



RESEARCH LETTER

10.1029/2020GL090547

Transformation of Generation Processes From Small Runoff Events to Large Floods

L. Tarasova¹ , S. Basso¹ , and R. Merz¹ ¹Department Catchment Hydrology, Helmholtz Centre for Environmental Research-UFZ, Halle (Saale), Germany

Key Points:

- We analyzed transformation of processes from small runoff events to larger floods using a process-based framework for event characterization
- A substantial transformation of the frequency of processes from small runoff events to frequent and upper tail floods is observed
- Differences in trajectories of process transformation among catchments suggest regionally variable predictability of extremes

Supporting Information:

- Supporting Information S1

Correspondence to:

L. Tarasova,
larisa.tarasova@ufz.de

Citation:

Tarasova, L., Basso, S., & Merz, R. (2020). Transformation of generation processes from small runoff events to large floods. *Geophysical Research Letters*, 47, e2020GL090547. <https://doi.org/10.1029/2020GL090547>

Received 26 AUG 2020

Accepted 9 NOV 2020

Accepted article online 16 NOV 2020

Abstract Mixture of runoff generation processes poses a challenge for predicting upper flood quantiles. We examined transformations of generation processes from all identifiable runoff events to frequent and upper tail floods for a large set of mesoscale catchments and observed a substantial change of the dominant processes. Two trajectories of transformation were detected. In regions where floods occur almost exclusively in winter the dominance of processes related to snowmelt consistently increases from small events to frequent and upper tail floods. In catchments characterized by frequent winter-spring floods and occasional summer-autumn flood events triggered by rare meteorological phenomena (e.g., Vb cyclones), processes that dominate upper tails are not adequately represented in the sample of frequent floods. Predictions of extremes and projections of flood changes might remain highly uncertain in the latter cases.

Plain Language Summary Prediction of floods remains a challenging task for the engineering practice. Floods triggered by different physical mechanisms have contrasting statistical attributes. Mixture of these processes hinders reliable prediction of the largest floods. In this study we classified a large number of streamflow events observed in a wide range of German river basins according to their generation processes. We analyzed changes in the frequency of occurrence of different generation processes, from the smallest identifiable runoff events to annual floods to rarer events corresponding to larger river flows. Interestingly, for some river basins certain processes tend to consistently increase their frequency from small streamflow events to common and larger floods. In other cases, we observed an opposite tendency. Certain processes become less important for the generation of annual floods compared to small streamflow events but then dominate generation of the largest floods. This has important implications for our ability to predict extreme floods and their possible changes.

1. Introduction

Predicting magnitude and frequency of large (extreme) floods and their possible changes is critical for hazard assessment, making the topic of their origins of significant practical importance (Smith et al., 2018). Detailed studies of extreme and catastrophic floods (e.g., Blöschl et al., 2013; Hirschboeck, 1987a; Nakamura et al., 2013; Rogger et al., 2012; Smith et al., 2018) and specific atmospheric mechanisms generating the largest flood peaks (e.g., Barth et al., 2017; Doswell et al., 1996; De Luca et al., 2017; Lima et al., 2017; Ralph et al., 2006; Villarini et al., 2011) provided evidences on the heterogeneity of processes in flood samples. These works fueled debates on the origins and hence the choice of appropriate methods for estimating upper tail (i.e., the largest) floods. Moreover, recent process-based investigations of flood changes (e.g., Berghuijs et al., 2019; Blöschl et al., 2019; Kemter et al., 2020) showed that their observed disparate regional patterns might in fact be driven by changes in the mix of flood generation processes, thus further complicating reliable estimations of future flood hazards.

Discussions on the origins and predictability of upper tail and extreme flood events from the sample of annual floods and ordinary runoff events have a long history (e.g., Klemes, 1993; Rossi et al., 1984). Several studies on the hydroclimatic nature of flood events in the United States (e.g., Barth et al., 2019; Hirschboeck, 1987b; Smith et al., 2011, 2018; Villarini & Smith, 2010) showed that the nature of upper tail events is often different from that of annual floods. On the other hand, recently developed approaches that utilize the whole range of observed streamflows to derive flood frequency curves (Basso et al., 2016; Claps & Laio, 2003; Miniussi et al., 2020) delivered more accurate predictions of upper tails, thus supporting the hypothesis that the largest floods (statistically) might indeed originate from ordinary runoff events.

©2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Regional and global surveys of processes generating annual and partial duration flood series (Hirschboeck, 1987b; Merz & Blöschl, 2003; Sikorska et al., 2015; Stein et al., 2019) showed that a mixture of processes is the norm. However, it is still unclear if the mixture of processes detected in frequent floods (e.g., annual maxima or peaks-over-threshold) is representative for upper tail events. A consistent process-based characterization of all observed runoff events spanning the whole range of magnitudes might help to detect possible transformation of processes from ordinary runoff events to annual floods and further to events with higher return periods. Such analyses in a wide set of catchments can provide valuable insights on regional variations in the transformation of processes and highlight regions where special care should be taken when predicting extremes or projecting possible future changes of floods.

Therefore, this study aims to investigate the nature and composition of processes generating different event samples, and to examine their transformations going from all identifiable runoff events to frequent (e.g., peaks-over-thresholds and annual maxima) and upper tail floods (e.g., events with return periods >10 years) for a large set of mesoscale catchments.

2. Data and Methods

The event characterization was performed on a set of 203,852 runoff events that occurred in 172 German catchments (31–23,700 km², median = 516 km², Figure 1a) and were separated from daily hydrometeorological time series by means of the automated method of Tarasova et al. (2018). The study period spans from 1951 to 2013 with a median observation length of 61 years. Only catchments with limited effects from reservoir operation were considered. Land use changes during the study period are limited (<2% of the area for most catchments) and do not considerably affect peak discharges (Tarasova et al., 2018).

In this study we primarily focused on four event samples: ordinary (all identifiable) runoff events isolated by the above mentioned event separation method (see Text S1 in the supporting information for details), two alternative samples of frequent floods defined as series of maximum annual floods (MAFs) and peak-over-threshold events (POT4, the threshold was selected iteratively to sample on average 4 events per year), and upper tail floods (flood events with discharges corresponding to return periods of at least 10 years [HQ10]; Smith et al., 2018; Villarini & Smith, 2010). Discharges corresponding to specific empirical return periods were estimated using Weibull plotting position of annual maxima. A median number of 1,214 ordinary, 216 POT4, 61 MAF, and 6 HQ10 events per catchment were identified. Event samples were identified and analyzed catchment-wise, so that less frequent events always correspond to larger discharges in each catchment. Hence, in this study we term less frequent events as large events and more frequent ones as small events.

