A Process-Based Framework to Characterize and Classify Runoff Events: The Event Typology of Germany

Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR026951

A Process-Based Framework to Characterize and Classify Runoff Events: The Event Typology of Germany

L. Tarasova[1], S. Basso[1], D. Wendi[2], A. Viglione[3,5], R. Kumar[4], and R. Merz[1]

[1]Department Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale), Germany, [2]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam, Germany, [3]Institute of Hydraulic Engineering and Water Recourses Management, Vienna University of Technology, Vienna, Austria, [4]Department of Computational Hydrosystems, Helmholtz Centre for Environmental Research, Leipzig, Germany, [5]Now at: Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Turin, Italy

Key Points:
- A new process-based framework for characterizing runoff events of diverse sizes is proposed and applied to a wide set of catchments
- Novel indicators combining the space–time dynamics of event precipitation and catchment state are used for classification
- The derived event typology captures distinct shapes of event hydrographs even if streamflow data are not directly used for classification

Supporting Information:
- Supporting Information S1

Correspondence to:
L. Tarasova,
larisa.tarasova@ufz.de

Citation:
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

Received 13 DEC 2019
Accepted 13 APR 2020
Accepted article online 16 APR 2020

Abstract This study proposes a new process-based framework to characterize and classify runoff events of various magnitudes occurring in a wide range of catchments. The framework uses dimensionless indicators that characterize space–time dynamics of precipitation events and their spatial interaction with antecedent catchment states, described as snow cover, distribution of frozen soils, and soil moisture content. A rigorous uncertainty analysis showed that the developed indicators are robust and regionally consistent. Relying on covariance- and ratio-based indicators leads to reduced classification uncertainty compared to commonly used (event-based) indicators based on absolute values of metrics such as duration, volume, and intensity of precipitation events. The event typology derived from the proposed framework is able to stratify events that exhibit distinct hydrograph dynamics even if streamflow is not directly used for classification. The derived typology is therefore able to capture first-order controls of event runoff response in a wide variety of catchments. Application of this typology to about 180,000 runoff events observed in 392 German catchments revealed six distinct regions with homogeneous event type frequency that match well regions with similar behavior in terms of runoff response identified in Germany. The detected seasonal pattern of event type occurrence is regionally consistent and agrees well with the seasonality of hydroclimatic conditions. The proposed framework can be a useful tool for comparative analyses of regional differences and similarities of runoff generation processes at catchment scale and their possible spatial and temporal evolution.

1. Introduction

Understanding event runoff generation processes at catchment scale is crucial for gaining insights about event-to-event variability of runoff response and solutes export from catchments (McGlynn & McDonnell, 2003). This knowledge is instrumental in a wide variety of applications ranging from hydrologically sound flood frequency predictions (Merz & Blöschl, 2008) to the derivation of effective indicators for water quality assessment at management (i.e., catchment) scale (Minaudo et al., 2019). In a changing world, understanding generation processes of runoff events is pivotal for elucidating mechanisms behind intensification and shifts in the seasonality of river flows (Blöschl et al., 2017), trends in long-term flood series (Slater & Wilby, 2017), and possible changes in the magnitude of flood hazard in the future (Turkington et al., 2016). Therefore, the development of a general framework for understanding and comparing typical runoff event generation processes and event runoff response at large (i.e., from catchment to regional) scales is an important task. Such a framework would be a valuable tool to link a wide variety of event runoff generation processes to individual catchment responses at the event time scale. It would also enable a detailed analysis of the transformation of dominant processes from small to large runoff events, thus unveiling possible differences between processes responsible for the occurrence of unexpected extreme and ordinary events (e.g., Rogger et al., 2012; Smith et al., 2018).

Classification frameworks are essential for reducing information into a manageable amount of classes by grouping together similar objects, and are widely used in hydrological science (e.g., Knoben et al., 2018; Sawicz et al., 2014; Sivakumar & Singh, 2012; Wagener et al., 2007). Process-based classification of hydrological events, such as river floods and hydrological droughts, also gained momentum in the last decades (e.g., Merz & Blöschl, 2003; van Loon & van Lanen, 2012) as it provides a basis for understanding the observed

variety of hydrological behaviors and their causes. Large scale studies dealing with process-based classification of hydrological events usually focus on large runoff events (e.g., maximum annual floods). Applications at regional (e.g., Diezig & Weingartner, 2007; Hirschboeck, 1987; Nied et al., 2014; Sikorska et al., 2015), country-wide (Berghuijs et al., 2016; Merz & Blöschl, 2003), continental (Berghuijs et al., 2019), and even global scales (Stein et al., 2019) exist. A comprehensive overview of existing causative flood classification approaches is given by Tarasova et al. (2019). Yet a consistent framework for process-based characterization and classification of runoff events that, beside flood events, analyzes a wide range of runoff magnitudes and is suitable for varied catchment sizes is still lacking.

