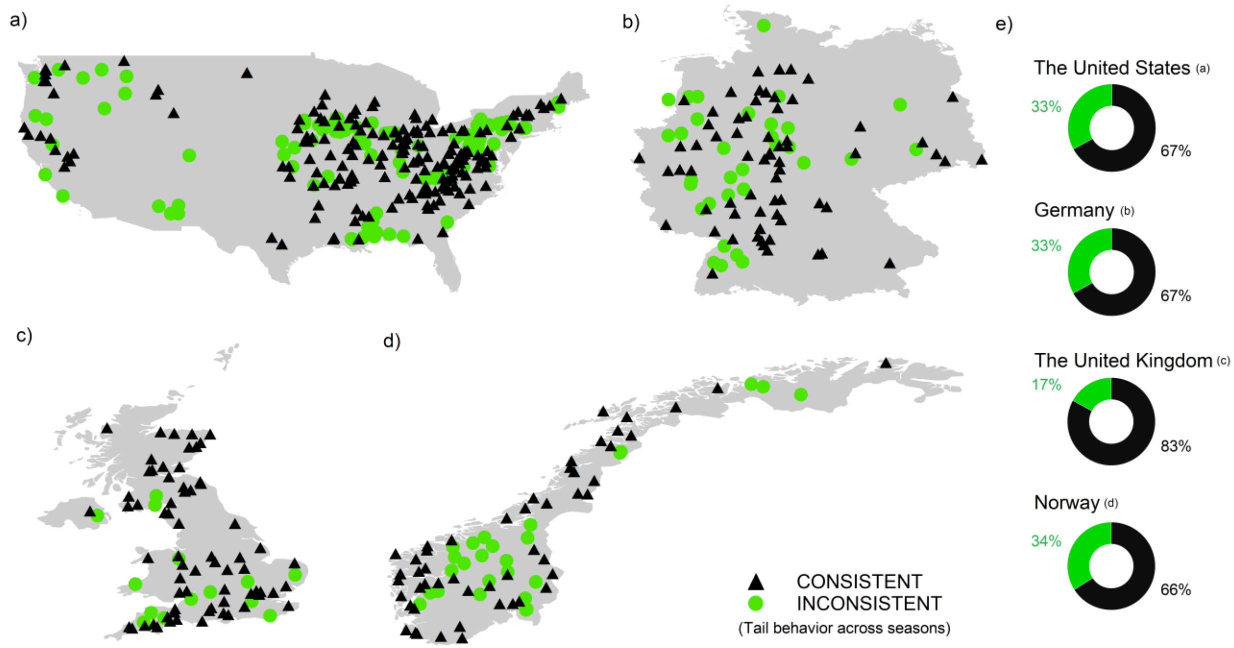


Figure 5. Consistency of inferred flood tail behavior across seasons. (a) 290 catchments in the US. (b) 98 catchments in Germany. (c) 81 catchments in the UK. (d) 79 catchments in Norway. (e) Percentage of consistent and inconsistent catchments in each country.

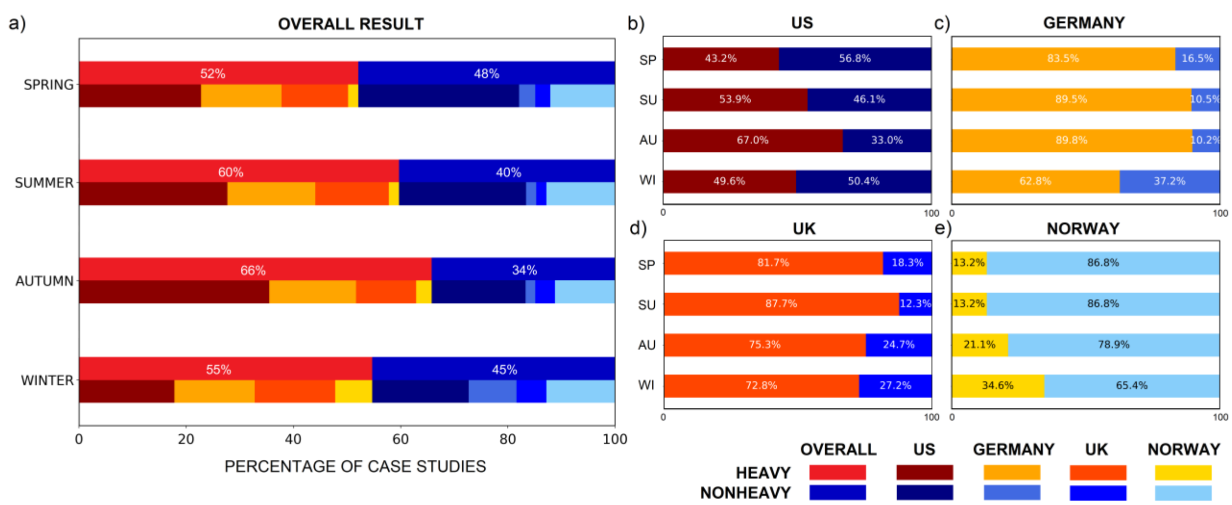
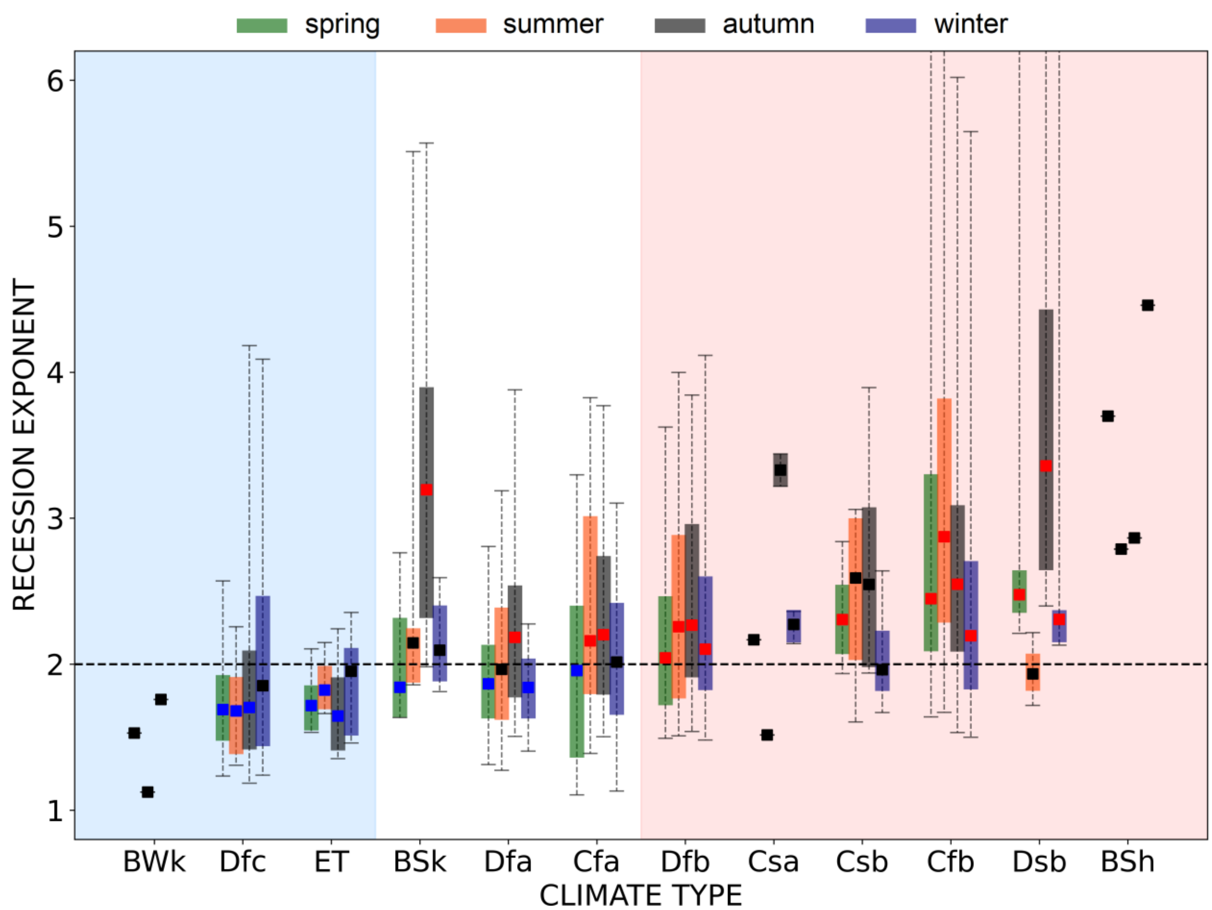