To enhance robustness of the possibly detected transformations of dominant processes in the above samples, we also considered maximum seasonal floods (MSF) (i.e., maximum flows in predefined seasons) and flood events with return periods of at least 2, 5, 15, and 20 years (HQ2, HQ5, HQ15, and HQ20). Samples of flood events with larger return periods were not considered due to the limited length of the available discharge series.

Moreover, we examined sensitivity of the results to specific sampling approaches by using different POT samples (two, three, four, and five events per year), different MSF samples (two or four seasons) and deriving samples based on empirical return periods (e.g., HQ10) from both annual and partial duration series. Maximum, median, and minimum numbers of events per catchment corresponding to each event sample are reported in Table S1.

For characterizing event conditions we used daily gridded rainfall data from the REGNIE data set (Rauthe et al., 2013), gridded snowmelt and soil moisture series simulated by the mesoscale Hydrological Model (Kumar et al., 2013; Samaniego et al., 2010), and air temperature interpolated using external drift kriging (Zink et al., 2017).

All events were categorized according to the multi-layer process-based framework of Tarasova et al. (2020). The framework uses indicators that characterize space-time dynamics of inducing precipitation events and their spatial interaction with antecedent catchment states, such as snow cover, frozen soils, and soil moisture content (Figure 1b). Each event was categorized at each layer based on predefined thresholds applied to the layer-specific dimensionless indicators (see Text S2 and Table S2) and labeled accordingly (Figure 1b). The states categorized at each layer of the framework can be combined hierarchically (Figure 1c) to obtain a

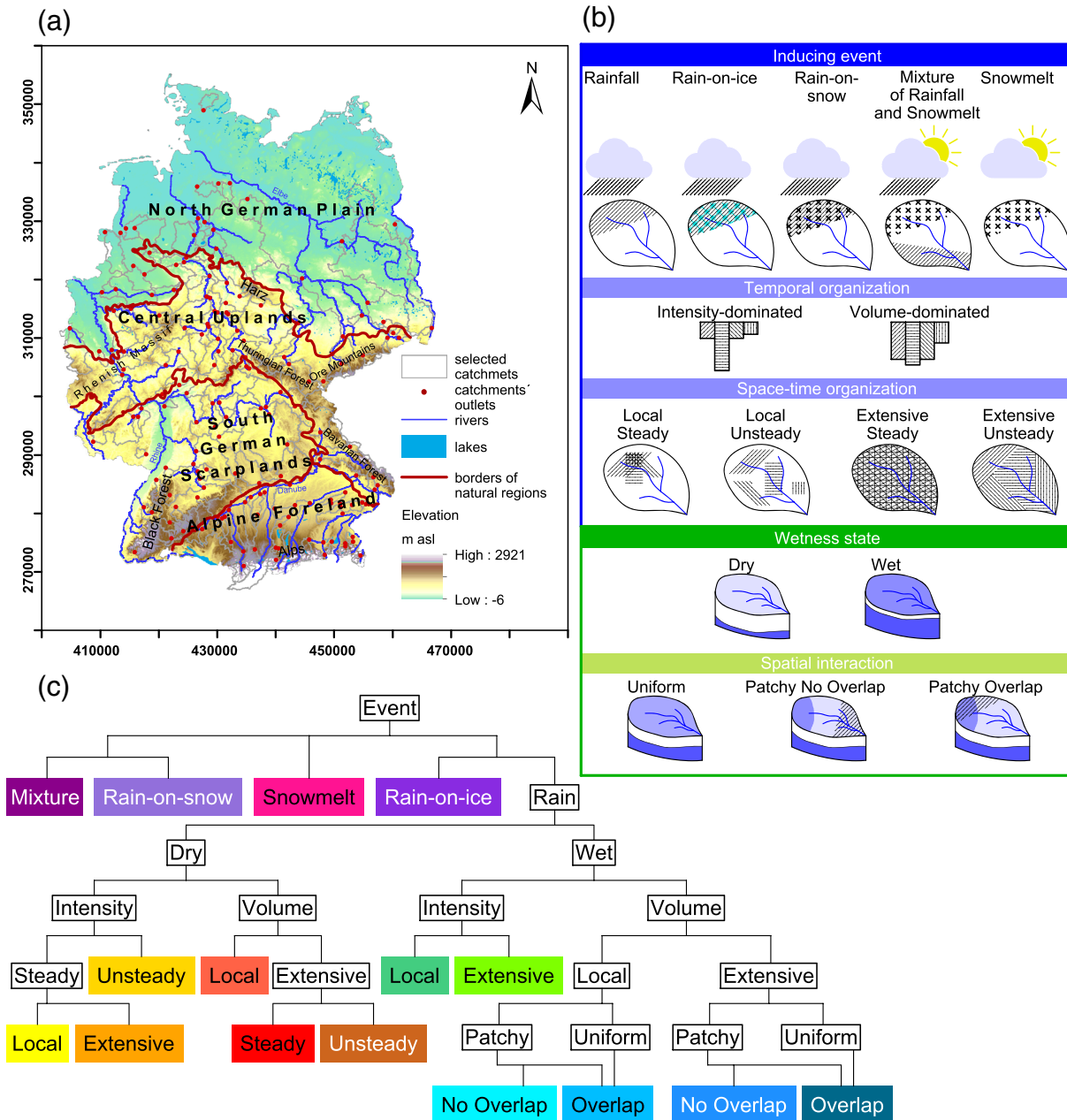


Figure 1. (a) Study catchments and German natural regions. (b) A framework for process-based event characterization. The nature of the inducing event is analyzed using the volumetric ratio of rainfall and snowmelt, the spatial covariance of rainfall volume, pre-event snow cover, and degree of soil freezing. *Intensity*- or *Volume*-dominated events are categorized by means of the temporal coefficient of variation of the precipitation rates and the ratio of intensity and volume, while their space-time organization is stratified into *Local Steady*, *Local Unsteady*, *Extensive Steady*, and *Extensive Unsteady* using the spatial coefficient of variation of the precipitation volume and the spatial covariance of precipitation fields between consecutive time steps. Events are labeled according to the state of the antecedent soil moisture into *Wet* and *Dry*. The spatial distribution of soil moisture is quantified by its spatial coefficient of variation (*Uniform* or *Patchy*). In case of *Patchy* (uneven) distribution, the spatial interaction of soil moisture and event precipitation is evaluated (*Overlap* or *No Overlap*) by means of their spatial covariance. Indicators and thresholds used to categorize events at each layer are described in Text S2 and Table S2. (c) A decision tree for hierarchical event classification that combines all layers of the framework into event types (Text S3). Modified from Tarasova et al. (2020) CC BY.