Although the above-mentioned flood classifications could theoretically be used to sort generation processes of a larger sample of runoff events (i.e., including small events), they mostly adopt spatially and temporally lumped characteristics of precipitation events (e.g., catchment- and event-averaged amount of rainfall), often neglect pre-event wetness states, and are therefore unsuitable to capture the spatiotemporal interactions of rainfall and soil moisture (Tarasova et al., 2019). This is, however, known to be a key driver of runoff generation. For example, Freyberg et al. (2014) found that only a small fraction (1–26%) of the total catchment area generates streamflow during rainfall events in a small pre-alpine catchment. Similarly, McGlynn and McDonnell (2003) showed that contribution from riparian zones dominates runoff for smaller events, while hillslope runoff generation increases for larger events in a New Zealand catchment. According to Seo et al. (2012) and Mei et al. (2014), space–time dynamics might also be crucial to correctly mimic magnitude and timing of flood events.

In recent decades, high-resolution observed and simulated hydrometeorological data sets became available (e.g., Zink et al., 2017), thus providing tools to analyze variations of runoff generation processes at catchment scale and their drivers (Viglione, Chirico, Komma, et al., 2010; Woods & Sivapalan, 1999). This advances aim to bridge the gap between small scale studies that investigate runoff generation processes by means of detailed observations (e.g., Jencso et al., 2009; Rinderer et al., 2019; Tromp-van Meerveld & McDonnel, 2006) and large scale studies which offer predictions of future changes of these processes purely based on modeling approaches (e.g., Donnelly et al., 2017; Gosling et al., 2017).

Another drawback of most existing flood classification frameworks is the use of classification thresholds that are only valid for specific regions. This limits the transferability of the frameworks to other regions and hinders their application at larger (e.g., global) scales. To overcome this issue, Stein et al. (2019) introduced thresholds expressed as catchment-specific percentiles of the considered indicators. However, the application of percentiles might be less justified when a wider sample of runoff events is considered and variability of indicators based on absolute values (e.g., maximum precipitation intensity and volume of precipitation events) are much higher than for samples that include only maximum annual flood events. Therefore, it is important to develop a framework that relies primarily on dimensionless indicators and is transferable across climatic conditions and to catchments and events of different sizes (Tarasova et al., 2019).

The quality of data-based characterization and classification schemes largely depends on the quality of the available data and the robustness of the applied classification methods with respect to several aspects of uncertainty (Tarasova et al., 2019). In fact, classification results are sensitive to the uncertainty of input data from different sources (e.g., different observed temperature or precipitation data (Kampf & Lefsky, 2016), various soil moisture, and snowmelt simulations (Stein et al., 2019), to the choice of indicators used to characterize event types and to the values of thresholds applied to attribute events to different classes (Sikorska et al., 2015; Stein et al., 2019). All these aspects must be examined to provide a robust characterization framework and guidelines for selection of input data, indicators, and their corresponding thresholds.

Finally, existing automated event classification frameworks (e.g., Sikorska et al., 2015; Stein et al., 2019) utilized predefined classification trees, whose structures were devised based on hydrological reasoning about possible runoff generation processes. The underlying assumption is that similar hydrometeorological and catchment state conditions within an individual catchment (i.e., the blocks of the classification tree) result in similar hydrological responses (e.g., Sikorska et al., 2015). Although this assumption is rarely tested (Tarasova et al., 2019), proving its soundness would confirm the validity of the classification and lay the foundation for transferring classification results to ungauged locations.

This study aims to amend gaps and limitations of the existing classification frameworks discussed above. First, we propose a new multilayer framework (section 3.1) for process-based characterization of runoff events of various sizes (i.e., a large numbers of events are analyzed in addition to peaks over thresholds or events with large magnitude and recurrence intervals). The proposed framework relies on a set of transferable dimensionless indicators which describe the nature of the inducing meteorological event, its spatiotemporal dynamics and the interaction of precipitation and catchment wetness state, and does not use streamflow data. We test the effects of the uncertainty of input data and chosen indicators, as well as their sensitivity to the choice of classification thresholds, and compared their performance to that of traditional indicators relying on absolute values and lumped characteristics of inducing events and catchment states (sections 3.3 and 4.1). By using the proposed characterization framework, we then derive an event typology for Germany (section 3.2). We investigate the adequacy of the hypothesized runoff generation processes by testing whether the derived event types group events with similar hydrograph dynamics (section 3.4). Finally, we examine spatial and temporal (seasonal) patterns of event type occurrence in a wide set of German catchments (section 3.5) and thus reveal spatial heterogeneity and temporal variability of dominant catchment-scale streamflow generation processes across Germany (section 5).

2. Data and Study Area

The event classification is performed on the set of 183,955 runoff events occurred in 392 German catchments (Figure 1) during the course of 22 years (1979–2002) of observations reported by Tarasova, Basso, Poncelet, et al. (2018). Catchment sizes range from 36 to 23,843 km^2. The median area of catchments in the data set is 494 km^2. Runoff events were identified from continuous hydroclimatic time series using the procedure illustrated in Tarasova, Basso, Zink, et al. (2018). First we separated base and quick flows using a simple smoothing algorithm (IH, 1980); then we attributed precipitation to each quick flow event if it occurred within the seasonal lag time of the respective catchment (Mei & Anagnostou, 2015); finally we identified single-peak components of multiple-peak events by using a set of iteratively adjusted refinement thresholds. The iterative procedure is based on an equality test between the two distributions of event runoff coefficients derived from a reference single-peak and a refined single-peak event groups. For a detailed description of the event separation procedure, we refer the reader to Tarasova, Basso, Zink, et al. (2018).