Figure 6. Seasonal variations in the percentage of inferred flood tail behavior between heavy and nonheavy case studies. (a) The aggregated results encompass all study regions, while the second line provides a breakdown by country. In total, there are 1,997 case studies composed by

581 540 in spring, 520 in summer, 543 in autumn, and 394 in winter. (b)-(e) Results for each study
582 country (see Table A1 for detailed case numbers across seasons in each country).



583

584 **Figure 7. Seasonal variations in recession exponents (inferring flood tail behavior) across**
585 **diverse climate regions.** Case studies grouped by climate regions based on seasons. Medians of
586 recession exponents in each group are compared with a value of two using Wilcoxon signed-rank
587 test (significance level: 0.05). Red squares indicate significantly heavy-tailed (recession exponents
588 > 2) groups, blue squares indicate significantly nonheavy-tailed (recession exponents < 2) groups,
589 and black squares denote insignificance. Climate regions are categorized as follows: the red area
590 denotes regions with prominent heavy tails across seasons, the blue area denotes regions with
591 prominent nonheavy tails across seasons, and the white area denotes regions with significant
592 seasonality in flood tail behavior.

593 4.4 Factors associated with catchment scales and their role in flood tail behavior

594 It remains unclear how flood tail behavior varies across catchment scales and what the underlying
595 drivers and mechanisms are (Merz et al., 2022). We employ catchment nonlinearity, represented
596 by recession exponents, to explore the influence of catchment scales on flood tail behavior, as
597 depicted in Figure 8. We utilize the categorization of regions characterized by distinct controls on

flood tail behavior, primarily influenced by characteristic runoff generation processes (as three groups identified in Figure 7), to elucidate the underlying mechanisms. Case studies are categorized into bins based on catchment areas, with the median values represented by squares, interquartile intervals depicted by vertical bars, and catchment area ranges indicated by horizontal dashed bars. Panels a, b, c, and d present results for all regions, regions exhibiting significant heavy tails across seasons, regions with a neutral propensity and seasonal variations, and regions displaying pronounced nonheavy tails across seasons, respectively. Each panel comprises a total of 30 bins, with approximately 67, 33, 24, and 10 case studies in panels a, b, c, and d, respectively (with minor variations due to rounding).

From the perspective of all case studies (Figure 8a), the pattern appears somewhat unclear. Apart from the case studies involving extremely small and large catchment areas, there seems to be a decrease in nonlinearity as catchment areas increase. Nevertheless, the relationship is rather weak and lacks clarity. These findings align with previous discussions on this matter (e.g., Merz and Blöschl, 2009; Villarini and Smith, 2010; Smith et al., 2018), which have suggested a relatively weak inverse correlation between catchment area and the occurrence of heavy-tailed flood behavior.

However, we can evidently clarify this relationship by considering the distinct runoff generation processes recognized in different regions. Panel b illustrates that catchment area plays no significant role in catchment nonlinearity in regions characterized by prominent heavy tails. Whereas a clear inverse relationship between nonlinearity and catchment area is shown in panel c, representing regions characterized by a neutral propensity for heavy and nonheavy tails. In contrast, a proportional relationship between nonlinearity and catchment area is identified in panel d, representing regions characterized by prominent nonheavy tails.

