

Geophysical Research Letters[®]



RESEARCH LETTER

10.1029/2023GL105429

Prediction of Extraordinarily High Floods Emerging From Heterogeneous Flow Generation Processes

Sumra Mushtaq¹ , Arianna Miniussi^{1,2} , Ralf Merz¹ , Larisa Tarasova¹ , Francesco Marra^{3,4}, and Stefano Basso^{1,5}

¹Department Catchment Hydrology, Helmholtz Centre for Environmental Research—UFZ, Halle (Saale), Germany, ²P&C International, General Reinsurance, Cologne, Germany, ³Department of Geosciences, University of Padova, Padova, Italy, ⁴Institute of Atmospheric Sciences and Climate, National Research Council (CNR-ISAC), Bologna, Italy, ⁵Norwegian Institute for Water Research (NIVA), Oslo, Norway

Key Points:

- We develop a statistical model to predict extraordinarily high floods occurring in river basins with variable flood extremes
- The approach leverages knowledge of various hydrologic processes generating runoff events to estimate the magnitude of extreme floods
- By accounting for multiple processes in flood magnitude-frequency analysis, we improve the estimation of extraordinarily high floods

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. Mushtaq,
sumra.mushtaq@ufz.de

Citation:

Mushtaq, S., Miniussi, A., Merz, R., Tarasova, L., Marra, F., & Basso, S. (2023). Prediction of extraordinarily high floods emerging from heterogeneous flow generation processes. *Geophysical Research Letters*, 50, e2023GL105429. <https://doi.org/10.1029/2023GL105429>

Received 17 JUL 2023

Accepted 27 SEP 2023

Author Contributions:

Conceptualization: Sumra Mushtaq, Arianna Miniussi, Ralf Merz, Larisa Tarasova, Francesco Marra, Stefano Basso

Data curation: Larisa Tarasova

Formal analysis: Sumra Mushtaq

Funding acquisition: Ralf Merz, Stefano Basso

Methodology: Sumra Mushtaq, Arianna Miniussi, Larisa Tarasova, Francesco Marra, Stefano Basso

Project Administration: Ralf Merz, Stefano Basso

Software: Sumra Mushtaq

Abstract River floods are generated by various processes that, if disregarded, may induce errors in flood hazard assessment. This is particularly relevant where events extraordinarily larger than the typical floods have been observed, that is, for rivers with a flood divide in their flood-frequency curves. We identify 11 such cases in a large set of German catchments and test a statistical approach that accounts for different runoff-generation processes to predict the magnitude and frequency of extraordinarily high floods. We observe that in catchments with a flood divide, ordinary peaks are generated by different runoff-generation processes and the distribution of at least one process is heavy-tailed. By accounting for the different tail behaviors of multiple processes, we can reproduce flood-frequency curves in these catchments. Our findings shed light on the origin of flood divides and set a method to improve the estimation of high flood quantiles in these high-risk cases.

Plain Language Summary River floods can be caused by many different processes. Neglecting this fact may impair our capability to predict the probability of occurrence of the most extreme floods. In this study, we combine knowledge of runoff-generation mechanisms across different events with a statistical method which uses information on these ordinary events to infer the magnitude and frequency of floods. We apply the method to 11 rivers in Germany where extraordinarily high floods have been observed. By accounting for ordinary runoff events and their diverse runoff-generation processes we obtain accurate estimates of the magnitude of rare and extreme floods, substantially improving results provided by traditional methods.

1. Introduction

Reliable estimates of the magnitude and frequency of river floods are crucial for societal activities such as engineering design, urban planning and risk management (Barth et al., 2017; Smith et al., 2018). In particular, flood magnitude-frequency analyses establish a functional relation between the magnitude of floods and their exceedance probability, and represent a basic tool for quantifying flood hazard (e.g., Cunnane, 1973; Gotvald et al., 2012). However, river floods remain common perils that cause fatalities and damages worldwide (Bevere & Remondi, 2021; François et al., 2019).

Accurate estimation of flood magnitude and frequency is especially challenging when a few events are significantly larger in magnitude than the rest of the observed floods (Smith et al., 2018). These unexpectedly high floods can result in large socio-economic impacts (Davenport et al., 2021; Kreibich et al., 2022; Merz et al., 2021). This phenomenon, displayed in flood-frequency curves as a marked increase in the magnitude of rare events, was recently termed as flood divide (Basso et al., 2023). This behavior is pervasive, although difficult to detect due to limited data availability. Past studies show that the number of flood divides increases with the length of the data series (e.g., Miniussi et al., 2023; Smith et al., 2018). The appearance of flood divides has been often related to data scarcity preventing a good characterization of the tail of flood distributions (e.g., Miniussi et al., 2023), an issue possibly exacerbated by non-stationarity of hydrologic processes caused by climate change (Yu et al., 2022). However, previous studies also linked them to non-linearities in the catchment response (Roger et al., 2012) arising from distinct physioclimatic characteristics of river basins (Basso et al., 2023) and to the existence of different runoff-generation processes in a catchment (Merz et al., 2022). Runoff events and floods may indeed be triggered by different processes (e.g., rainfall on wet or dry soils, snowmelt, a combination of both; Hirschboeck, 1987b;

© 2023. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Supervision: Arianna Miniussi, Ralf Merz, Stefano Basso

Validation: Sumra Mushtaq, Arianna Miniussi, Ralf Merz, Larisa Tarasova, Francesco Marra, Stefano Basso

Visualization: Sumra Mushtaq, Arianna Miniussi, Ralf Merz, Larisa Tarasova, Francesco Marra, Stefano Basso

Writing – original draft: Sumra Mushtaq, Arianna Miniussi, Larisa Tarasova, Francesco Marra, Stefano Basso

Writing – review & editing: Sumra Mushtaq, Arianna Miniussi, Ralf Merz, Larisa Tarasova, Francesco Marra, Stefano Basso

Villarini & Smith, 2010; Sikorska et al., 2015) and thus be categorized into different types based on the inducing mechanisms (Stein et al., 2020; Tarasova, Basso, & Merz, 2020).