The initial rainfall-runoff event attribution described above essentially defines the lag time from the starting point of the considered runoff event backwards within which rainfall events are searched at each individual location (i.e., gauge). However, inducing events (i.e., precipitation) in the headwaters might start earlier than the observed streamflow reactions at the outlet of large river networks (Diederen et al., 2019). Therefore, precipitation events considered for the same runoff event at upstream and downstream locations might differ when the start of the precipitation event lays inside the search distance defined by the lag time for the upstream location, but is outside of this distance for the downstream location (see supporting information Text S1 and Figure S1 for a visual explanation). To remedy this issue in the case of nested catchments, individual runoff events were organized along the river network they belong to and the starting points of runoff events at downstream locations were adjusted in a way that the earliest starting date among upstream locations was assigned to all the downstream locations. These starting dates of runoff events were then used to search for the corresponding precipitation event for each individual catchment according to the seasonal median lag time as described above (see also supporting information, Figure S1b). The search time range backwards was bounded by the starting date of the precipitation event at the upstream location. The above-described procedure ensures that no precipitation is accidentally discarded in larger catchments and allows for characterizing nested catchments consistently. We used the identified beginning and end dates of the attributed precipitation event to subset daily gridded products of hydrometeorological variables and compute event-based indicators according to the proposed characterization framework.

We used daily gridded rainfall data (1-km grid) from the REGNIE (Regionalisierte Niederschlagshöhen) data set (Rauthe et al., 2013) provided by the German Weather Service, and snowmelt and soil moisture time series (on a 4-km grid) simulated by the mesoscale Hydrologic Model (mHM; Samaniego et al., 2010; Kumar et al., 2013) along with air temperature interpolated using external drift kriging (Zink et al., 2017). For uncertainty analysis and robustness check, we additionally used a precipitation data set interpolated by means of external drift kriging (as reported in Zink et al. [2017]), a temperature data set interpolated by using the

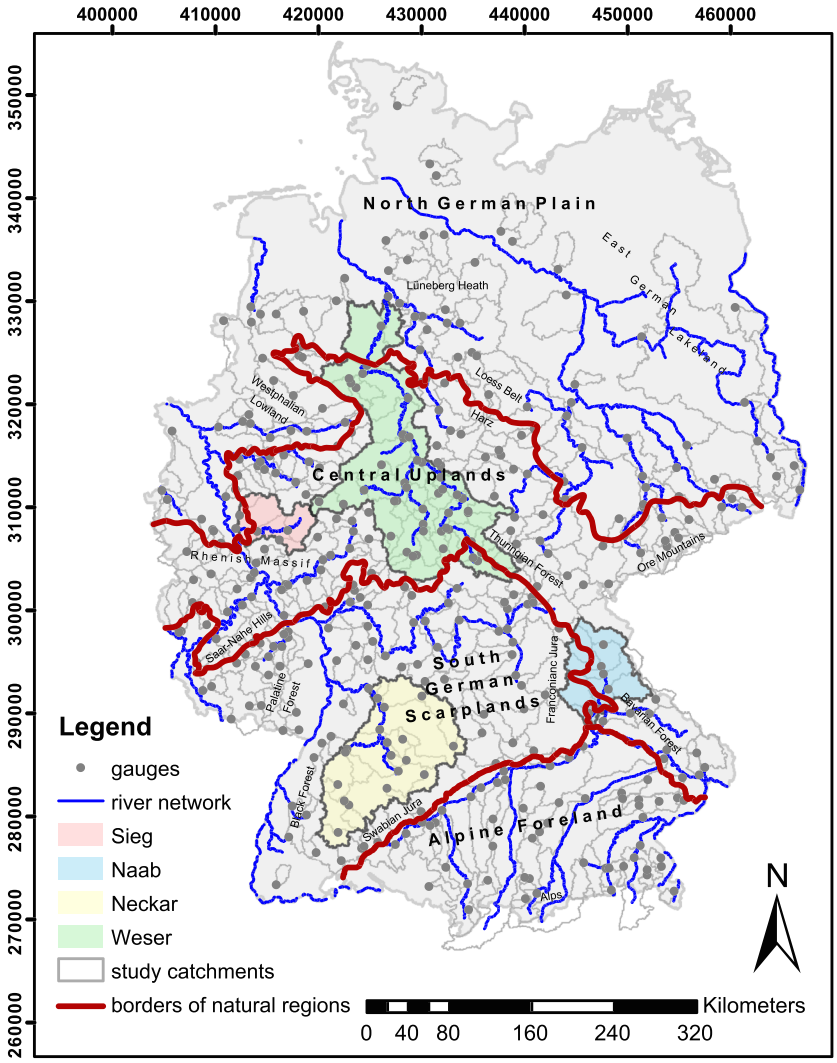

Figure 1. Study catchments (thin black outline) and four test river basins: the Sieg, Naab, Neckar, and Weser Rivers. Gray points indicate location of gauging stations. Red lines indicate the borders of four main German natural regions: the North German Plain, Central Uplands, South German Scarplands, and Alpine Forelands.

inverse distance method, 100 equifinal realizations (i.e., 100 different simulations resulted from using 100 different representative parameter sets) of the mHM (Zink et al., 2017), as well as simulations of the alternative distributed regionally calibrated hydrological model SALTO (Merz et al., 2020), which has a structure similar to the conceptual Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1995).