As shown by the previous sections, nonlinearity in neutral regions is primarily driven by high evapotranspiration facilitated by high temperatures. When the catchment area increases, it has a higher chance of encompassing diverse terrain types, including areas with higher altitudes, such as mountainous regions. Increased altitude tends to result in lower temperatures and evapotranspiration rates, negating the evapotranspiration variation and its impact on catchment nonlinearity, which is the main driver of nonlinearity in this region and thus substantiates an inverse relationship (Figure 8c). In regions with prominent heavy tails (Figure 8b), nonlinearity is generated from the interplay of uneven rainfall and evapotranspiration dynamics, and the enlargement of catchments does not substantively change this relationship. For regions with prominent nonheavy tails (Figure 8d), the underlying mechanisms are similar to the neutral regions but work in the opposite direction due to the differently dominant mechanism. Recall that the runoff process in this region is generally dominated by snow dynamics. The region is mainly located in high mountain or high latitude areas. As catchments expand, more diverse terrain is encompassed, potentially introducing a mixture of flood generation processes due to the incorporation of lowland or coastal areas. Particularly, more rain-on-snow events or rainfall-driven events may be encompassed in a same catchment together with snowmelt-driven events (Vormoor et al., 2016). Therefore, an increase in nonlinearity is facilitated due to the mixture of flood types, favoring the emergence of heavier tails in flood distributions (Tarasova et al., 2020). It should be noted that the tail patterns, based on Figure 8d, are still more likely to be nonheavy tails (i.e., recession exponents below two), even though nonlinearity indeed appears to show an increasing tendency along with catchment area.

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

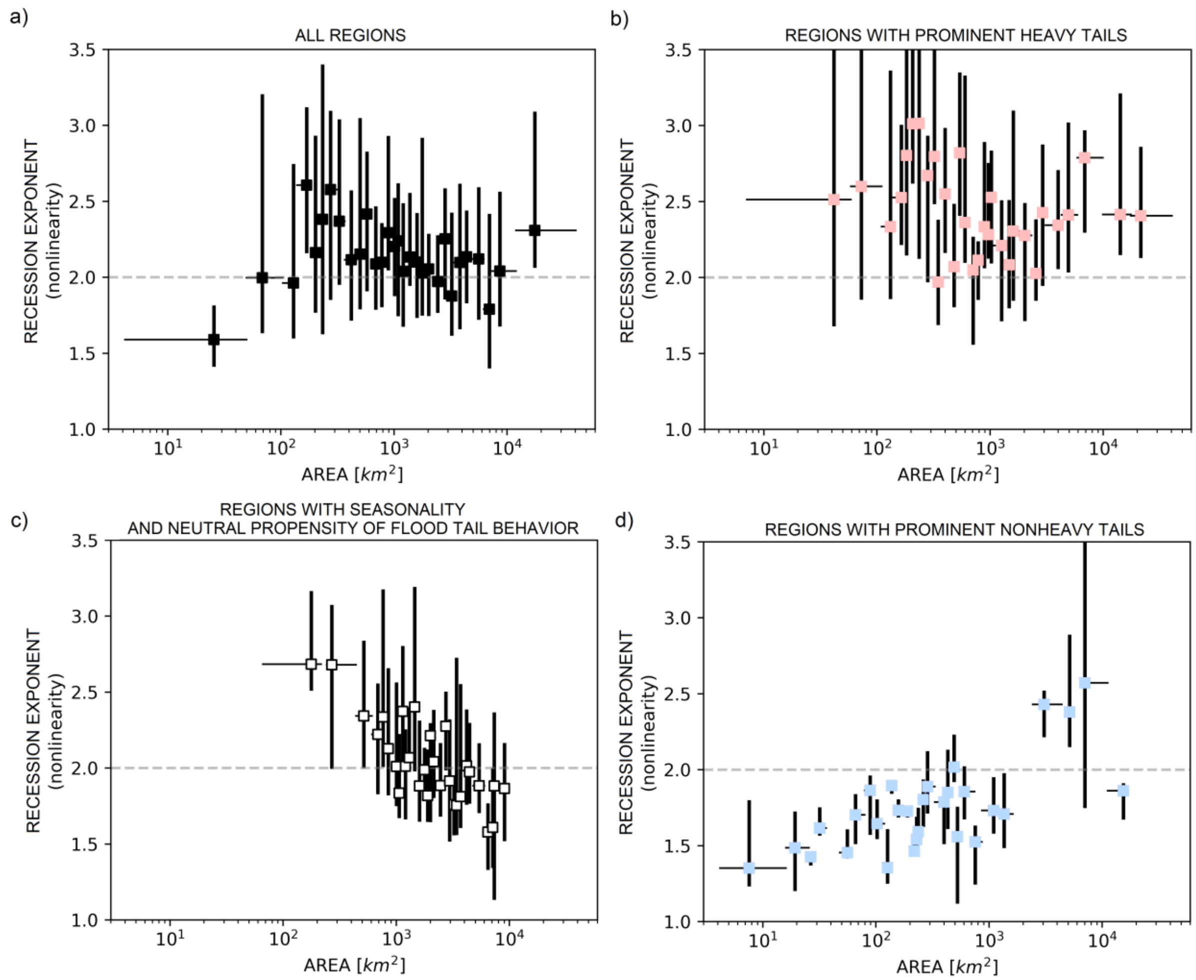


Figure 8. Catchment nonlinearity as a function of catchment area. The recession exponents, representing catchment nonlinearity, have been evenly grouped into bins based on catchment areas. The squares denote the median values, vertical bars represent the interquartile intervals of the recession exponents, and horizontal dashed bars indicate the catchment area ranges for each bin. (a) All regions (encompassing case studies, *n*=1997). (b)-(d) show case studies separately according to categorization recognized in Figure 7. (b) Regions with prominent heavy tails (*n*=978). (c) Regions with seasonality and neutral propensity of flood tail behavior (*n*=733). (d) Regions with prominent nonheavy tails (*n*=286). In each panel, there are a total of 30 bins, each containing

approximately 67, 33, 24, and 10 case studies in panels a, b, c, and d, respectively (with slight variations due to rounding).