Traditional flood magnitude-frequency analyses (e.g., Sivapalan & Samuel, 2009) are unsuitable to handle heterogeneous flood records (Hirschboeck, 1987a), especially when floods differ among each others to the point that flood divides are observed. Mixed extreme value distributions have been proposed for these cases, but the limited availability of large flood observations introduces considerable estimation uncertainty (Alila & Mtraoui, 2002; Barth et al., 2019; Waylen & Woo, 1982). To tackle this issue, past studies (e.g., Fischer, 2018; Hirschboeck, 1987b) proposed a mixed distribution approach using peaks over a threshold (i.e., by using all observed peaks above a threshold).

These approaches improved hazard estimation compared to methods which assume homogeneity of the flood record. However, peaks over threshold may not be a representative sample of the processes that generate extraordinary events in catchments where the generation processes of the largest and most common floods are strikingly different (Tarasova, Basso, & Merz, 2020). In these cases, using ordinary peaks (i.e., all the independent streamflow peaks) may allow for sampling the full variety of generation processes while constraining prediction uncertainty. Recently, Miniussi et al. (2020) adopted such an approach to model floods occurring during different ENSO phases. However, studies that consider specific runoff-generation processes which produced ordinary peaks to estimate flood frequencies and magnitudes are lacking.

In this study, we aim to fill this gap by addressing the issue of reliably estimating flood frequencies and magnitudes in basins with a flood divide in the empirical flood magnitude-frequency curve. To do so, we rely on the analysis of ordinary peaks by explicitly considering the presence of multiple runoff-generation processes in a mixed non-asymptotic extreme value model. The advantage of this approach over traditional methods lies in the extension of information obtained by using a larger sample of ordinary peaks and accounting for the underlying physical processes leading to runoff generation in flood frequency analysis.

2. Methodology

2.1. Study Area and Data

This study uses daily streamflow records of 169 mesoscale river basins in Germany with drainage area between 30 and 23,000 km² (median: 581 km²; Figure S1 in Supporting Information S1). The analyzed data series range from 1951 to 2013, encompassing 36–62 (median: 59) hydrological years per basin (November to October). From this data set, we identified river basins exhibiting a sharp increase in the magnitude of rare floods (i.e., a flood divide in their empirical flood magnitude-frequency curve, *sensu* Miniussi et al., 2023; Basso et al., 2023) by visually examining empirical flood-frequency curves. In order to facilitate their identification from a large data set, we calculated the L-skewness of the annual maxima sample (Hosking, 1990), which is a general and universal index for tail heaviness and robust to the presence of outliers (Vogel & Fennessey, 1993), we noticed that, for Germany, all the cases with L-skewness exceeding 0.3 were also presenting this issue (Figure S2 in Supporting Information S1). Notably, the threshold of 0.3 is selected for German basins and may vary for other regions. Visual inspection of the empirical flood-frequency curves remains the main tool to confirm the presence of flood divides in a general case. In fact, flood frequency analyses are typically performed case by case by practitioners, who realize the presence of a flood divide in their data without relying on quantitative metrics, sometimes just because traditional approaches seem not to fit the data in any way. The selection yields 11 river basins with a clear flood divide in the empirical flood magnitude-frequency curve (Figure S1 in Supporting Information S1). All the following analyses are performed individually for each single basin.

2.2. Simplified Metastatistical Value Approach Applied to Heterogeneous Processes

We adopt the Simplified Metastatistical Extreme Value (SMEV) approach (Marani & Ignaccolo, 2015; Marra et al., 2019) to model the magnitude and frequency of extreme floods emerging from different runoff-generating processes (e.g., rain-on-wet or dry soils, snowmelt) in each single basin. The SMEV approach enables us to account for the presence of events triggered by multiple runoff-generation mechanisms (also termed event types in the following) through a parsimonious parametrization and to derive the compound extreme value distribution emerging from these types (Marra et al., 2019). Indicating with $i = 1 \dots S$ the event type in any of the basins, the mixed-SMEV cumulative distribution function ζ_{SMEV} can be written as

$$\zeta_{\text{SMEV}}(x) = \prod_{i=1}^S [F_i(x; \theta_i)]^{n_i} \quad (1)$$

where θ_i are the parameters of the cumulative distribution F_i of ordinary events of the i th type and n_i is the average yearly number of ordinary events of type i . Notice that only two event types are considered in each basin.

We compared estimates obtained through the mixed-SMEV approach against those of a single-SMEV and Generalized Extreme Value (GEV) distributions, a widely used approach for flood magnitude-frequency analyses (Katz et al., 2002; Petrow et al., 2007). To this end, we estimate the GEV distribution parameters from the sample of annual maxima of each basin using the L-moments method (Hosking, 1990). For the single-SMEV approach, we used a log-normal distribution to describe all ordinary events in each basin. The choice of a log-normal distribution is supported by previous study on the same data set of river basins in Germany (Mushtaq et al., 2022). We performed resampling with replacement (i.e., bootstrap) across years (1,000 realizations of all available years; Overeem et al., 2008) to assess the estimation uncertainty of all the methods. We finally quantified the model accuracy by means of non-dimensional error, computed as $\epsilon = \frac{x_{\text{est}} - x_{\text{obs}}}{x_{\text{obs}}}$ between estimated (x_{est}) and observed maxima (x_{obs}) (Zorzetto et al., 2016). Notice that such an error metric favors the GEV distribution, which is explicitly parameterized to match the observed maxima, as opposed to mixed-SMEV and single-SMEV approaches, whose parameter estimation is performed on a larger sample, of which just a subset belongs to the annual maxima.

2.3. Process-Based Classification of Ordinary Events

We employed the method of Lang et al. (1999) to identify the independent ordinary peaks from the streamflow record required to apply the SMEV approach, as previously done by Miniussi et al. (2020) and Mushtaq et al. (2022). The number of independent ordinary peaks obtained for each catchment ranges from 370 to 933 (median value: 796).