unique type for each event. The framework and hierarchical decision tree applied for classification of runoff events in Germany showed regionally consistent performance allowing comparison of classified events in the study catchments (Tarasova et al., 2020).

Frequency of occurrence (fraction) of each possible category was calculated for each layer of the framework (Figure 1b) for all considered event samples (e.g., ordinary events, MAF, HQ10) separately.

To evaluate the statistical significance of possibly detected transformations of processes, we applied the chi-square (χ^2) test for categorical data in two different fashions. First, we checked if generation processes characterized at individual layers of the framework (e.g., *Intensity*- or *Volume*-dominated events) are evenly distributed in a particular event sample (e.g., MAF). This goodness-of-fit test examines if the frequency distribution of processes is statistically different from an expected (i.e., equal) distribution and rejects the null hypothesis if at least one of the processes is more likely to occur. In the second setting, we used the χ^2 test as a homogeneity test to check if the distributions of generation processes are different between event samples (e.g., between MAF and HQ10). The null hypothesis that the distributions of generation processes are homogeneous (i.e., similar) is rejected if the tested event samples have statistically different distribution of generation processes. p values were estimated using Monte Carlo simulations with 10,000 replicates (Hope, 1968) to avoid issues with asymptotic approximations that might arise due to imbalanced event samples (Agresti, 2007).

The dissimilarity between distributions of hierarchical event types for different event samples was quantified using the Euclidean distance:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (1)$$

where event type frequency distributions of two different event samples are given as $\mathbf{p} = (p_1, p_2, \dots, p_n)$ and $\mathbf{q} = (q_1, q_2, \dots, q_n)$ in the Euclidean n space, with n total number of possible hierarchical event types (i.e., the combinations of the categories from each layer of the framework in Figure 1c). Accordingly, p_i is the frequency of one event type (e.g., *Rain-on-snow*) in one event sample (e.g., MAF) and q_i is the frequency of the same event type in the alternative event sample (e.g., HQ10). Euclidean distance is 0 when two event samples have identical distributions of event types.

3. Results

Figures 2 and 3 show a substantial transformation of dominant generation processes from small runoff events to larger floods. Pairwise χ^2 homogeneity tests indicated significant differences (Table S3) in the frequency of generation processes between samples of ordinary events and frequent floods (Figures 2 and 3, All-POT4; for 98% of the catchments on average for the different layers of the framework), between ordinary and upper tail events (HQ10-All, 36% of catchments), between frequent and upper tail floods (MAF-HQ10, 11% of catchments), and between annual floods and peak-over-threshold events (POT4-MAF, 7% of catchments).

The χ^2 tests for goodness-of-fit indicated that the frequency of different generation processes is distributed unequally for most event samples (Table S4). In the following, we describe the emerged patterns in the distribution of generation processes among different event samples in the study catchments.

Ordinary events are mostly *Rainfall*-induced, while a considerable portion of frequent floods is classified as *Rain-on-snow* or *Mixture of rainfall and snowmelt* (Figure 2, All). The emphasis shifts back to *Rainfall*-induced events for upper tail floods, especially in the southern regions and in mountainous catchments in Eastern Germany. *Rain-on-snow* floods instead dominate upper tails in catchments draining the Central Uplands. *Rain-on-ice* events generally occur very rarely (Figure 2, All) but can lead to large floods in several catchments (Figure 2, HQ10).

Among ordinary events, *Volume*-dominated events are more common than *Intensity*-dominated ones (Figure 2, All). This is not surprising, since intense rainfall storms occur relatively seldom in temperate climates like Germany, where long synoptic precipitation events related to frontal activities prevail (Mediero et al., 2015). Their dominance is even more pronounced for frequent floods (Figure 2, POT4, MAF).