5 Discussion

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

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

We further identify key drivers of heavy-tailed flood behavior by conducting large scale physioclimatic analyses. Specifically, our findings reveal that regions with a pronounced propensity for heavy tails exhibit distinct characteristics: the presence of a dry period and higher temperatures (as shown in Figure 4 and Table 1). This aligns with previous studies based on the mathematical analysis which associates heavier-tailed flood behavior with a lower frequency of streamflow-triggering rainfall events. Such lower frequency often results from erratic rainfall patterns and higher rates of evapotranspiration, leading to drier catchment conditions (Botter, 2010; Basso et al., 2016). In line with this theory, our large scale analysis provides evidence by showing a prevalent propensity for heavy tails in regions characterized by uneven rainfall patterns throughout the year (i.e., more erratic rainfall), contributing to the presence of dry periods, along with higher potential evapotranspiration rates, as indicated by higher temperatures.

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

The seasonality of flood tail behavior has been suggested in previous studies but remains less understood (Basso et al., 2015; Smith et al., 2018; Macdonald et al., 2022). It's noteworthy that more than one-third of catchments appear to exhibit inconsistent flood tail behavior across seasons (Figure 5). In these catchments, some seasons show a tendency toward nonheavy tails, while others tend to display heavy tails. Identifying these catchments and understanding the factors driving them to exhibit heavy tails is vital for hazard assessment. This understanding allows us to pinpoint catchments where extreme floods could potentially occur, even if methods solely based on annual maximum floods might estimate the flood tail as nonheavy based on annual maxima, when heavy tails can still occur within a single season. We have identified that regions characterized by stronger evapotranspiration dynamics across seasons favor this seasonality of flood tail behavior, as it leads to larger variations in water storage during higher evapotranspiration seasons, such as the growing seasons (highlighted in white in Figure 7). This finding aligns with previous studies that have observed similar seasonal dynamics in the nonlinearity of hydrological responses (Tashie et al., 2019; Tarasova et al., 2018).

In this study, we also found that the relationship between flood tail behavior and the expansion of catchment scales can be explained by changes in catchment nonlinearity, which are influenced by distinct flood generation processes. Previous studies have presented diverse perspectives on the relationship between flood tail behavior and catchment scales. While some studies have suggested that smaller catchments tend to exhibit heavier tails (e.g., Meigh et al., 1997; Pallard et al., 2009), others have noted a similar trend but with only a weak correlation (Merz and Blöschl, 2009; Villarini and Smith, 2010). Meanwhile, some studies have found no significant relationship between these two variables (Morrison and Smith, 2002; Smith et al., 2018). These studies have explored this topic without reaching a consensus, and many conclusions lack sufficient evidence and a clear understanding. In contrast, our findings (Figure 8) distinctly differentiate between various patterns by considering region classifications based on distinct dominant flood generation processes, thereby providing a mechanistic understanding. As a catchment expands, it encompasses more diverse terrain, which in turn facilitates a wider range of altitudes and flood types. In regions where tail behavior is primarily influenced by evapotranspiration dynamics (Figure 8c), the presence of diverse altitudes tends to moderate the effect of higher temperatures, reducing the influence of high evapotranspiration on the emergence of heavy tails. In regions where tail behavior is primarily controlled by snowmelt (Figure 8d) (mainly composed of catchments in Norway in this study), it has been shown that larger catchments are more likely to encompass a mix of flood types, including snowmelt-driven and rainfall-driven floods (Vormoor et al., 2016). Merz et al. (2022) suggested that heavier-tailed behavior in rainfall-driven floods tends to dominate in such mixed conditions. Our findings support this hypothesis by demonstrating an increase in tail heaviness as catchment area enlarges. In regions where heavy tails are pronounced due to the strong nonlinearity resulting from the interplay of uneven rainfall and high evapotranspiration, there is no significant relationship between catchment nonlinearity and catchment area (Figure 8b). This lack of relationship may be because the expansion of the catchment area does not appear to significantly enhance or reduce this interplay.

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

"Hypothesis 2 (of the review paper): The Characteristic Flood Generation Process Shapes the Upper Flood Tail Catchments." While previous studies have hinted at the possibility that regions where flood generation is dominated by rainfall-driven floods tend to exhibit heavier-tailed flood behavior compared to regions dominated by snowmelt (Bernardara et al., 2008; Thorarinsdottir et al., 2018), more explicit process explanations are desired. In line with this hypothesis, we present further evidence showing that the specific nonlinearity inherent in each flood generation process is the primary driver of flood tail behavior. Specifically, we show that in snowmelt-dominated regions, such as the case studies in Norway, hydrological responses closely resemble linear behavior and thus floods tend to exhibit pronounced nonheavy-tailed behavior. Conversely, heavy-tailed floods are more prominent in regions like the UK, where hydrological responses display nonlinearity (as indicated by recession exponents above two). In these areas, flood generation processes are primarily driven by rainfall events. Furthermore, our study reveals that flood generation processes are significantly influenced by the interplay between regional terrain and meteorological features. These factors, in turn, impact the nonlinearity of hydrological responses and can determine the occurrence of heavy or nonheavy tails in flood distributions (Figure 3e).