We further classified the ordinary peaks into process-based types by using the classification of Tarasova, Basso, Wendi, et al. (2020), which labels streamflow events corresponding to the identified ordinary peaks according to their runoff-generation processes, which are assessed based on the nature of the inducing events (i.e., rainfall vs. snowmelt) and the catchment wetness states (i.e., wet or dry). This process-based approach uses observed daily precipitation (Rauthe et al., 2013) as well as daily snow water equivalent and soil moisture simulated by the mesoscale Hydrological Model (Kumar et al., 2013; Samaniego et al., 2010). Tarasova, Basso, Wendi, et al. (2020) employed dimensionless indicators to differentiate between inducing events and catchment wetness states and observed the small uncertainties associated with model structure and parametric uncertainty (Figure 4 in Tarasova, Basso, Wendi, et al., 2020). However, it is worth noting that employing different indicators or classification frameworks could lead to substantial differences in event classification (Tarasova et al., 2019), consequently affecting the accuracy of flood estimates.

In order to efficiently incorporate distinct event types in flood magnitude-frequency analyses and to ensure sufficient sample sizes for each of them, we aggregate the event types by Tarasova, Basso, Wendi, et al. (2020) into two major groups ($S = 2$): processes related to dry antecedent conditions (rain-on-dry events—Type-1) and wet antecedent conditions (rain-on-wet and snowmelt events—Type-2). The probability distributions of the magnitude of these event types are significantly different ($p < 0.05$), as evaluated through a pairwise two-sided Kolmogorov-Smirnov test (Massey Jr, 1951).

2.4. Selection of Ordinary Distributions

For each event type, we choose a suitable ordinary distribution. The presence of a clear flood divide in the empirical flood magnitude-frequency distribution suggests that one of the event types is characterized by a heavy-tailed distribution (Merz et al., 2022). We used Weibull plotting positions (Weibull, 1939) to derive the empirical cumulative distributions of ordinary events of each type and distinguish between those either exhibiting or not a heavy-tailed behavior, here intended as a power-law tail (i.e., a linear behavior in double-logarithmic coordinates; Newman, 2005). This process can in principle be automated, for example, by using tests like the one applied by Marra et al. (2023). However, since the number of basins considered in this study is limited, it is preferred here to proceed with the accuracy of human supervision. Visual inspection of the empirical cumulative distribution

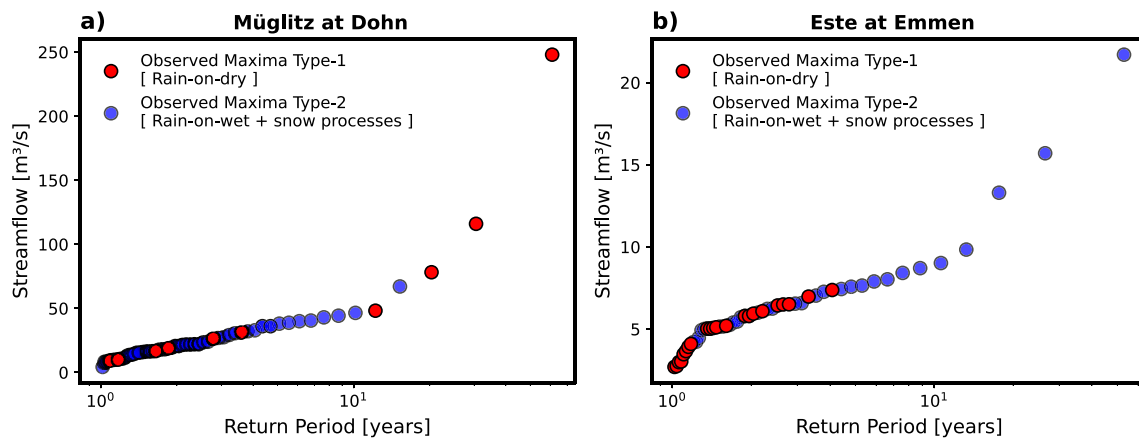


Figure 1. Empirical flood magnitude-frequency curves for two exemplary case studies exhibiting a flood divide: (a) the Müglitz River at Dohn (Gauge-ID: 550940, area = 196 km²) and (b) the Este River at Emmen (Gauge-ID: 6338260, area = 171 km²). Floods are classified into two major event types: rain-on-dry (red dots—Type-1) and combination of rain-on-wet and snow processes (blue dots—Type-2).

functions of ordinary events of the two types showed that, for each of the 11 study catchments, either Type-1 or Type-2 events clearly exhibit power-law behavior (see respectively Figures 2a and 2e), as manifested by the linear form of the distribution in a log-log probability plot.

If the empirical distribution of ordinary events of the examined type has a heavy tail, we use a power-law distribution (Malamud & Turcotte, 2003, 2006). The cumulative distribution function of the power-law is $P(x) = (x/x_{\min})^{1-\alpha}$, where, x is the analyzed variable, x_{\min} is a left-censoring threshold (i.e., the threshold above which the power-law behavior is manifest) and α is the scaling parameter. We estimated the parameters α and x_{\min} through the method of Clauset et al. (2009). If the empirical cumulative distribution function of the ordinary events of the examined type is not heavy-tailed (i.e., it does not show a linear behavior in double-logarithmic coordinates), we model it using a two parameter log-normal distribution (Bobe et al., 1993). The cumulative distribution function of the log-normal distribution is expressed as $F(x;\mu,\sigma) = \Phi(\log(x)-\mu)/\sigma$, where σ and μ are respectively its shape and scale parameters and Φ is the standard normal distribution function. We fit the log-normal distribution using the method of L-moments (Hosking, 1990) by left-censoring the lower portion of the ordinary events for the mixed-SMEV (i.e., considering Type-1 and Type-2) and single-SMEV (i.e., by considering all ordinary events without types). The left-censoring method enables us to characterize the tail of the distribution with few parameters, by ignoring the magnitudes of the censored part while retaining their probability (Marra et al., 2019). For both the power-law and log-normal distributions, we selected the left-censoring threshold by minimizing the root mean square error between predicted and observed magnitudes of ordinary events in the upper twentieth percentile (Ritter & Munoz-Carpena, 2013).

3. Results and Discussion

In this study, we investigate river basins that experienced extraordinarily high floods with much larger magnitudes than the bulk of recorded annual maxima. To better illustrate the properties of these rare floods we focus on two exemplary case studies, the Müglitz and Este River basins (Figure 1), which exhibit marked flood divides in their empirical flood magnitude-frequency curves. In fact, the largest annual maxima (roughly those with a return period exceeding 10 years) grow to considerably larger magnitudes than the smaller ones (e.g., the largest observed annual maxima are 3–8 times and 2–4 times larger than their mean values for the Müglitz and Este River basins, respectively).