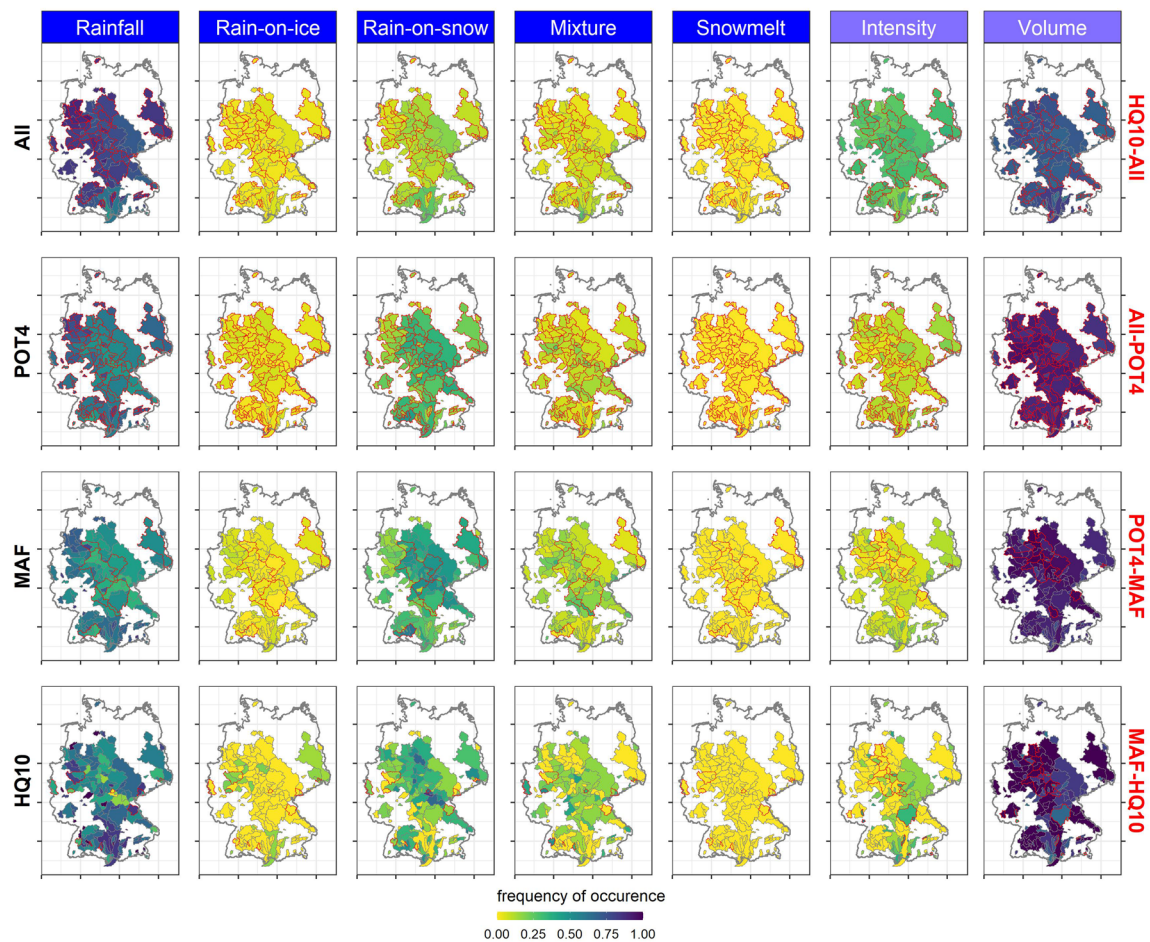


Figure 2. Fraction of different inducing events (dark blue column headings on the left side) and their temporal organization (light blue column headings on the right side) corresponding to the first two layers of the characterization framework (Figure 1b) in the event samples indicated in black font on the left side of the panels: ordinary events (*All*), peak-over-threshold events (*POT4*), maximum annual floods (*MAF*), and events with return periods of at least 10 years (*HQ10*). For each layer the relative frequency of occurrence of the corresponding categories sums up to unity. Significant difference ($p < 0.1$) in the distribution of categories between pairs of event samples (indicated in red font on the right side of the panels) was confirmed by χ^2 tests of homogeneity (Table S3) for catchments displayed with a red outline in the maps. The results of the χ^2 tests for goodness-of-fit, that test the hypothesis of equal frequency distribution of different categories within each event sample, are available in Table S4.

Nonetheless, floods with higher return periods can be also generated by *Intensity*-dominated events in the southern and eastern regions (Figure 2, HQ10).

Equal portions of ordinary events are attributed to *Local* and *Extensive* categories. Instead, *Steady* events (i.e., precipitation occurring over the same portion of a catchment during consecutive days) are more frequent than *Unsteady* ones (Figure 3, All). This tendency becomes more prominent for frequent floods, which mainly exhibit *Extensive* spatial structures (Figure 3, POT4, MAF). Similarly, the majority of upper tail events have an *Extensive Steady* organization (Figure 3, HQ10). However, *Local Steady* events are as likely to trigger larger floods (Figure 3, HQ10) in small mountainous catchments of the Central Uplands and Alpine Forelands.

Dry catchment conditions slightly prevail among all events (Figure 3, All), whereas frequent and upper tail floods are mainly generated during *Wet* conditions (Figure 3, POT4, MAF, HQ10). Nonetheless, a portion of the largest floods can be generated during *Dry* conditions as well (Figure 3, HQ10).

For what concerns the spatial interactions between catchment wetness states and inducing precipitation, runoff events are rarely generated by rainfall that does not overlap with wet areas of the catchment (Figure 3, *Patchy No Overlap*). When considering larger events, a *Uniform* spatial distribution of catchment wetness

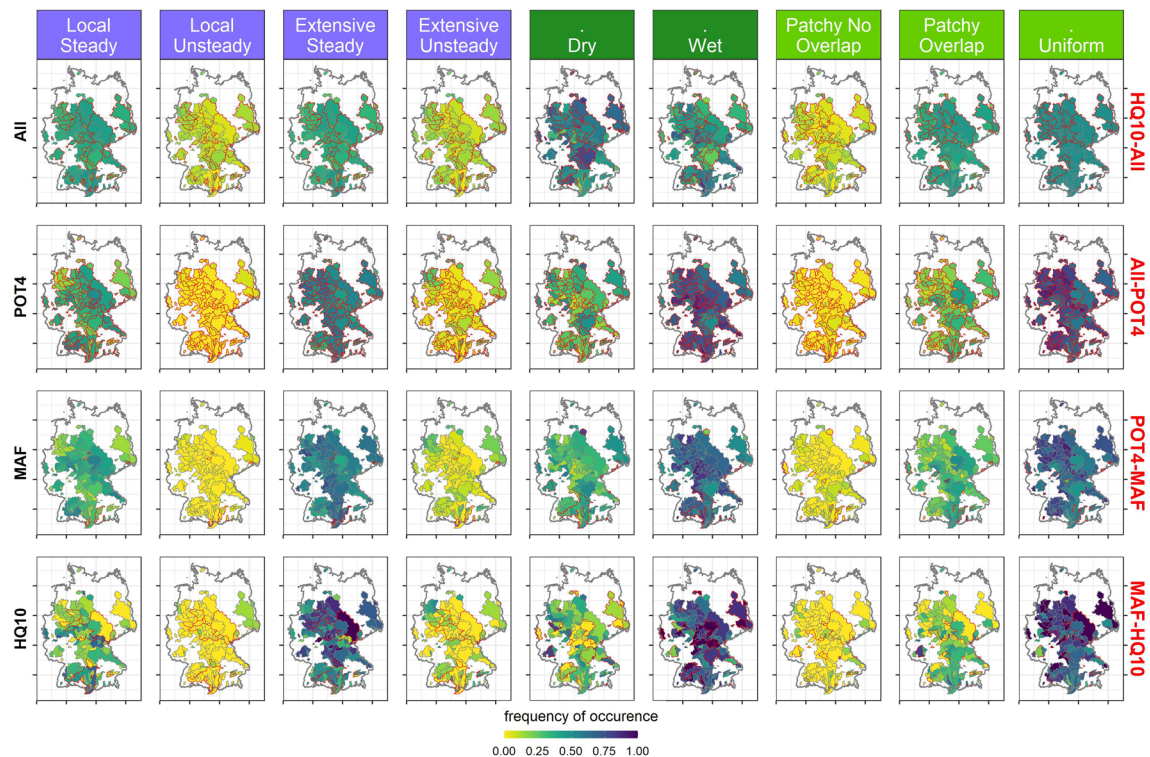


Figure 3. Fraction of different space-time organizations of inducing events (light blue column headings on the left side), wetness states (dark green column headings), and spatial interactions between wetness state and inducing event (light green column headings on the right side), corresponding to the remaining layers of the characterization framework (Figure 1b) in different event samples. For details see caption of Figure 2.