The uncertainty analysis (section 3.3) was performed in 75 catchments (19% of all study catchments) belonging to four large river basins in Germany (i.e., the Sieg, Naab, Neckar, and Weser Rivers, Figure 1). Overall, it considered 38,289 individual runoff events (21% of all runoff events in the data set). The size of catchments ranges from 218 to 2,825 km^2 for the Sieg River (4 catchments, 2,462 events), from 588 to 5,096 km^2 for the Naab River (7 catchments, 2,937 events), from 135 to 12,670 km^2 for the Neckar River (19 catchments, 10,711 events), and from 36 to 21,583 km^2 for the Weser River (45 catchments, 22,179 events). The selected catchments represent the whole range of catchments sizes and hydroclimatic conditions in the German data set.

Similarity of hydrograph dynamics for each event pair in each individual catchment (section 3.4) was calculated for events that are at least 5 days long (see supporting information, Text S3). In total, 100,750 events (55% of all runoff events) were analyzed through this procedure.

3. Methods

3.1. A Process-Based Framework for Characterization of Runoff Events

Here we propose a multilayer framework for process-based characterization of runoff events based on a consistent set of indicators which capture space–time dynamics of rainfall and snowmelt, spatial organization of antecedent catchment wetness states, snow cover and soil freezing conditions, as well as the spatial interaction of precipitation and soil moisture within catchments (Figure 2a).

The idea underlying this characterization approach is that runoff events can be categorized by a set of indicators, each describing a single aspect of the transformation of rainfall into runoff. For example, one indicator characterizes the spatial extent of the precipitation field during the event, while another describes the temporal evolution of the precipitation rate during the event. The set of observed runoff events is split at each layer based on crisp predefined thresholds (e.g., Brunner et al., 2017) applied to the layer-specific dimensionless indicators (Table 1), and runoff events are labeled accordingly (Figure 2a). The combination of states categorized at each layer composes the overall character of runoff generation processes during a given event.

In the proposed framework, streamflow data are not utilized for characterization, but only used to define start and end date of the events (see section 2). In addition, we used streamflow data to validate the proposed classification (section 3.4) by analyzing the similarity of event hydrographs of the derived event types (section 3.2).

In the first layer of the characterization framework, we analyzed the nature of the meteorological inducing event. An event was considered a pure snowmelt (*Snowmelt*) event if the volumetric ratio of catchment-averaged event snowmelt and total precipitation is larger than a prescribed threshold value (i.e., 0.95). Such high threshold value was selected to identify events that are solely triggered by melting processes and where the contribution of rainfall is minimal. This choice as well as the choice of later described thresholds is unavoidably subjective. The sensitivity of all thresholds was therefore examined in a thorough uncertainty analysis (see sections 3.3 and 4.1). Similarly to snowmelt events, we used the volumetric ratio between catchment-averaged event rainfall and total precipitation to identify purely rainfall events (Table 1 and Figure 2a). Among their mixtures (i.e., when the volumetric portion of both rainfall and snowmelt in the total precipitation event volume was higher than 0.05, *Mixture of Rainfall and Snowmelt*), *Rain-on-snow* events were identified by using the spatial covariance of pre-event snow water equivalent and event rainfall volume normalized by the product of their average values (Table 1). Spatial covariance was calculated as a grid-wise covariance of two gridded variables for a certain time step (i.e., in this case positive values of spatial covariance indicate that, within a catchment more rainfall fell on the areas where more snow water equivalent was accumulated before the beginning of the event).

Rain-on-ice (e.g., Wendi et al., 2019) occurs when rain falls on frozen surfaces (e.g., frozen soils) and was here distinguished from rainfall events using the spatial covariance of event-averaged rainfall volume and pre-event degree of soil freezing (SF_0) defined as $SF_0 = \sum_{i=0}^{n} |T_i|$, where T is air temperature (°C) of a snow-free pixel and n is the number of days with air temperature below −2 °C during the 5 days preceding the start of a rainfall event (Flerchinger et al., 2005).

We also considered the temporal organization of the inducing event by analyzing the coefficient of variation in time of the catchment-averaged precipitation rate and the ratio of maximum intensity to total volume of the inducing event (rainfall, snowmelt, or mixture depending on the nature of the inducing event; Table 1). For this ratio-based indicator the classification threshold was set to its mean value (i.e., 0.5). When the ratio of maximum intensity to total volume of the inducing event is higher than 0.5, it indicates that more than 50% of event precipitation volume occurred in a single time step. Therefore, with regard to their temporal organization events can be intensity-dominated or volume-dominated (Figure 2a), essentially corresponding to different runoff generation mechanisms (i.e., infiltration-excess and saturation-excess; Dunne, 1978; Horton, 1933).