In Figure 1, annual maxima are color-coded based on the type of processes that generated the events. We observe that floods with runoff-generation processes, different from those mostly observed for common floods (i.e., those on the left hand sides of panels a, b) may strongly affect the upper tail behavior, as also highlighted by Tarasova, Basso, and Merz (2020). For instance, extraordinarily high floods in the Müglitz River (Figure 1a) mainly belong to the rain-on-dry type (4 out of 6 maxima with a return period greater than 10 years are characterized as Type-1), while this type is hardly present in lower flood peaks. Conversely, extraordinarily high floods in the Este basin

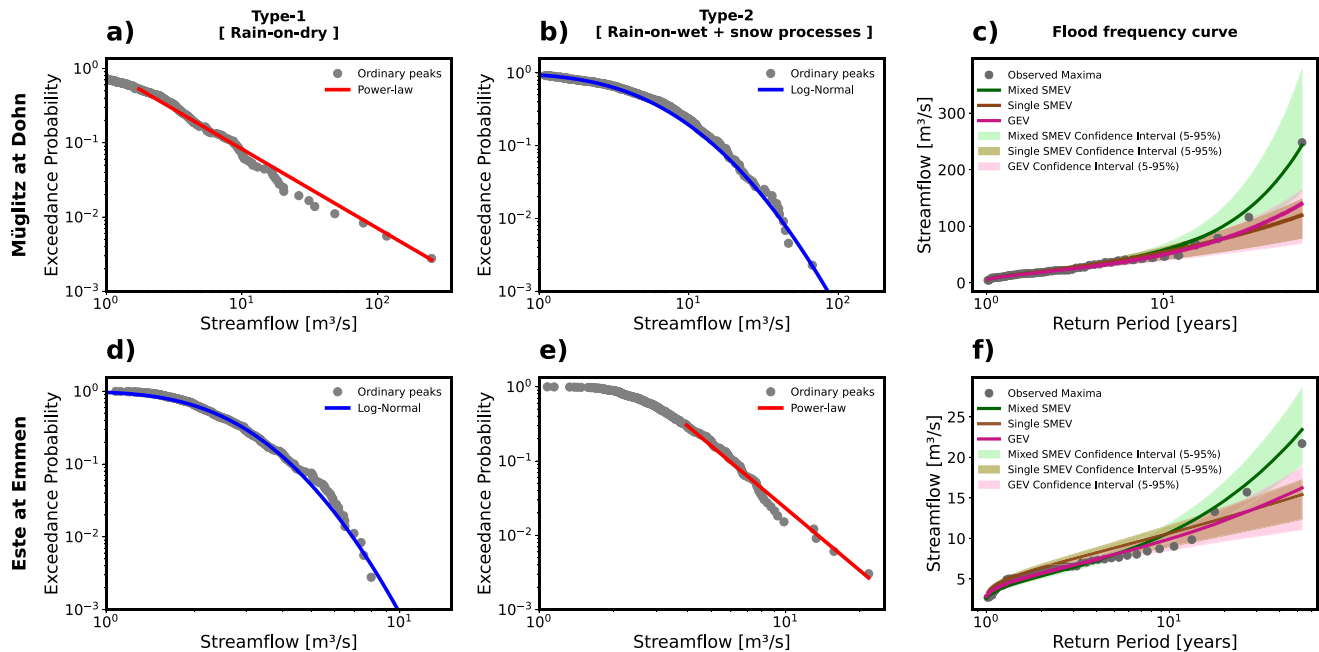


Figure 2. Exceedance cumulative distributions of ordinary peaks (a, b, d, e) and flood magnitude-frequency curves resulting from a mixed-SMEV, single-SMEV and a standard Generalized Extreme Value (GEV) approach (c, f) for the Müglitz River at Dohn (a–c) and the Este River at Emmen (d–f). Panels (a, d) show the ordinary distributions for Type-1 events, and panels (b, e) show the ordinary distributions for Type-2 events. Blue (log-normal) and red (power-law) lines in panels (a, b, d, e) display the probability distributions describing ordinary events of Type-1 and Type-2. Green, orange and pink curves in panels c and f show the median values for the corresponding quantiles of 1,000 resample values with replacement for mixed-SMEV, single-SMEV and GEV estimates, respectively. Green, orange and pink shaded areas indicate the related confidence intervals (5th–95th percentiles).

(Figure 1b) are caused by rain-on-wet conditions and snowmelt processes (5 maxima with a return period greater than 10 years are characterized as Type-2).

As described in Section 2.4, we examine the empirical cumulative distribution functions of the different types of ordinary events (gray dots in Figures 2a, 2b, 2d, and 2e). We notice that in all cases in which a flood divide is present one event type displays heavy-tailed behavior (Figures 2a and 2e). Hence, we fit the empirical ordinary distribution of these events with a power-law distribution (see red lines in Figures 2a and 2e), and use a log-normal distribution for the other type (blue lines in Figures 2b and 2d). While previous studies identified both heavy-tailed and light-tailed behaviors in flood records (Bernardara et al., 2008; Mushtaq et al., 2022), here we move a step further and show that different tail behaviors may be associated with distinct runoff-generation processes (Yu et al., 2022).

The mixed-SMEV flood magnitude-frequency curves obtained for the two exemplary case studies are displayed in Figures 2c and 2f (green solid lines). Gray dots in Figures 2c and 2f represent the empirical frequencies of the sample of annual maxima. Our analysis reveals that mixing different types of ordinary events with distinct tail behavior through the mixed-SMEV framework allows for capturing the upper tail of the distribution of annual maxima. This is opposed to standard methods that rely on identical distribution assumptions, which tend to underestimate the largest floods (results for the single-SMEV and GEV are shown for comparison in Figures 2c and 2f). It is relevant to mention that the mixed-SMEV approach ensures improved estimation of upper tail quantiles compared to single-SMEV and GEV, as the latter generally fail to capture the upper tail behavior (Figures 2c and 2f; Figures S3–S11 in Supporting Information S1). The shaded areas in Figures 2c and 2f depict the 95% confidence intervals calculated via bootstrap with replacement across years for mixed-SMEV, single-SMEV and GEV distributions. The single-SMEV has lower uncertainty than both the mixed-SMEV and GEV distributions (Miniussi & Marra, 2021). Despite the use of a larger portion of the data, the uncertainty of mixed-SMEV and GEV are instead similar. The underlying reason is twofold: by using a mixed-SMEV we estimate a larger number of parameters than with a single-SMEV approach, and often one of the event types is markedly less populated than the other, leading to an increase in uncertainty.