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

"Hypothesis 4: Non-Linear Response to Precipitation Causes Heavy Flood Tails." Studies have consistently highlighted the significance of nonlinearity in hydrological processes within catchments as a key determinant in the emergence of heavy-tailed flood behavior (e.g., Struthers and Sivapalan, 2007; Rogger et al., 2012; Basso et al., 2015). In our research, we contribute by introducing a quantitative approach that employs hydrograph recession exponents as a measure of nonlinearity in flood tail analyses and validate its effectiveness in identifying heavy-tailed flood behavior in a large scale analysis. While nonlinearity has long been acknowledged as a contributing factor, our works uniquely utilizes this driver as a reliable index by establishing a specific recession exponent threshold that robustly discriminates heavy-tailed distributions, characterized by power-law tails, from nonheavy ones, offering a valuable tool to the field.

Furthermore, our large scale analysis identifies rainfall unevenness and high temperatures as crucial drivers behind the observed nonlinearity in flood responses. Specifically, they intensify catchment soil dryness and amplify water balance storage variations, thereby facilitating both very small and very large runoff events, translating into heavy-tailed flood behavior.

"Hypothesis 5: Drier Catchments Have Heavier Flood Tails Due To Interaction of Water Balance Processes." In alignment with previous studies that suggest the water balance processes in drier catchments contribute to the emergence of heavy-tailed flood behavior (e.g., Molnar et al., 2006; Merz and Blöschl, 2009; Guo et al., 2014), we emphasize the critical interplay between uneven rainfall and evapotranspiration dynamics in facilitating these processes and shaping such the behavior. Specifically, our findings show that heavy-tailed flood behavior is more likely to occur in catchments characterized by lower rainfall and higher evapotranspiration in one season (e.g., summer), contrasted with more rainfall and lower evapotranspiration in another season (e.g., winter). When one of these conditions is lacking, heavy-tailed behavior may be less pronounced. For example, regions classified as BSh and BSk, both of which exhibit semi-arid characteristics based on their rainfall patterns, exhibit differences in the prevalence of heavy-tailed cases. BSk regions, despite their semi-arid status, exhibit fewer pronounced heavy-tailed cases due to colder temperatures (Table 1) and only show a higher rate of heavy-tailed cases during the summer (Figure 7). This interplay highlights the importance of considering the seasonality of flood tail behavior, particularly in regions that do not experience significant dry periods based on their rainfall patterns. In such regions, heavy tails are still likely to occur in seasons with higher evapotranspiration rates (indicated by the white area in Figure 7).

"Hypothesis 6: Smaller Catchments Have Heavier Flood Tails Due To Less Pronounced Spatial Aggregation Effects." A commonly debated question among hydrologists is whether the roles identified in large catchments are applicable to smaller ones, and vice versa. This issue has also arisen in discussions regarding flood tail heaviness, but evidence on the matter has been scattered. While smaller catchments have been suggested to exhibit heavier tails (Meigh et al., 1997; Pallard et al., 2009), previous research has revealed weak (Merz and Blöschl, 2009; Villarini and Smith, 2010) to no (Morrison and Smith, 2002; Smith et al., 2018) correlations between catchment size and tail heaviness. Our findings (Figure 8) help clarify the relationship between catchment nonlinearity (used as an indicator of tail heaviness) and catchment sizes. We observe distinct patterns among regions characterized by strong, neutral, and weak conditions of heavy tail behavior. These findings underscore the importance of considering the dominant flood generation processes in each region and elucidate how catchment size interacts with flood tail behavior by influencing these dominant processes—either amplifying, reducing, or having no significant effect.

6 Conclusions

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

1. **Capability of Recession Exponents for Detecting Heavy-Tailed Flood Behavior:** The capability of this index to discern between case studies which display heavy-tailed flood

distributions and those exhibit nonheavy-tailed behavior is validated by using empirical data from catchments across Germany, Norway, the UK, and the US. This extensive analysis provides a well-rounded evaluation due to the inclusion of regions with divergent conditions, such as rainfall-driven floods (Germany, the UK, and the US) versus snowmelt-driven floods (Norway), as well as regions characterized by single/dominant hydroclimates (Germany, the UK, and Norway) versus those with mixed hydroclimates (the US).

2. **Regional Propensity for Heavy-Tailed Flood Behavior:** Germany and the UK are characterized by a propensity for heavy-tailed flood behavior, which is prevalent in these regions. Conversely, a tendency for nonheavy-tailed flood behavior is predominant in Norway under current hydroclimatic conditions, as indicated by the degree of catchment nonlinearity in each region. The US exhibits a mixture of heavy- and nonheavy-tailed behavior. This is likely the results of overarching climatic characteristics, which also shape river network morphology, interacting with diverse regional physioclimatic settings. We emphasize that the relatively more uniform climates in Germany, the UK, and Norway contribute to a dominant presence of heavy or nonheavy-tailed behaviorsin these countries, while the US experiences more complex regional patterns due to more diverse hydroclimatic conditions.
3. **Factors Influencing Heavy-Tailed Flood Behavior:** The presence of simultaneous dry periods (defined by uneven rainfall throughout the year) and higher temperatures emerge as the pivotal conditions favoring heavy-tailed flood behavior. Drier catchments alter the runoff generation process, resulting in higher nonlinearity of catchment responses, while higher temperatures elevate evapotranspiration rates, enhancing nonlinearity but also maintaining atmospheric moisture preventing precipitation limitations. The absence of either condition diminishes the prevalence of heavy-tailed flood behavior. More generalized climate categorizations like Arid, Temperate, and Continental exhibit minimal influence on our results.
4. **Seasonality of Flood Tail Behavior:** We contribute to a better understanding of the seasonality of flood tail behavior. Around two-thirds of catchments exhibit consistent behavior across seasons, with the remaining one-third demonstrating seasonality. Heavy-tailed flood behavior is more likely during the growing season (spring to autumn) and diminishes during the dormant season (autumn to winter). These findings hint at the role of temperature-driven evapotranspiration dynamics for the emergence of heavy-tailed flood behavior, which are particularly important in regions which do not experience simultaneous dry conditions and high temperatures.
5. **Influences of Catchment Area on Flood Tail Behavior:** We elucidate that the impacts of catchment size on flood tail behavior are primarily contingent on the dominant flood generation processes within each region. Specifically, the expansion of catchment area tends to have three distinct effects: (1) It diminishes tail heaviness in regions with moderate nonlinearity, characterized by strong evapotranspiration dynamics and relatively even rainfall throughout the year. This reduction is attributed to the smoothing of evapotranspiration variations. (2) Conversely, in regions with low nonlinearity, characterized by snowfall dynamics, increasing catchment area intensifies tail heaviness. This effect results from the inclusion of diverse flood types, particularly rainfall-driven floods. (3) In regions with high nonlinearity, characterized by simultaneous strong evapotranspiration dynamics and uneven rainfall throughout the year, catchment size expansion appears to have no significant impact on tail heaviness. This lack of effect is likely due to the absence of significant influence on rainfall patterns, which are critical in determining the presence of drier soil conditions.