Moreover, we analyzed the space–time organization of the inducing event using the coefficient of variation in space of precipitation volume and event-averaged spatial covariance of precipitation fields between consecutive time steps (Table 1). Using these two indicators, an event can be characterized as *Local Steady*, *Local Unsteady*, *Extensive Steady*, and *Extensive Unsteady* (Figure 2a). For a *Local* event precipitation, volume is

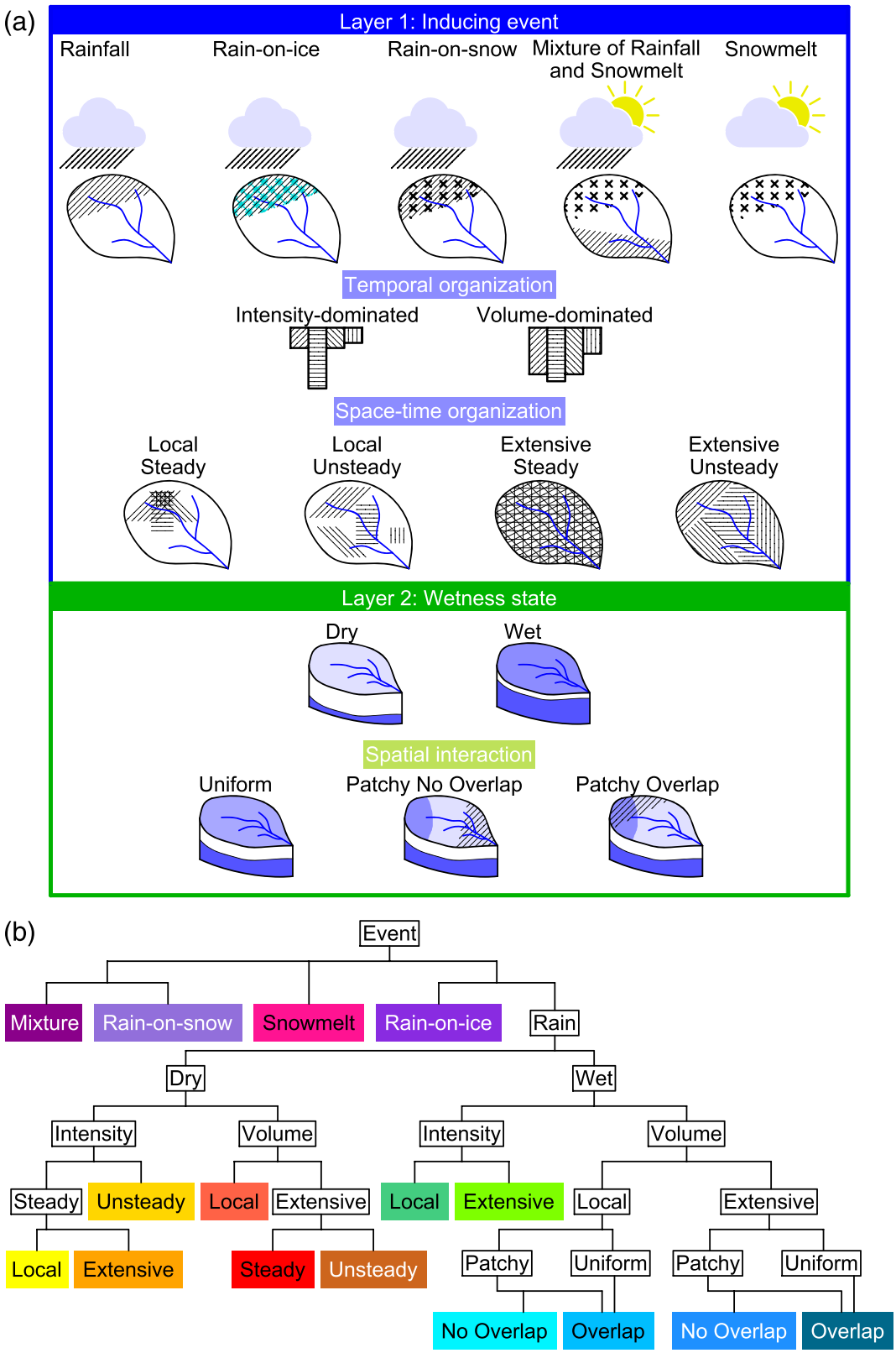

Figure 2. (a) A multilayer framework for process-based characterization and categorization of runoff events. Indicators and categorization thresholds used for each layer of the framework are summarized in Table 1. (b) A decision tree to perform hierarchical classification of runoff events. The amount of possible combinations (i.e., event types) was reduced by lumping together hydrologically similar (i.e., by hypothesizing possible runoff generation processes at catchment scale that correspond to each type of runoff events, see Table 2) and rarely occurring combinations.