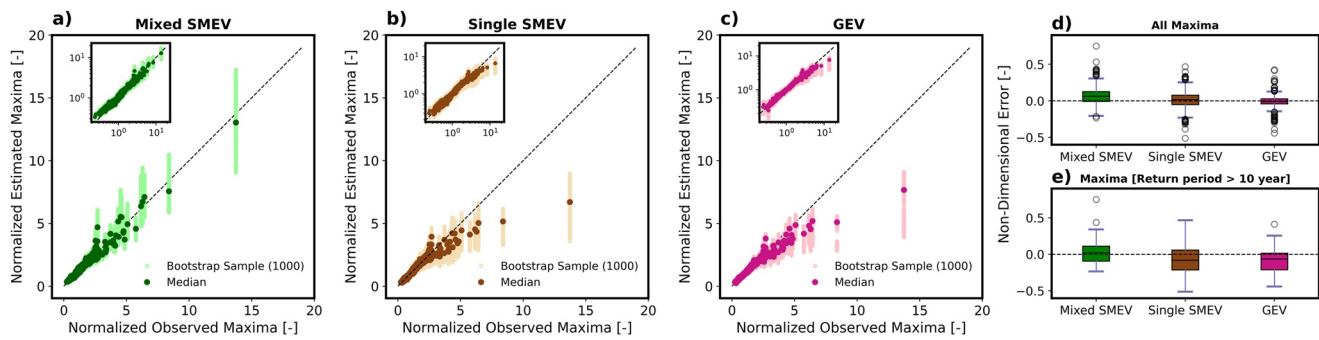


Figure 3. Estimated versus observed normalized (i.e., divided by their median value) annual maxima for (a) mixed-SMEV (green dots) (b) single-SMEV (orange dots) and (c) Generalized Extreme Value (pink dots) for 11 river basins in the data set which exhibit flood divides. Light and dark colors indicate results for 1,000 realizations of the bootstrap and their median values, respectively. Insets of panels (a–c) show the same results plotted on a double-logarithmic scale. Panels d and e show non-dimensional error between observed and estimated maxima of the analyzed statistical distributions computed for the median of 1,000 bootstrap values (i.e., dark green, orange and pink dots in a–c respectively) with the same return period: (d) all quantiles; (e) the quantiles corresponding to return periods >10 years.

These results demonstrate that relying solely on a pre-defined statistical distribution and disregarding the heterogeneity of the sample arising from different underlying physical processes may lead to an erroneous estimation of extreme floods. The proposed approach could thus improve the estimation of upper tail quantiles in basins where a flood divide is observed in the flood magnitude-frequency curves. These results confirm that heavy tails (i.e., flood divides) can originate from a mixture of flood-generating processes (Merz et al., 2022). It is worth noting that rain-on-dry events dominate the upper tail in the Müglitz River basin (Figure 2a), while rain-on-wet and snowmelt events dominate the upper tail in the Este River basin (Figure 2e). Past studies indicate that various runoff-generation processes might be associated with heavier tails of the distribution of floods in different regions (Tarasova et al., 2023), as we observe in our set of case studies. The flood divide arises from the fact that the mixture contains a heavy-tailed process, whose upper tail tends to dominate the mixture distribution at large quantiles (Figures 2c and 2f). These results are coherent with previous studies, which show that the tail of a mixture distribution is influenced by the component with the most pronounced tail (e.g., Cavanaugh et al., 2015). Results for all other case studies are reported in Figures S3–S11 in Supporting Information S1. These results are broadly consistent with what was found for the exemplary case studies discussed in Figures 1 and 2, despite a few cases showing a less clear distinction among event types. As expected, this leads to comparable performances of the three methods in these catchments.

Figure 3 summarizes the findings obtained for all 11 cases which exhibit flood divides in the set of 169 German river basins. Observed versus estimated annual maxima resulting from the proposed mixed-SMEV, the single-SMEV and the GEV approaches are evaluated in a bootstrap fashion (see Methods) and respectively displayed in panels a, b and c. This overall comparison confirms the results discussed above for the two exemplary case studies and highlights the capability of the mixed-SMEV approach to provide reliable estimates of floods for a wide range of quantiles in all river basins with a flood divide (Figure 3a). The comparison with the single-SMEV (Figure 3b) and the GEV (Figure 3c) distribution reveals that mixed-SMEV estimates (Figure 3a) are characterized by a considerably smaller bias, especially for the upper tail quantiles.

To summarize the performance of the mixed-SMEV approach to estimate flood magnitude-frequency in catchments with a flood divide, we computed the non-dimensional errors (see Section 2.2) between observed annual maxima and estimates of the corresponding empirical quantiles (Figures 3d and 3e). Figure 3d displays a boxplot of non-dimensional errors for all observed maxima in the 11 case studies. Here, the mixed-SMEV approach tends to overestimate the bulk of floods, mainly due to small floods with a return period less than 10 years, whereas single-SMEV and GEV distributions show the same degrees of over and underestimation. However, when we focus our analysis on quantiles with return period greater than 10 years (Figure 3e), which are the most relevant for flood hazard assessment, the non-dimensional errors provided by the mixed-SMEV approach are lower compared to single-SMEV and GEV distributions.