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

Appendix A Information on Study Regions

Table A1. Daily Hydrological Data Information

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

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Project number 421396820 "Propensity of rivers to extreme floods: climate-landscape controls and early detection (PREDICTED)" and Research Group FOR 2416 "Space-Time Dynamics of Extreme Floods (SPATE)". We also acknowledge the financial support provided by the Helmholtz Centre for Environmental Research - UFZ. The manuscript and supporting information contain all the necessary details to replicate the results.

Data Availability Statement

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

References

Arai, R., Toyoda, Y., & Kazama, S. (2020). Runoff recession features in an analytical probabilistic streamflow model. Journal of Hydrology, 597, 125745. https://doi.org/10.1016/j.jhydrol.2020.125745

Basso, S., Botter, G., Merz, R., & Miniussi, A. (2021). PHEV! The PHysically-based Extreme Value distribution of river flows. Environmental Research Letters, 16(12). https://doi.org/10.1088/1748-9326/ac3d59

Basso, S., Schirmer, M., & Botter, G. (2015). On the emergence of heavy-tailed streamflow distributions. Advances in Water Resources, 82, 98–105. https://doi.org/10.1016/j.advwatres.2015.04.013

Basso, S., Schirmer, M., & Botter, G. (2016). A physically based analytical model of flood frequency curves. Geophysical Research Letters, 43(17), 9070–9076. https://doi.org/10.1002/2016GL069915

Basso, S., Merz, R., Tarasova, L., & Miniussi, A. (2023). Extreme flooding controlled by stream network organization and flow regime. Nature Geoscience, 16(April), 339–343. https://doi.org/10.1038/s41561-023-01155-w

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data, 5, 180214. https://doi.org/10.1038/sdata.2018.214

Bernardara, P., Schertzer, D., Sauquet, E., Tchiguirinskaia, I., & Lang, M. (2008). The flood probability distribution tail: How heavy is it? Stochastic Environmental Research and Risk Assessment, 22(1), 107–122. https://doi.org/10.1007/s00477-006-0101-2

Bevere, L., & Remondi, F. (2022). Natural catastrophes in 2021: the floodgates are open. Swiss Re Institute sigma research.

Biswal, B. (2021). Decorrelation is not dissociation: There is no means to entirely decouple the Brutsaert-Nieber parameters in streamflow recession analysis. Advances in Water Resources, 147, 103822. https://doi.org/https://doi.org/10.1016/j.advwatres.2020.103822

Biswal, B., & Marani, M. (2010). Geomorphological origin of recession curves. Geophysical Research Letters, 37(24), 1–5. https://doi.org/10.1029/2010GL045415

Botter, G. (2010). Stochastic recession rates and the probabilistic structure of stream flows. Water Resources Research, 46(12). https://doi.org/10.1029/2010WR009217

Botter, G., Basso, S., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). Natural streamflow regime alterations: Damming of the Piave river basin (Italy). Water Resources Research, 46(6), 1–14. https://doi.org/10.1029/2009WR008523

Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007). Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. climate regimes. Water Resources Research, 43(11), 1–10. https://doi.org/10.1029/2007WR006162

Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. Water Resources Research, 43(2), 1–14. https://doi.org/10.1029/2006WR005043

Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2009). Nonlinear storage-discharge relations and catchment streamflow regimes. Water Resources Research, 45(10), 1–16. https://doi.org/10.1029/2008WR007658

Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resources Research, 13(3), 637–643. https://doi.org/10.1029/WR013i003p00637

Ceola, S., Botter, G., Bertuzzo, E., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2010). Comparative study of ecohydrological streamflow probability distributions. Water Resources Research, 46(9), 1–12. https://doi.org/10.1029/2010WR009102

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. SIAM Review, 51(4), 661–703. https://doi.org/10.1137/070710111

Cunderlik, J. M., & Burn, D. H. (2002). Utilisation d’une information sur le régime des crues dans une analyse fréquentielle régionale des crues. Hydrological Sciences Journal, 47(1), 77–92. https://doi.org/10.1080/02626660209492909

Derrick, B., Toher, D., & White, P. (2016). Why Welchs test is Type I error robust. The Quantitative Methods for Psychology, 12(1), 30–38. https://doi.org/doi:10.20982/tqmp.12.1.p030

Doulatyari, B., Betterle, A., Basso, S., Biswal, B., Schirmer, M., & Botter, G. (2015). Predicting streamflow distributions and flow duration curves from landscape and climate. Advances in Water Resources, 83, 285–298. https://doi.org/10.1016/j.advwatres.2015.06.013

Durrans, S. R., Eiffe, M. A., Thomas, W. O., & Goranflo, H. M. (2003). Joint Seasonal /Annual Flood Frequency Analysis. Journal of Hydrologic Engineering, 8(4), 181–189. https://doi.org/10.1061/(asce)1084-0699(2003)8:4(181)