Water Resources Research

10.1029/2019WR026951

Table 1

Indicators for Each Characterization Layer (section 3.1 and Figure 2a) and Alternative Indicators Used for Robustness Check of the Proposed Characterization Framework (section 3.3)

Layer	Selected indicator (dimensionless)	Expression	Selected thresholds and their possible intervals	Alternative indicators and their units	Thresholds and intervals for alternative indicators	Performed split
a) Inducing event	Ratio of event snowmelt volume and total event precipitation volume	$\frac{M_{x,y,t}}{P_{x,y,t}}$	0.95; [0, 1]	Absolute values of snowmelt volume (mm)	Q_1[a]; (0, +∞)	*Snowmelt* vs. any other type of inducing event
	Ratio of event rainfall volume and total event precipitation volume	$\frac{R_{x,y,t}}{P_{x,y,t}}$	0.95; [0, 1]	Absolute values of rainfall volume (mm)	Q_1; (0, +∞)	*Rainfall* or *Rain-on-ice* vs. *Mixture* or *Rain-on-snow*
	Normalized spatial covariance of event-averaged snow cover and rainfall	$\frac{\mathrm{cov}_{x,y} SWE_t R_t}{SWE_{x,y,t} \times R_{x,y,t}}$	0; (−∞, +∞)	Extent of snow covered areas (−)	0.5; (0, 1)	*Rain-on-snow* vs. *Mixture*
	Normalized spatial covariance of pre-event level of soil freezing and event rainfall volume	$\frac{\mathrm{cov}_{x,y} SFt_0 R_t}{SF_{x,y,t} \times R_{x,y,t}}$	0; (−∞; +∞)	Portion of overlap of areas with frozen soils and rainfall field (−)	0.25; (0, 1)	*Rain-on-ice* vs. *Rainfall*
Temporal organization	Temporal coefficient of variation of precipitation rate	$\frac{\sqrt{\mathrm{var}_t\left(P_{x,y}\right)}}{P_{x,y,t}}$	1; (0, +∞)	Duration of precipitation event (days)	Q_2; (0, +∞)	*Volume-dominated* vs. *Intensity-dominated*
	Ratio of maximum precipitation rate and total precipitation volume	$\frac{\max\left(P_{x,y}(t)\right)}{P_{x,y,t}}$	0.5; (0, 1]	Maximum precipitation rate (mm/d)	Q_3; (0, +∞)	
Space–time organization	Spatial coefficient of variation of precipitation event volume	$\frac{\sqrt{\mathrm{var}_{x,y}(P_t)}}{P_{x,y,t}}$	Q_2; (0, +∞)	Extent of precipitation field (−)	0.5; (0, 1)	*Local* vs. *Extensive*
	Mean normalized spatial covariance of precipitation rates between consecutive time steps	$\frac{1}{t+1} \times \sum_t \frac{\mathrm{cov}_{x,y} Pt, Pt+1}{P_{x,y,t} \times P_{x,y,t}}$	0; (−∞; +∞)	Portion of overlap of precipitation fields during consecutive time steps (−)	0.25; (0, 1)	*Steady* vs. *Unsteady*
b) Wetness state	Catchment-averaged antecedent soil moisture	$SM_{x,y}(t_0)$	max(κ) [b]; [0, 1]	30-day antecedent rainfall and snowmelt index (mm)	Q_2; (0, +∞)	*Wet* vs. *Dry*
Spatial interaction of precipitation and soil moisture	Spatial coefficient of variation of antecedent soil moisture	$\frac{\sqrt{\mathrm{var}_{x,y}\left(SM_{t_0}\right)}}{SM_{x,y,t}}$	Q_2; (0, +∞)	Extent of wet areas (−)	0.5; (0, 1)	*Uniform* vs. *Patchy*
	Normalized spatial covariance of precipitation event volume and antecedent soil moisture	$\frac{\mathrm{cov}_{x,y} SMt_0 P_t}{SM_{x,y,t} \times P_{x,y,t}}$	0; (−∞; +∞)	Portion of overlap of wet areas and precipitation field (−)	0.25; (0, 1)	*Patchy No Overlap* vs. *Patchy Overlap*

Note. Input data are raster-based, and for each grid (x,y) and time step t during an event, the following variables are provided: event rainfall rate $R(x,y,t)$ (mm/day); event snowmelt rate $M(x,y,t)$ (mm/day); total event precipitation rate (i.e., sum of rainfall and snowmelt rates) $P(x,y,t)$ (mm/day); snow water equivalent $SWE(x,y,t)$ (mm). Antecedent soil moisture $SM(x,y,t_0)$ (−) and antecedent degree of soil freezing $SF(x,y,t_0)$ (°C; see section 3.1) are only provided for the pre-event time step t_0. As an example of notation catchment-averaged (in space) rainfall rate, R variable for time step t is noted as $R_{x,y}(t)$, while event-averaged (in time) raster of event rainfall volumes is referred as $R_t(x,y)$. $R_{x,y,t}$ is a catchment-averaged variable corresponding to the event rainfall total volume precipitated in a catchment. All other variables are noted in the same fashion. Here we adopt the notation of analytical framework of Viglione, Chirico, Woods, et al. (2010), that quantifies space–time dynamics of runoff events and refer to spatial covariance as $\mathrm{cov}_{x,y}$. Similarly, we indicate spatial and temporal variances as $\mathrm{var}_{x,y}$ and var_t. All covariance-based indicators are normalized by the product of the average values of the corresponding variables.