Separately fitting different probability distributions to samples of ordinary peaks with distinct tail behaviors increases the physical basis of flood magnitude-frequency analyses. In fact, the proposed method allows for accounting for different statistical properties of events triggered by various runoff-generation processes. It is,

however, worth noting that separating the processes yields significant improvements only if event types are characterized by markedly different tail behaviors (e.g., Marra et al., 2019). If the different processes have similar tail behaviors, the mixed distribution will not be distinguishable from the one obtained by using all the data together, since uncertainty will be predominant due to the need to estimate a larger number of parameters on a smaller amount of data. The practice of classifying event types evolved in recent years (Tarasova, Basso, Wendi, et al., 2020; Turkington et al., 2016; Vormoor et al., 2016) and is deemed to grow (Merz et al., 2022; Tarasova et al., 2019), increasing the availability of information required to apply the proposed approach.

A substantial difference among runoff-generation processes was suggested as a possible cause of flood divides by past studies. Rogger et al. (2012, 2013) showed, by means of extensive field surveys and modelling analyses, that flood divides may emerge from strong non-linearities in runoff-generation processes resulting from progressive saturation of the catchment (i.e., a shift from dry to wet conditions). Basso et al. (2016) provided a mechanistic explanation of this phenomenon and the resulting appearance of flood divides in flood magnitude-frequency curves by linking it to the catchment water balance. Catchments in wetter climates experience sustained water supply which determines unvaried runoff-generation processes. Conversely, river basins in drier areas, where longer lag times between rainfall events allow for the catchment to dry, undergo transient conditions leading to varied runoff-generation processes (Basso et al., 2023). Our results confirm that explicitly accounting for the existence of different runoff-generation processes enables us to adequately model the tail of flood distributions in catchments exhibiting flood divides.

The above discussion hints at the interplay between various hydro-meteorological drivers and the hydroclimatic settings leading to the occurrence of extraordinarily high floods in a region. In fact, river basins exhibiting flood divides (i.e., the exemplary Müglitz River and other eight case studies in our data set) are mostly located in the Central-Alpine region of Europe (Blöschl et al., 2017). In this region, the distributions of precipitation volumes for rain-on-wet and snowmelt floods have slightly heavier tails than for rain-on-dry events, leading to the possible occurrence of high floods of Type-2 (as exemplified by one river basin in our data set, see Figure S7 in Supporting Information S1). However, the distribution of precipitation intensity of rain-on-dry events in this area is remarkably heavier than for the other event types (Tarasova et al., 2023), mostly due to the occurrence of Vb-cyclones (Hofstätter et al., 2016). This feature likely underlies the occurrence of heavy-tailed distributions of rain-on-dry ordinary events and extraordinarily high floods of Type-1, which are mostly responsible (8 case studies) for the appearance of flood divides in flood-magnitude frequency curves. In contrast, the Este River and one additional catchment among the 11 showing a flood divide are located in the Atlantic region, where rain-on-wet flood events exhibit heavier tails than rain-on-dry events (Tarasova et al., 2023).

Variable distributions of key factors contributing to runoff generation, such as rainfall intensity, volume, soil moisture, snow accumulation and release are mirrored by differences in the tail behavior of ordinary distributions across regions, which can be leveraged by means of the proposed approach to improve the estimation of the hazard posed by extraordinarily high floods. However, climate change induces shifts in the mixture of flood generation processes, such as a decrease in snowmelt events and an increase in intense rainfall events (Hall et al., 2014; Huo et al., 2022). These changes inevitably influence the flood probabilities. Our method explicitly incorporates the presence of diverse flow-generating processes, and can therefore be used to predict changes in flood frequency based on the projected changes in the frequency of the runoff-generation processes.

4. Conclusions

We provide a framework to derive accurate estimates of flood magnitude and frequency for basins where extraordinarily high floods occur. The approach leverages knowledge of the heterogeneity of runoff-generation processes by means of a process-informed mixed non-asymptotic statistical method, the SMEV framework. We employ the mixed-SMEV to estimate flood magnitude and frequency for 11 river basins in Germany featuring a flood divide. In these cases, at least one runoff-generation process is characterized by a heavy-tailed empirical distribution. Thanks to the explicit consideration of various runoff-generation processes, the proposed approach enables us to accurately predict the magnitude of rare floods, outweighing the performance of single-SMEV and GEV distributions. The approach relies on a classification of event types and is only worth using when runoff events generated by different hydrologic processes are characterized by distributions with distinct tail behaviors. These requirements may constrain the use of the approach in practice. Nonetheless, classifications of event types are becoming more common, and the method provides a process-based solution to estimate large flood quantiles

in contexts for which current methods fail. As climate change may alter the frequency of different event types, our approach also offers a way to account for climate change impacts on flood hazards in a physically sound manner.

Data Availability Statement

For providing the discharge data for Germany, we are grateful to the Bavarian State Office of Environment (LfU, <https://www.gkd.bayern.de/de/fluesse/abfluss/tabellen>). Readers can directly access to river discharge data used in this study by clicking on the provided link. We are also thankful to the Global Runoff Data Centre (GRDC) prepared by the Federal Institute for Hydrology (BfG, <https://portal.grdc.bafg.de/applications/public.html?publicuser=PublicUser#dataDownload/Home>) for their valuable data resources. To access the data repository, please use the provided link and click on “Download by Station” option, select Germany as the country of interest, and locate the table in the top-left corner to access discharge gauges for Germany.

Acknowledgments

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—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).” The financial support of the Helmholtz Centre for Environmental Research—UFZ is also well acknowledged. The first author is also thankful to the Higher Education Commission of Pakistan (HEC) and the German Academic Exchange services (DAAD) for providing financial support to this study (DAAD- 91716446). FM was partially funded by the CARIPARO Foundation through the Excellence Grant 2021 to the “Resilience” project. Open Access funding enabled and organized by Projekt DEAL.