El Adlouni, S., Bobée, B., & Ouarda, T. B. M. J. (2008). On the tails of extreme event distributions in hydrology. Journal of Hydrology, 355(1–4), 16–33. https://doi.org/10.1016/j.jhydrol.2008.02.011

Fiorentino, M., Manfreda, S., & Iacobellis, V. (2007). Peak runoff contributing area as hydrological signature of the probability distribution of floods. Advances in Water Resources, 30(10), 2123–2134. https://doi.org/10.1016/j.advwatres.2006.11.017

Fischer, S., & Schumann, A. (2016). Robust flood statistics: comparison of peak over threshold approaches based on monthly maxima and TL-moments. Hydrological Sciences Journal, 61(3), 457–470. https://doi.org/10.1080/02626667.2015.1054391

Gaume, E. (2006). On the asymptotic behavior of flood peak distributions. Hydrology and Earth System Sciences, 10(2), 233–243. https://doi.org/10.5194/hess-10-233-2006

Gioia, A., Iacobellis, V., Manfreda, S., & Fiorentino, M. (2008). Runoff thresholds in derived flood frequency distributions. Hydrology and Earth System Sciences, 12(6), 1295–1307. https://doi.org/10.5194/hess-12-1295-2008

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

Huntingford, C., Marsh, T., Scaife, A. A., Kendon, E. J., Hannaford, J., Kay, A. L., et al. (2014). Potential influences on the United Kingdom's floods of winter 2013/14. Nature Climate Change, 4(9), 769–777. https://doi.org/10.1038/nclimate2314

Jachens, E. R., Rupp, D. E., Roques, C., & Selker, J. S. (2020). Recession analysis revisited: Impacts of climate on parameter estimation. Hydrology and Earth System Sciences, 24(3), 1159–1170. https://doi.org/10.5194/hess-24-1159-2020

Katz, R. (2002). Statistics of Extremes in Climatology and Hydrology. Advances in Water Resources, 25, 1287–1304.

Kondor, D., Pósfai, M., Csabai, I., & Vattay, G. (2014). Do the rich get richer? An empirical analysis of the Bitcoin transaction network. PLoS ONE, 9(2). https://doi.org/10.1371/journal.pone.0086197

Laio, F., Porporato, A., Fernandez-Illescas, C. P., & Rodriguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and responce to water stress IV. Discussion of real cases. Advances in Water Resources, 24(7), 745–762. https://doi.org/10.1016/S0309-1708(01)00007-0

Lehner, B., Liermann, C. R., Revenga, C., Vörömsmarty, C., Fekete, B., Crouzet, P., et al. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment, 9(9), 494–502. https://doi.org/10.1890/100125

Lins, H. F. (2008). Challenges to hydrological observations. WMO Bulletin, 57(January), 55–58.

Lu, P., Smith, J. A., & Lin, N. (2017). Spatial characterization of flood magnitudes over the drainage network of the Delaware river basin. Journal of Hydrometeorology, 18(4), 957–976. https://doi.org/10.1175/JHM-D-16-0071.1

Macdonald, E., Merz, B., Guse, B., Wietzke, L., Ullrich, S., Kemter, M., et al. (2022). Event and Catchment Controls of Heavy Tail Behavior of Floods. Water Resources Research, 58(6), 1–25. https://doi.org/10.1029/2021wr031260

Malamud, B. D. (2004). Tails of natural hazards. Physics World, 17(8), 31–35. https://doi.org/10.1088/2058-7058/17/8/35

Malamud, B. D., & Turcotte, D. L. (2006). The applicability of power-law frequency statistics to floods. Journal of Hydrology, 322(1–4), 168–180. https://doi.org/10.1016/j.jhydrol.2005.02.032

McCuen, R. H., & Smith, E. (2008). Origin of Flood Skew. Journal of Hydrologic Engineering, 13(9), 771–775. https://doi.org/10.1061/(asce)1084-0699(2008)13:9(771)

McDermott, T. K. J. (2022). Global exposure to flood risk and poverty. Nature Communications, 13(1), 6–8. https://doi.org/10.1038/s41467-022-30725-6

Meigh, J. R., Farquharson, F. A. K., & Sutcliffe, J. V. (1997). A worldwide comparison of regional flood estimation methods and climate. Hydrological Sciences Journal, 42(2), 225–244. https://doi.org/10.1080/02626669709492022

Mejía, A., Daly, E., Rossel, F., Javanovic, T., & Gironás, J. (2014). A stochastic model of streamflow for urbanized basins. Water Resources Research, 50, 1984–2001. https://doi.org/10.1002/2013WR014834

Merz, B., Basso, S., Fischer, S., Lun, D., Blöschl, G., Merz, R., et al. (2022). Understanding heavy tails of flood peak distributions. Water Resources Research, 1–37. https://doi.org/10.1029/2021wr030506

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. Nature Reviews Earth and Environment, 2(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3

Merz, R., & Blöschl, G. (2009). Process controls on the statistical flood moments - a data based analysis. Hydrological Processes, 23(5), 675–696. https://doi.org/10.1002/hyp

Molnar, P., Anderson, R. S., Kier, G., & Rose, J. (2006). Relationships among probability distributions of stream discharges in floods, climate, bed load transport, and river incision. Journal of Geophysical Research: Earth Surface, 111(2), 1–10. https://doi.org/10.1029/2005JF000310

Morrison, J. E., & Smith, J. A. (2002). Stochastic modeling of flood peaks using the generalized extreme value distribution. Water Resources Research, 38(12), 41-1-41–12. https://doi.org/10.1029/2001wr000502

Müller, M. F., Dralle, D. N., & Thompson, S. E. (2014). Analytical model for flow duration curves in seasonally dry climates. Water Resources Research, 50, 5510–5531. https://doi.org/10.1002/2014WR015301