occurs more and more frequently. Nevertheless, a considerable portion of upper tail floods still occurs under *Patchy* distribution of the soil moisture (Figure 3, HQ10).

Similar findings on the transformation of generation processes from small runoff events to larger floods emerge when alternative samples of events (e.g., maximum seasonal floods and events with return periods of at least 2, 5, 15, and 20 years; Figures S1 and S2) are considered, further supporting the identified transformations of dominant generation processes.

Different sampling approaches for MSF (i.e., different number of seasons) and POT events (i.e., different number of events selected yearly) seem to have no major effect on the derived distributions of generation processes (no statistically significant differences among these samples were detected in most catchments, Table S5). Deriving samples of events with specific return periods (e.g., HQ10) either using annual maxima or peaks-over-threshold has no effect on the identified distributions of processes (Table S6).

4. Discussion

Our results showed presence of substantial transformations of dominant generation processes from small events to larger floods along two contrasting trajectories. For some catchments the portion of a particular generation process (e.g., *Rain-on-snow* or *Volume*-dominated events) consistently increases from the sample of ordinary events to frequent and upper tail floods. For others, the dominance of certain processes (e.g., *Rainfall*-induced events or events during *Dry* conditions) only emerges for upper tail events, while their portions in frequent floods are smaller than in the sample of all events.

The detected transformations of dominant generation processes can be summarized by the difference in the distributions of event types between the samples, here quantified by the Euclidean distance between these distributions and evaluated statistically by means of a χ^2 test of homogeneity. Since the sample of all events includes the whole range of possible event types (e.g., Figures 4d and 4e) and is expected differ from the sample of upper tail floods, we chose the Euclidean distance between these samples as a reference and evaluated the relative change in the distance between the samples of frequent and upper tail floods.