References

- Alila, Y., & Miraoui, A. (2002). Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques. *Hydrological Processes*, *16*(5), 1065–1084. <https://doi.org/10.1002/hyp.346>
- Barth, N. A., Villarini, G., Nayak, M. A., & White, K. (2017). Mixed populations and annual flood frequency estimates in the western United States: The role of atmospheric rivers. *Water Resources Research*, *53*(1), 257–269. <https://doi.org/10.1002/2016WR019064>
- Barth, N. A., Villarini, G., & White, K. (2019). Accounting for mixed populations in flood frequency analysis: Bulletin 17C perspective. *Journal of Hydrologic Engineering*, *24*(3), 04019002. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001762](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001762)
- Basso, S., Merz, R., Tarasova, L., & Miniussi, A. (2023). Extreme flooding controlled by stream network organization and flow regime. *Nature Geoscience*, *16*(4), 339–343. <https://doi.org/10.1038/s41561-023-01155-w>
- 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>
- Bernardara, P., Schertzer, D., Eric, S., 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>
- Bever, L., & Remondi, F. (2021). Natural catastrophes in 2021: The floodgates are open (Swiss Re Institute, 2022). Retrieved from <https://www.swissre.com/institute/research/sigma-research/sigma-2022-01.html>
- Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A., Merz, B., Arheimer, B., et al. (2017). Changing climate shifts timing of European floods. *Science*, *357*(6351), 588–590. <https://doi.org/10.1126/science.aan2506>
- Bobee, B., Cavadias, G., Ashkar, F., Bernier, J., & Rasmussen, P. (1993). Towards a systematic approach to comparing distributions used in flood frequency analysis. *Journal of Hydrology*, *142*(1–4), 121–136. [https://doi.org/10.1016/0022-1694\(93\)90008-W](https://doi.org/10.1016/0022-1694(93)90008-W)
- Cavanaugh, N. R., Gershunov, A., Panorska, A. K., & Kozubowski, T. J. (2015). The probability distribution of intense daily precipitation. *Geophysical Research Letters*, *42*(5), 1560–1567. <https://doi.org/10.1002/2015GL063238>
- Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. <https://doi.org/10.48550/arXiv.0706.1062>
- Cunnane, C. (1973). A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. *Journal of Hydrology*, *18*(3–4), 257–271. [https://doi.org/10.1016/0022-1694\(73\)90051-6](https://doi.org/10.1016/0022-1694(73)90051-6)
- Davenport, F. V., Burke, M., & Diffenbaugh, N. S. (2021). Contribution of historical precipitation change to US flood damages. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(4), e2017524118. <https://doi.org/10.1073/pnas.2017524118>
- Fischer, S. (2018). A seasonal mixed-POT model to estimate high flood quantiles from different event types and seasons. *Journal of Applied Statistics*, *45*(15), 2831–2847. <https://doi.org/10.1080/02664763.2018.1441385>
- François, B., Schlef, K. E., Wi, S., & Brown, C. M. (2019). Design considerations for riverine floods in a changing climate—a review. *Journal of Hydrology*, *574*, 557–573. <https://doi.org/10.1016/j.jhydrol.2019.04.068>
- Gotvald, A. J., Barth, N. A., Veilleux, A. G., & Parrett, C. (2012). Methods for determining magnitude and frequency of floods in California, based on data through water year 2006. *US Geological Survey Scientific Investigations Report*, *5113*(38). <http://pubs.usgs.gov/sir/2012/5113/>
- Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., et al. (2014). Understanding flood regime changes in Europe: A state-of-the-art assessment. *Hydrology and Earth System Sciences*, *18*(7), 2735–2772. <https://doi.org/10.5194/hess-18-2735-2014>
- Hirschboeck, K. K. (1987a). Catastrophic flooding and atmospheric circulation anomalies. In L. Mayer & D. Nash (Eds.), *Catastrophic flooding* (pp. 23–56). Allen & Unwin. <https://doi.org/10.4324/9781003020325-2>
- Hirschboeck, K. K. (1987b). Hydroclimatically-defined mixed distributions in partial duration flood series. In *hydrologic frequency modeling: Proceedings of the international symposium on flood frequency and risk analyses, 14–17 May 1986, Louisiana State University, Baton Rouge, USA* (pp. 199–212). Springer Netherlands. https://doi.org/10.1007/978-94-009-3953-0_13
- Hofstätter, M., Chimani, B., Lexer, A., & Blöschl, G. (2016). A new classification scheme of European cyclone tracks with relevance to precipitation. *Water Resources Research*, *52*(9), 7086–7104. <https://doi.org/10.1002/2016WR019146>
- Hosking, J. R. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B*, *52*(1), 105–124. <https://doi.org/10.1111/j.2517-6161.1990.tb01775.x>
- Huo, R., Li, L., Engeland, K., Xu, C. Y., Chen, H., Paasche, Ø., & Guo, S. (2022). Changing flood dynamics in Norway since the last millennium and to the end of the 21st century. *Journal of Hydrology*, *613*, 128331. <https://doi.org/10.1016/j.jhydrol.2022.128331>
- Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8), 1287–1304. [https://doi.org/10.1016/S0309-1708\(02\)00056-8](https://doi.org/10.1016/S0309-1708(02)00056-8)
- Kreibich, H., Van Loon, A. F., Schröter, K., Ward, P. J., Mazzoleni, M., Sairam, N., et al. (2022). The challenge of unprecedented floods and droughts in risk management. *Nature*, *608*(7921), 80–86. <https://doi.org/10.1038/s41586-022-04917-5>

- Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, 49(1), 360–379. <https://doi.org/10.1029/2012WR012195>
- Lang, M., Ouarda, T. B., & Bobee, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3–4), 103–117. [https://doi.org/10.1016/S0022-1694\(99\)00167-5](https://doi.org/10.1016/S0022-1694(99)00167-5)
- Malamud, B. D., & Turcotte, D. L. (2003). Shelf record of climatic changes in flood magnitude and frequency, north-coastal California: Comment. *Geology*, 31(3), 288. [https://doi.org/10.1130/0091-7613\(2003\)031<0288:SROCCI>2.0.CO;2](https://doi.org/10.1130/0091-7613(2003)031<0288:SROCCI>2.0.CO;2)
- Malamud, B. D., & Turcotte, D. L. (2006). The applicability of power-law frequency statistics to floods. *Journal of Hydrology*, 322(1–4), 168–180. <https://doi.org/10.1016/j.jhydrol.2005.02.032>
- Marani, M., & Ignaccolo, M. (2015). A metastatistical approach to rainfall extremes. *Advances in Water Resources*, 79, 121–126. <https://doi.org/10.1016/j.advwatres.2015.03.001>
- Marra, F., Amponsah, W., & Papalexiou, S. M. (2023). Non-asymptotic Weibull tails explain the statistics of extreme daily precipitation. *Advances in Water Resources*, 173, 104388. <https://doi.org/10.1016/j.advwatres.2023.104388>
- Marra, F., Zoccatelli, D., Armon, M., & Morin, E. (2019). A simplified MEV formulation to model extremes emerging from multiple nonstationary underlying processes. *Advances in Water Resources*, 127, 280–290. <https://doi.org/10.1016/j.advwatres.2019.04.002>
- Massey, F. J., Jr. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253), 68–78. <https://doi.org/10.1080/01621459.1951.10500769>
- 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*, 58(6), e2021WR030506. <https://doi.org/10.1029/2021WR030506>
- Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth & Environment*, 2(9), 592–609. <https://doi.org/10.1038/s43017-021-00195-3>
- Miniussi, A., Marani, M., & Villarini, G. (2020). Metastatistical extreme value distribution applied to floods across the continental United States. *Advances in Water Resources*, 136, 103498. <https://doi.org/10.1016/j.advwatres.2019.103498>
- Miniussi, A., & Marra, F. (2021). Estimation of extreme daily precipitation return levels at-site and in ungauged locations using the simplified MEV approach. *Journal of Hydrology*, 603, 126946. <https://doi.org/10.1016/j.jhydrol.2021.104085>
- Miniussi, A., Merz, R., Kaule, L., & Basso, S. (2023). Identifying discontinuities of flood frequency curves. *Journal of Hydrology*, 617, 128989. <https://doi.org/10.1016/j.jhydrol.2022.128989>
- 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, 104127. <https://doi.org/10.1016/j.advwatres.2022.104127>
- Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323–351. <https://doi.org/10.1080/0010751050052444>
- Overeem, A., Buishand, A., & Holleman, I. (2008). Rainfall depth-duration-frequency curves and their uncertainties. *Journal of Hydrology*, 348(1–2), 124–134. <https://doi.org/10.1016/j.jhydrol.2007.09.044>
- Petrow, T., Merz, B., Lindenschmidt, K.-E., & Thieken, A. H. (2007). Aspects of seasonality and flood generating circulation patterns in a mountainous catchment in south-eastern Germany. *Hydrology and Earth System Sciences*, 11(4), 1455–1468. <https://doi.org/10.5194/hess-11-1455-2007>
- Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., & Gratzki, A. (2013). A Central European precipitation climatology—Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, 22(3), 235–256. <https://doi.org/10.1127/0941-2948/2013/0436>
- Ritter, A., & Munoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, 480, 33–45. <https://doi.org/10.1016/j.jhydrol.2012.12.004>
- Rogger, M., Pirkel, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., et al. (2012). Step changes in the flood frequency curve: Process controls. *Water Resources Research*, 48(5), W05544. <https://doi.org/10.1029/2011WR011187>
- Rogger, M., Viglione, A., Derr, J., & Blöschl, G. (2013). Quantifying effects of catchments storage thresholds on step changes in the flood frequency curve. *Water Resources Research*, 49(10), 6946–6958. <https://doi.org/10.1002/wrcr.20553>
- Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46(5), W05523. <https://doi.org/10.1029/2008WR007327>
- Sikorska, A. E., Viviroli, D., & Seibert, J. (2015). Flood-type classification in mountainous catchments using crisp and fuzzy decision trees. *Water Resources Research*, 51(10), 7959–7976. <https://doi.org/10.1002/2015WR017326>
- Sivapalan, M., & Samuel, J. M. (2009). Transcending limitations of stationarity and the return period: Process-based approach to flood estimation and risk assessment. *Hydrological Processes*, 23(11), 1671–1675. <https://doi.org/10.1002/hyp.7292>
- 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>
- Stein, L., Pianosi, F., & Woods, R. (2020). Event-based classification for global study of river flood generating processes. *Hydrological Processes*, 34(7), 1514–1529. <https://doi.org/10.1002/hyp.13678>
- Tarasova, L., Basso, S., & Merz, R. (2020). Transformation of generation processes from small runoff events to large floods. *Geophysical Research Letters*, 47(22), e2020GL090547. <https://doi.org/10.1029/2020GL090547>
- Tarasova, L., Basso, S., Wendi, D., Viglione, A., Kumar, R., & Merz, R. (2020). A process-based framework to characterize and classify runoff events: The event typology of Germany. *Water Resources Research*, 56(5), e2019WR026951. <https://doi.org/10.1029/2019WR026951>
- 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 & Environment*, 4(1), 49. <https://doi.org/10.1038/s43247-023-00714-8>
- Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., et al. (2019). “Causative classification of river flood events.” Wiley interdisciplinary reviews: Water (Vol. 6, p. e1353). <https://doi.org/10.1002/wat2.1353>
- Turkington, T., Breinl, K., Ettema, J., Alkema, D., & Jetten, V. (2016). A new flood type classification method for use in climate change impact studies. *Weather and Climate Extremes*, 14, 1–16. <https://doi.org/10.1016/j.wace.2016.10.001>
- Villarini, G., & Smith, J. A. (2010). Flood peak distributions for the eastern United States. *Water Resources Research*, 46(6). <https://doi.org/10.1029/2009WR008395>
- Vogel, R. M., & Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. *Water Resources Research*, 29(6), 1745–1752. <https://doi.org/10.1029/93WR00341>
- 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>

- Waylen, P., & Woo, M. K. (1982). Prediction of annual floods generated by mixed processes. *Water Resources Research*, *18*(4), 1283–1286. <https://doi.org/10.1029/WR018i004p01283>
- Weibull, W. (1939). A statistical theory of strength of materials. *IVB-Handl*. <https://cir.nii.ac.jp/crid/1573668925268278656>
- Yu, G., Wright, D. B., & Davenport, F. V. (2022). Diverse physical processes drive upper-tail flood quantiles in the US Mountain West. *Geophysical Research Letters*, *49*(10), e2022GL098855. <https://doi.org/10.1029/2022GL098855>
- Zorzetto, E., Botter, G., & Marani, M. (2016). On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters*, *43*(15), 8076–8082. <https://doi.org/10.1002/2016GL069445>