Müller, M. F., Roche, K. R., & Dralle, D. N. (2021). Catchment processes can amplify the effect of increasing rainfall variability. Environmental Research Letters, 16(8). https://doi.org/10.1088/1748-9326/ac153e

Mushtaq, S., Miniussi, A., Merz, R., & Basso, S. (2022). Reliable estimation of high floods: A method to select the most suitable ordinary distribution in the Metastatistical extreme value framework. Advances in Water Resources, 161(September 2021), 104127. https://doi.org/10.1016/j.advwatres.2022.104127

Pallard, B., Castellarin, A., & Montanari, A. (2009). A look at the links between drainage density and flood statistics. Hydrol. Earth Syst. Sci (Vol. 13). Retrieved from www.hydrol-earth-syst-sci.net/13/1019/2009/

Papalexiou, S. M., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? An assessment of daily rainfall distribution tails. Hydrology and Earth System Sciences, 17(2), 851–862. https://doi.org/10.5194/hess-17-851-2013

Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. American Naturalist, 164(5), 625–632. https://doi.org/10.1086/424970

Pumo, D., Viola, F., La Loggia, G., & Noto, L. V. (2014). Annual flow duration curves assessment in ephemeral small basins. Journal of Hydrology, 519(PA), 258–270. https://doi.org/10.1016/j.jhydrol.2014.07.024

Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., & Merz, R. (2012). Step changes in the flood frequency curve : Process controls. Water Resources Research, 48, 1–15. https://doi.org/10.1029/2011WR011187

Santos, A. C., Portela, M. M., Rinaldo, A., & Schaefli, B. (2018). Analytical flow duration curves for summer streamflow in Switzerland. Hydrology and Earth System Sciences, 22(4), 2377–2389. https://doi.org/10.5194/hess-22-2377-2018

Sartori, M., & Schiavo, S. (2015). Connected we stand: A network perspective on trade and global food security. Food Policy, 57, 114–127. https://doi.org/https://doi.org/10.1016/j.foodpol.2015.10.004

Schaefli, B., Rinaldo, A., & Botter, G. (2013). Analytic probability distributions for snow-dominated streamflow. Water Resources Research, 49(5), 2701–2713. https://doi.org/10.1002/wrcr.20234

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

Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L., & Bates, P. (2018). Strange Floods: The Upper Tail of Flood Peaks in the United States. Water Resources Research, 54(9), 6510–6542. https://doi.org/10.1029/2018WR022539

Spearman, C. (1904). The proof and measurement of association between two things. American Journal of Psychology, 15(1), 72–101. https://doi.org/10.2307/1412159

Struthers, I., & Sivapalan, M. (2007). A conceptual investigation of process controls upon flood frequency: Role of thresholds. Hydrology and Earth System Sciences, 11(4), 1405–1416. https://doi.org/10.5194/hess-11-1405-2007

Tarasova, L., Basso, S., & Merz, R. (2020). Transformation of Generation Processes From Small Runoff Events to Large Floods. Geophysical Research Letters, 47(22). https://doi.org/10.1029/2020GL090547

Tarasova, L., Basso, S., Zink, M., & Merz, R. (2018). Exploring Controls on Rainfall-Runoff Events: 1. Time Series-Based Event Separation and Temporal Dynamics of Event Runoff Response in Germany. Water Resources Research, 54(10), 7711–7732. https://doi.org/10.1029/2018WR022587

Tarasova, L., Lun, D., Merz, R., Blöschl, G., Basso, S., Bertola, M., et al. (2023). Shifts in flood generation processes exacerbate regional flood anomalies in Europe. Communications Earth and Environment, 4(1), 1–12. https://doi.org/10.1038/s43247-023-00714-8

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

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

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

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

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

Villarini, G., Smith, J. A., Baeck, M. L., Marchok, T., & Vecchi, G. A. (2011). Characterization of rainfall distribution and flooding associated with U.S. landfalling tropical cyclones: Analyses of Hurricanes Frances, Ivan, and Jeanne (2004). Journal of Geophysical Research Atmospheres, 116(23). https://doi.org/10.1029/2011JD016175

Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D., & Wong, W. K. (2016). Evidence for changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway. Journal of Hydrology, 538, 33–48. https://doi.org/10.1016/j.jhydrol.2016.03.066

Wang, H.-J., Merz, R., Yang, S., & Basso, S. (2023). Inferring heavy tails of flood distributions from common discharge dynamics. EGUsphere, 2023(May), 1–24. https://doi.org/https://doi.org/10.5194/egusphere-2023-660

Wang, H., Merz, R., Yang, S., Tarasova, L., & Basso, S. (2022). Emergence of heavy tails in streamflow distributions: the role of spatial rainfall variability. Advances in Water Resources Journal, 171(104359). https://doi.org/10.1016/j.advwatres.2022.104359

Ward, A. S., Wondzell, S. M., Schmadel, N. M., & Herzog, S. P. (2020). Climate Change Causes River Network Contraction and Disconnection in the H.J. Andrews Experimental Forest, Oregon, USA. Frontiers in Water, 2(April), 1–10. https://doi.org/10.3389/frwa.2020.00007

Welch, B. L. (1947). The generalization of “Student’s” problem when several different population variances are involved. Biometrika, 34(1–2), 28–35. https://doi.org/doi:10.1093/biomet/34.1-2.28. MR 0019277. PMID 20287819

Wietzke, L. M., Merz, B., Gerlitz, L., Kreibich, H., Guse, B., Castellarin, A., & Vorogushyn, S. (2020). Comparative analysis of scalar upper tail indicators. Hydrological Sciences Journal, 65(10), 1625–1639. https://doi.org/10.1080/02626667.2020.1769104

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics Bulletin, 1(6), 80–83. https://doi.org/10.2307/3001968

Wu, Q., Ke, L., Wang, J., Pavelsky, T. M., Allen, G. H., Sheng, Y., et al. (2023). Satellites reveal hotspots of global river extent change. Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-37061-3