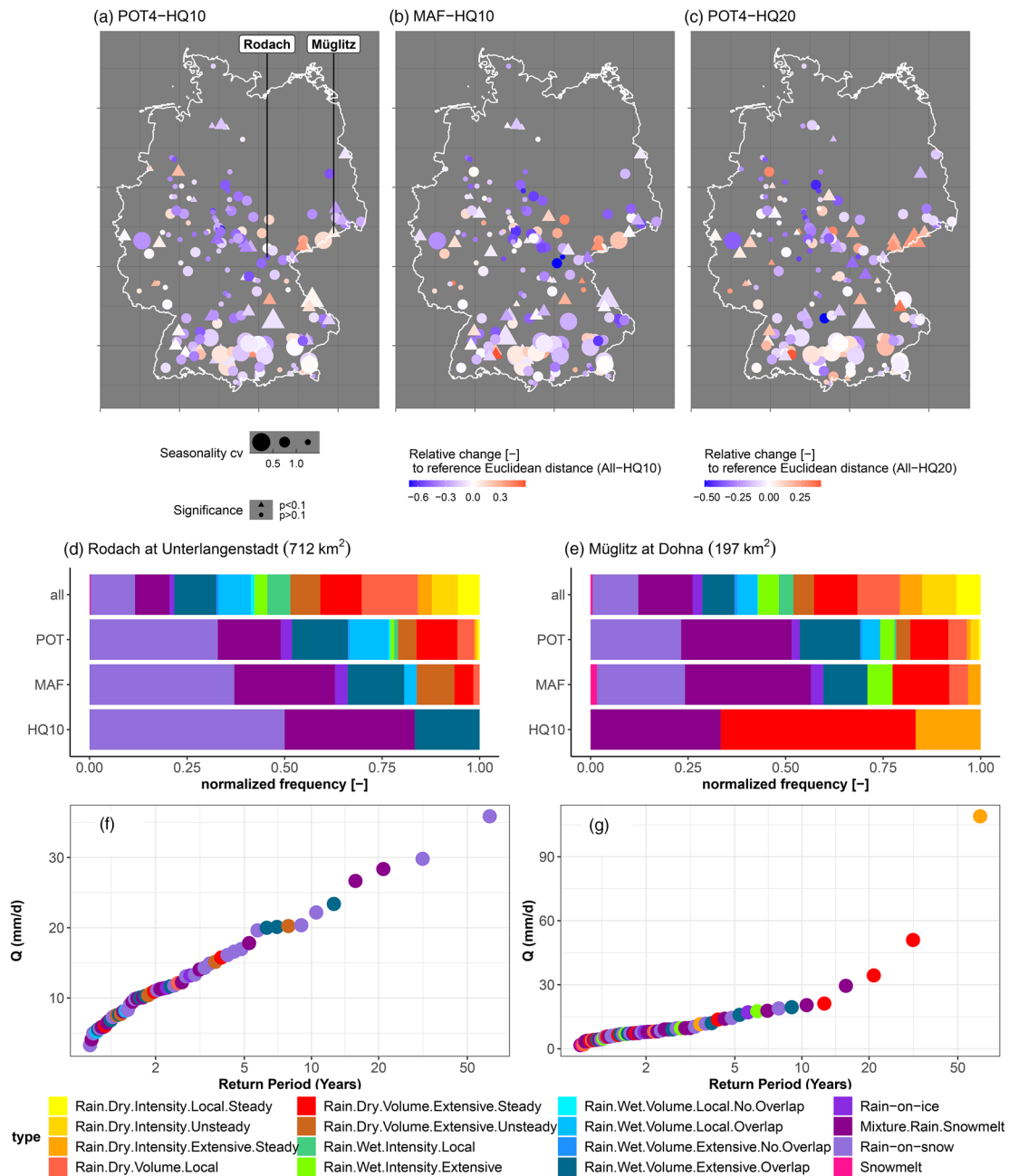


Figure 4. (a–c) Dissimilarity between the distribution of event types (i.e., hierarchical aggregates of all layers of characterization framework, Figure 1c) of upper tail floods ($HQ10$, $HQ20$) and samples of frequent floods (MAF , $POT4$) measured by their relative changes in Euclidean distance compared to the reference distance between samples of ordinary (All) and upper tail events in the study catchments (displayed as catchment centroids). Negative relative changes (blue colors) indicate that frequent floods are more similar to upper tail floods than ordinary events in terms of event type distribution. Positive relative changes (red colors) instead indicate that frequent floods are more dissimilar to upper tails than ordinary events. Catchments with significant differences in the distributions of event types ($p < 0.1$) between samples of frequent and upper tail floods are indicated as triangles. The size of the points reflects the seasonality of floods quantified as the coefficient of variation of the number of annual maxima occurred in different months. Examples of two catchments: (d) Rodach and (e) Müglitz with contrasting distributions of event types for various event samples. Flood frequency plots of observed annual maxima (62 years) for these two catchments (f and g).

A negative relative change in their distance implies that frequent floods are more similar to upper tails than ordinary events in term of event type distribution. This is observed in catchments with pronounced winter flood regime (western and central parts of Germany) (Figures 4a and 4b, small blue markers, pronounced seasonality). For most of these catchments no significant differences in the distribution of event types between frequent and upper tail floods were detected (Table S7 and Figures 4a and 4b, blue circles). This situation embodies the first detected trajectory, where the dominance of a certain process consistently increases from small to large events and sampling of frequent floods from ordinary events (e.g., by means of annual maxima or discharge thresholds) leads to prioritizing generation processes that are dominant for upper tails.

A positive relative change instead implies that frequent floods are more dissimilar to upper tails than ordinary events in terms of event type distributions. This is observed in catchments characterized by mixed flood seasonality (southern and eastern parts of Germany) (Figures 4a and 4b, large red markers, low seasonality). This situation corresponds to the second trajectory (detected in 23–27% of the catchments depending on the sample of frequent floods considered, Table S7), where the contribution of a certain process decreases from ordinary events to frequent floods, but dominates the generation of upper tails. Sampling of frequent floods from ordinary events in this case leads to prioritizing processes that are not dominant for upper tails. For at least a quarter of these catchments (24–44%, Table S7) the difference between event type distributions of frequent and upper tail floods is statistically significant (Figures 4a and 4b, red triangles). As expected, more catchments with statistically different distributions were detected (41–44%, Table S7; Figures 4c and S3, red triangles) if higher return periods (HQ15, HQ20) are considered.

These contrasting trajectories and their possible outcomes for predicting upper tails of flood frequency distributions are well illustrated by considering two exemplary catchments (Figures 4d–4g). In the Rodach catchment (located in Central Germany) the sample of frequent floods accurately represents the distribution of different events types among the largest events. In the Müglitz catchment (located in Eastern Germany), instead, the majority of frequent floods are categorized as *Rain-on-snow*, *Mixtures of rainfall and snowmelt* and *Rainfall*-induced events during *Wet* conditions (cold colors), whereas the largest floods are produced by *Rainfall*-induced events during *Dry* conditions (warm colors). In this particular mountainous catchment this is caused by the occurrence of rare but intense summer rainfall brought by Vb cyclones (Nied et al., 2014; Petrow et al., 2007), while the majority of annual floods is generated in winter and spring by less intense winter storms interacting with accumulated snowpack. As for most catchments (Figures 2 and 3), distribution of generation processes in the POT4 sample is very similar to that of MAF in this catchment (Figure 4d) and does not guarantee a more accurate sampling of generation processes for upper tails.

Most catchments attributed to the second trajectory (Figures 4a–4c, red markers) are located in Southern and Eastern Germany, areas characterized by mixed flood seasonality and affected by the above mentioned Vb cyclones (Hofstätter et al., 2016), and show a large variability in catchment sizes (Figure S4). The detected distributions of generation processes in these catchments (Figures 2 and 3) show that most annual floods triggered by to snowmelt processes (*Rain-on-snow*, *Mixture of rainfall and snowmelt*) or volume-dominated rainfall events occurring in *Wet* conditions during the winter-spring season (Beurton & Thielen, 2009), while *Intensity*-dominated rainfall events associated to specific meteorological phenomena like Vb cyclones (Hofstätter et al., 2016; Nied et al., 2014), blocking conditions (Grams et al., 2014) or convective storms (Bronstert et al., 2020) trigger rarer floods under *Dry* conditions in the summer-autumn season, especially in mountainous catchments.

This is likely to pose a problem for the homogeneity assumption required to fit theoretical distributions. Mixed distributions could be applied instead (Barth et al., 2019; DWA, 2012; England et al., 2018), but their usage with annual or partial duration series might result in highly uncertain estimates if only a few events of the types responsible for generating upper tails are sampled. Similar situations are likely to occur in other regions of the world where relatively infrequent climatic phenomena or conditions that are underrepresented in the annual or partial duration series dominate the upper tail while a large portion of floods is generated by processes of different nature. The most prominent examples of such phenomena, besides the mentioned Vb cyclones active in Central Europe (Hofstätter et al., 2016), are inland-penetrating atmospheric rivers in central Arizona, USA (Barth et al., 2017), tropical cyclones in the Eastern and Central United States (Smith et al., 2011; Villarini et al., 2014), as well as globally widespread organized convective thunderstorms (Doswell et al., 1996). Discrepancy of event types included in the samples of frequent and upper tail floods

support using ordinary events instead, as they capture the full variety of possible generation processes. In such catchments recently developed metastatistical and physically based approaches (Basso et al., 2016; Marani & Ignaccolo, 2015; Miniussi et al., 2020; Zorzetto et al., 2016) might be a viable alternative to expand information on rare generation processes using the whole sample of ordinary events.