[a]Q_1,Q_2,Q_3 are quartile of a variable that correspond to the 25th percentile, median, and 75th percentile, respectively. [b]κ is the curvature of a nonlinear function that describes the relation between event runoff coefficients and simulated soil moisture (see section 3.1 for details).

unevenly distributed in the catchment and mostly concentrated in a small portion of it, thus possibly hinting towards local runoff generation. For an *Extensive* event, precipitation volume is instead evenly distributed in the catchment, thus possibly suggesting extensive runoff generation. For a *Steady* event, daily precipitation volume mostly occurs in the same part of the catchment during consecutive days of a wet spell, with higher likelihood of having saturated areas in this part of the catchment. On the contrary, the location of the daily precipitation volumes during consecutive days of a wet spell varies (i.e., the centroid of the precipitation moves) for an *Unsteady* event, thus suggesting that different portions of the catchment are wetted.

The second characterization layer sorts events according to their corresponding pre-event catchment wetness state (Figure 2a). We labeled event conditions as *Wet* or *Dry* based on catchment-averaged antecedent soil moisture. By using simulated soil moisture, we inevitably introduce model-specific bias into the characterization framework. Simulated soil moisture essentially shows the current state of the soil moisture storage unit according to accounting schemes that are widely used in computational hydrology (e.g., Bergström, 1995). If the nonlinear storage-discharge relationship used in the model differs in different catchments (e.g., supporting information Figure S2), using percentiles of soil moisture as threshold to define wet and dry states might be inappropriate. To properly account for the nonlinear behavior of the soil moisture storage reservoir, we used the measure of maximum curvature κ (see e.g., Rogger et al., 2013) of a fitted exponential function that describes the nonlinear relation between event runoff coefficient (rc) and antecedent soil moisture (sm; Tarasova, Basso, Zink, et al., 2018 and Figure S2) as a classification threshold. For soil moisture states below the point of maximum curvature, event runoff coefficients tend to increase slowly with increasing soil moisture (representing dry catchment states), while beyond this point event runoff coefficients increase rapidly with soil moisture (representing wet catchment states; Rogger et al., 2013). Curvature at each point of the function $rc = f(sm)$ is calculated as $\kappa = \frac{\left|f(sm)''\right|}{\left[1+\left(f(sm)'\right)^2\right]^{\frac{3}{2}}}$, and its maximum value is calculated for each catchment in the data set. More details on the threshold used to define the wetness state of a catchment are provided in supporting information Text S2.

We also characterized the spatial interaction between inducing event and antecedent catchment wetness state. We used the spatial coefficient of variation as an indicator of *Uniform* or *Patchy* spatial organization of the soil moisture (Table 1), and in case of detected spatial inhomogeneity, the spatial covariance of event precipitation volume and pre-event soil moisture state (Table 1) was used to detect spatial overlapping of saturated areas and area wetted by the current precipitation event. In terms of spatial interaction, each event can be characterized as *Uniform*, *Patchy Overlap*, or *Patchy No Overlap* (Figure 2a).

For all covariance-based indicators, we set the threshold at zero, because positive values of spatial covariance correspond to the occurrence of spatial overlap of the investigated hydrometeorological variables. For indicators based on the temporal coefficient of variation the threshold was set to 1 (i.e., the case when the mean is equal to the standard deviation) for all catchments. The spatial coefficient of variation instead showed dependence on catchment size, with larger catchments always having higher values of the spatial coefficient of variation compared to smaller ones. Therefore, we selected the second quartile (i.e., the median) as threshold for the indicators based on the spatial coefficient of variation (Table 1).

3.2. Hierarchical Event Classification: The Event Typology of Germany

A goal of this study is to investigate the occurrence of different runoff generation processes, their spatial patterns and seasonality in Germany. The combination of states categorized at each layer of the proposed characterization framework (e.g., *Rain.Intensity.Local.Steady.Dry.Uniform*; section 3.1) thoroughly describes the runoff generation process occurring during a given event. Not every combination of possible occurred for each catchment in the data set (e.g., almost all rain-on-snow events are volume-dominated events occurring on wet conditions). Therefore, we reduced the amount of possible combinations of mechanisms by hypothesizing plausible runoff generation processes at catchment scale (Table 2), and lumping together similar and rarely occurring combinations during the hierarchical classification (Figure 2b).