Beside their relevance for the predictability of upper tail events, the revealed differences in the distribution of generation processes from small to larger events provide insights on the causes of observed disparate patterns of changes in small and large floods (Bertola et al., 2020). These alterations, possibly driven by the changing frequency of their corresponding generation processes (e.g., more frequent intensity-dominated events or less frequent rain-on-snow events), are likely to continue in the future (Kendon et al., 2014), thus altering the probabilistic structure of the extremes and modifying flood hazard in the affected catchments. Knowledge on dominant generation processes of the full range of runoff events might indicate directions toward suitable adaptations to future flood risk.

5. Conclusions

We examined transformations of generation processes from all identifiable runoff events to frequent (MAFs and peak-over-threshold events) and upper tail floods (i.e., with return period of at least 10 years) for a large set of mesoscale catchments, detecting substantial changes of the dominant processes along two contrasting trajectories. In regions where floods occur almost exclusively in winter the dominance of processes related to snowmelt consistently increases from small events to larger floods and the samples of frequent and upper tail floods show similar distributions of generation processes. In mountainous catchments characterized by frequent winter-spring floods related to snowmelt and occasional summer-autumn floods triggered by rare meteorological phenomena (e.g., Vb cyclones), instead, processes that dominate upper tails are not adequately represented in the sample of frequent floods. In these cases, conventional statistical approaches for the estimation of flood frequency from annual maxima and partial duration series, widely used in several fields of geoscience and in the engineering practice, should be applied with care. In fact, the prediction of upper tails in these catchments might remain highly uncertain even when mixed distributions are utilized. Future studies might advance in this direction by comparing the performance of traditional and emerging metastatistical approaches in catchments exhibiting different trajectories of process transformation and examine if unpredictable floods occur where the upper tail is dominated by rare processes underrepresented in the sample of frequent floods.

Data Availability Statement

The simulated soil moisture data are available at <https://www.ufz.de/index.php?en=41160> website. For providing the discharge data for Germany, we are grateful to Bavarian State Office of Environment (LfU, <https://www.gkd.bayern.de/de/fluesse/abfluss>) and Global Runoff Data Centre (GRDC) prepared by the Federal Institute for Hydrology (BfG, <https://www.bafg.de/GRDC>). Climatic data can be obtained from the German Weather Service (DWD; https://opendata.dwd.de/climate_environment/CDC/).

Acknowledgments

The financial support of the German Research Foundation (“Deutsche Forschungsgemeinschaft”, DFG) in terms of the research group FOR 2416 “Space-Time Dynamics of Extreme Floods (SPATE)” and grant number 421396820 “Propensity of rivers to extreme floods: climate-landscape controls and early detection (PREDICTED),” as well as the Helmholtz Centre for Environmental Research-UFZ, is gratefully acknowledged. We are grateful to the Editor Valeriy Ivanov and two anonymous reviewers for insightful comments that greatly improved the original manuscript. We thank Arianna Miniussi for useful discussions on metastatistical approaches. We are grateful to Rohini Kumar, Matthias Zink and Luis Samaniego for providing simulations of the mHM model.

References

- Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons.
- 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*, 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., Schirmer, M., & Botter, G. (2016). A physically based analytical model of flood frequency curves. *Geophysical Research Letters*, *43*, 9070–9076. <https://doi.org/10.1002/2016GL069915>
- Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The relative importance of different flood-generating mechanisms across Europe. *Water Resources Research*, *55*, 4582–4593. <https://doi.org/10.1029/2019WR024841>
- Bertola, M., Viglione, A., Lun, D., Hall, J., & Blöschl, G. (2020). Flood trends in Europe: Are changes in small and big floods different? *Hydrology and Earth System Sciences*, *24*(4), 1805–1822. <https://doi.org/10.5194/hess-24-1805-2020>
- Beurton, S., & Thieken, A. H. (2009). Seasonality of floods in Germany. *Hydrological Sciences Journal*, *54*(1), 62–76. <https://doi.org/10.1623/hysj.54.1.62>
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., et al. (2019). Changing climate both increases and decreases European river floods. *Nature*, *573*(7772), 108–111. <https://doi.org/10.1038/s41586-019-1495-6>
- Blöschl, G., Nester, T., Komma, J., Parajka, J., & Perdigão, R. A. P. (2013). The June 2013 flood in the Upper Danube Basin, and comparisons with the 2002, 1954 and 1899 floods. *Hydrology and Earth System Sciences*, *17*(12), 5197–5212. <https://doi.org/10.5194/hess-17-5197-2013>

- Bronstert, A., Crisologo, I., Heistermann, M., Ozturk, U., Vogel, K., & Wendi, D. (2020). Flash-floods: More often, more severe, more damaging? An analysis of hydro-geo-environmental conditions and anthropogenic impacts. *Climate change, hazards and adaptation options. climate change management* (pp. 225–244). https://doi.org/10.1007/978-3-030-37425-9_12
- Claps, P., & Laio, F. (2003). Can continuous streamflow data support flood frequency analysis? An alternative to the partial duration series approach. *Water Resources Research*, 39(8), 1216. <https://doi.org/10.1029/2002WR001868>
- DWA (2012). Ermittlung von Hochwasserwahrscheinlichkeiten - Estimation of flood probabilities, DWA-Guidelines M 522 (Tech. Rep.). Hennef, Germany.
- De Luca, P., Hillier, J. K., Wilby, R. L., Quinn, N. W., & Harrigan, S. (2017). Extreme multi-basin flooding linked with extra-tropical cyclones. *Environmental Research Letters*, 12(11), 114009. <https://doi.org/10.1088/1748-9326/aa868e>
- Doswell, C. A., Brooks, H. E., & Maddox, R. A. (1996). Flash flood forecasting: An ingredients-based methodology. *Weather and Forecasting*, 11(4), 560–581. [https://doi.org/10.1175/1520-0434\(1996\)011<0560:FFFAIB>2.0.CO;2](https://doi.org/10.1175/1520-0434(1996)011<0560:FFFAIB>2.0.CO;2)
- England, J. F. J., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas, W. O., J. r., Veilleux, A. G., et al. (2018). *Guidelines for determining flood flow frequency bulletin 17C*. Reston, Virginia: U.S. Geological Survey. <https://pubs.usgs.gov/tm/04/b05/tm4b5.pdf>
- Grams, C. M., Binder, H., Pfahl, S., Piaget, N., & Wernli, H. (2014). Atmospheric processes triggering the central European floods in June 2013. *Natural Hazards and Earth System Sciences*, 14(7), 1691–1702. <https://doi.org/10.5194/nhess-14-1691-2014>
- Hirschboeck, K. K. (1987a). Catastrophic flooding and atmospheric circulation anomalies. In L. Mayer & D. Nash (Eds.), *Catastrophic flooding* (pp. 23–56). Boston: Allen & Unwin.
- Hirschboeck, K. K. (1987b). Hydroclimatically-defined mixed distributions in partial duration flood series., *Hydrologic frequency modeling: Proceedings of the international symposium on flood frequency and risk analyses* (pp. 199–212). 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, 7086–7104. <https://doi.org/10.1002/2016WR019146>
- Hope, A. C. A. (1968). A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 30(3), 582–598. <https://www.jstor.org/stable/2984263>
- Kemter, M., Merz, B., Marwan, N., Vorogushyn, S., & Blöschl, G. (2020). Joint trends in flood magnitudes and spatial extents across Europe. *Geophysical Research Letters*, 47, e2020GL087464. <https://doi.org/10.1029/2020GL087464>
- Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., & Senior, C. A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, 4(7), 570–576. <https://doi.org/10.1038/nclimate2258>
- Klemes, V. (1993). Probability of extreme hydrometeorological events—A different approach. In *Extreme hydrological events: Precipitation, floods and droughts* (pp. 167–176). IAHS Publication no.213
- Kumar, R., Livneh, B., & Samaniego, L. (2013). Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. *Water Resources Research*, 49, 5700–5714. <https://doi.org/10.1002/wrcr.20431>
- Lima, C. H. R., Aghakouchak, A., & Lall, U. (2017). Classification of mechanisms, climatic context, areal scaling, and synchronization of floods: The hydroclimatology of floods in the Upper Paraná River basin, Brazil. *Earth System Dynamics*, 8(4), 1071–1091. <https://doi.org/10.5194/esd-8-1071-2017>
- 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>
- Mediero, L., Kjeldsen, T. R., Macdonald, N., Kohnova, S., Merz, B., Vorogushyn, S., et al. (2015). Identification of coherent flood regions across Europe by using the longest streamflow records. *Journal of Hydrology*, 528, 341–360. <https://doi.org/10.1016/j.jhydrol.2015.06.016>
- Merz, R., & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, 39(12), 1340. <https://doi.org/10.1029/2002WR001952>
- Miniussi, A., Marani, M., & Villarini, G. (2020). Metastatistical extreme value distribution applied to floods across the continental United States. *Advances in Water Resources*, 136. <https://doi.org/10.1016/j.advwatres.2019.103498>
- Nakamura, J., Lall, U., Kushnir, Y., Robertson, A. W., & Seager, R. (2013). Dynamical structure of extreme floods in the U.S. Midwest and the United Kingdom. *Journal of Hydrometeorology*, 14(2), 485–504. <https://doi.org/10.1175/JHM-D-12-059.1>
- Nied, M., Pardowitz, T., Nissen, K., Ulbrich, U., Hundecha, Y., & Merz, B. (2014). On the relationship between hydro-meteorological patterns and flood types. *Journal of Hydrology*, 519, 3249–3262. <https://doi.org/10.1016/j.jhydrol.2014.09.089>
- Petrow, T., Merz, B., Lindenschmidt, K.-E., & Thielen, H. (2007). Aspects of seasonality and flood generating circulation patterns in a mountainous catchment in south-eastern Germany. *Hydrology and Earth System Sciences Discussions*, 4(2), 589–625. <https://doi.org/10.5194/hessd-4-589-2007>
- Ralph, F. M., Neiman, P. J., Wick, G. A., Gutman, S. I., Dettinger, M. D., Cayan, D. R., & White, A. B. (2006). Flooding on California's Russian River: Role of atmospheric rivers. *Geophysical Research Letters*, 33, L13801. <https://doi.org/10.1029/2006GL026689>
- 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>
- 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, W05544. <https://doi.org/10.1029/2011WR011187>
- Rossi, F., Fiorentino, M., & Versace, P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resources Research*, 20(7), 847–856. <https://doi.org/10.1029/WR020i007p00847>
- Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46, 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, 7959–7976. <https://doi.org/10.1002/2015WR017326>
- 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, 6510–6542. <https://doi.org/10.1029/2018WR022539>
- Smith, J. A., Villarini, G., & Baeck, M. L. (2011). Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the Eastern United States. *Journal of Hydrometeorology*, 12(2), 294–309. <https://doi.org/10.1175/2010JHM1242.1>
- Stein, L., Pianosi, F., & Woods, R. (2019). Event-based classification for global study of river flood generating processes. *Hydrological Processes*, 34, 1514–1529. <https://doi.org/10.1002/hyp.13678>
- 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, e2019WR026951. <https://doi.org/10.1029/2019WR026951>

- 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*, 7711–7732. <https://doi.org/10.1029/2018WR022587>
- Villarini, G., Goska, R., Smith, J. A., & Vecchi, G. A. (2014). North atlantic tropical cyclones and U.S. flooding. *Bulletin of the American Meteorological Society*, *95*(9), 1381–1388. <https://doi.org/10.1175/BAMS-D-13-00060.1>
- Villarini, G., & Smith, J. A. (2010). Flood peak distributions for the eastern United States. *Water Resources Research*, *46*, 1–17. <https://doi.org/10.1029/2009WR008395>
- Villarini, G., Smith, J. A., Serinaldi, F., & Ntelekos, A. A. (2011). Analyses of seasonal and annual maximum daily discharge records for central Europe. *Journal of Hydrology*, *399*(3–4), 299–312. <https://doi.org/10.1016/j.jhydrol.2011.01.007>
- Zink, M., Kumar, R., Cuntz, M., & Samaniego, L. (2017). A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty. *Hydrology and Earth System Sciences*, *21*(3), 1769–1790. <https://doi.org/10.5194/hess-21-1769-2017>
- Zorzetto, E., Botter, G., & Marani, M. (2016). On the emergence of rainfall extremes from ordinary events. *Geophysical Research Letters*, *43*, 8076–8082. <https://doi.org/10.1002/2016GL069445>