The classification tree derived in this study is referred in the following as event typology for Germany (Figure 2b and Table 2). Notice that the complexity of the characterization framework can be adapted for the specific task at hand, for instance if a certain sample size is required for each event type (e.g., for deriving mixed distributions for flood frequency analysis; e.g., House & Hirschboeck, 1997), or if certain

Table 2
Event Types and Corresponding Hypothesized Runoff Generation Processes at Catchment Scale

Event type	Hypothesized runoff generation processes
Snowmelt	Radiation-induced snowmelt (usually *Wet.Volume.Steady*)
Rain-on-ice	Frozen soils prevent infiltration of rainfall (usually *Wet.Intensity.Uniform*)
Rain-on-snow	Several possible ways of runoff generation ranging from situations when snowpack prevents infiltration of rainfall or instead either stores substantial portion of rainfall water or is degraded due to rainfall-induced snowmelt (usually *Wet.Volume.Extensive.Steady.Uniform*)
Mixture of Rainfall and Snowmelt	Radiation-induced snowmelt and simultaneous in time but not in space rainfall (usually *Wet.Volume.Local.Unsteady*)
Rain.Wet.Intensity.Local	Local runoff generation; pre-event saturation or infiltration excess; possible connectivity[a]
Rain.Wet.Intensity.Extensive	Extensive runoff generation; pre-event saturation or infiltration excess; possible connectivity
Rain.Wet.Volume.Local.No.Overlap	Local runoff generation, event-fed saturation, possible connectivity
Rain.Wet.Volume.Local.Overlap	Local runoff generation, pre-event saturation, established connectivity
Rain.Wet.Volume.Extensive.No.Overlap	Extensive runoff generation, event-fed saturation, possible connectivity
Rain.Wet.Volume.Extensive.Overlap	Extensive runoff generation, pre-event saturation, established connectivity
Rain.Dry.Intensity.Local.Steady	Local runoff generation; possible infiltration excess or event-fed saturation; no connectivity
Rain.Dry.Intensity.Unsteady	Local runoff generation; possible infiltration excess; no connectivity
Rain.Dry.Intensity.Exrtensive.Steady	Possible extensive runoff generation; possible event-fed saturation and infiltration excess; possible event-induced connectivity
Rain.Dry.Volume.Local	Local event-fed saturation; no connectivity
Rain.Dry.Volume.Extensive.Steady	Extensive event fed-saturation; possible event-induced connectivity
Rain.Dry.Volume.Extensive.Unsteady	Extensive runoff generation; possible pre-event saturation; possible event-induced connectivity

Note. Due to the small number of events induced by snowmelt, rain-on-ice, rain-on-snow and mixtures of rainfall and snowmelt compared to pure rainfall-induced events, no further splitting was performed (see Figure 2b), but their most frequent attributes are noted in brackets.
[a]Here by connectivity, we refer to a catchment state when a direct linkage between runoff source areas and the stream is established (Jencso et al., 2009; Rinderer et al., 2019).

characteristics are crucial despite their rarity (as in the case of sediment transport by intensity-dominated events (Basso et al., 2015)).

3.3. Uncertainty Analysis and Robustness Check of the Proposed Framework

The proposed characterization framework (section 3.1) relies heavily on the quality of the observed and simulated input data. Therefore, we performed a rigorous uncertainty analysis and robustness check. We analyzed two different uncertainty aspects in the case of 75 test catchments belonging to four river basins that represent the variety of climatic conditions, catchment characteristics and sizes in the data set (see section 2). First we quantified the input data uncertainty by measuring the discrepancy (defined as the portion of inconsistently categorized events among all events) in the characterization obtained at each layer when varying one of the following components:

1. *Input data source (interpolation method):* we used two Germany-wide gridded precipitation products (i.e., Rauthe et al., 2013; Zink et al., 2017) that were produced through different interpolation methods of station observations. The test was performed at 4-km resolution due to the coarser resolution of the precipitation product provided by Zink et al. (2017). In a similar fashion, we analyzed effects on the results of the characterization due to different interpolation methods applied in the considered temperature data sets. In both cases only the interpolation method differed while the amount of data was not altered.
2. *Data resolution*: the original spatial resolution of observed and simulated data used for event characterization was upscaled by a factor of 2 (i.e., we analyzed the effect of a coarser spatial resolution).
3. *Parameter set*: we used snowmelt and soil moisture data simulated by 100 different representative parameter sets (i.e., 100 equifinal realizations of parameters in terms of streamflow efficiency for major German river catchments) of the mHM model (Kumar et al., 2013; Samaniego et al., 2010; Zink et al., 2017).
4. *Model structure and calibration technique:* we used simulated snowmelt and soil moisture of two regionally calibrated distributed conceptual hydrological models, namely, the mHM and SALTO (Merz

et al., 2020) models. The comparison was performed at 8-km resolution due to the coarser resolution of the latter model outputs. Both models were driven by the same input data.

In a second step we focused on the uncertainty of the categorization itself (i.e., the results of binary splits) evaluating the robustness of thresholds and selected indicators by means of Monte Carlo experiments similar to the procedure described by Sikorska et al. (2015):

1. *Thresholds uncertainty:* we performed a Monte Carlo experiment by substituting the former crisp thresholds (e.g., the original value of the threshold for the temporal coefficient of variation of the precipitation rate, which was equal to 1, see Table 1) with random variables with probability density functions defined as truncated normal distributions (Sikorska et al., 2015). The distribution was bounded at 0 for all thresholds since negative values are physically not possible for ratio- and coefficient of variation-based indicators. For covariance-based indicators, negative values are possible but disagree with the intended meaning of the indicators (i.e., spatial coincidence), and therefore, these thresholds were bounded at 0 as well. The experiment was run 10,000 times, and every time, a new value for each threshold was sampled randomly and used to perform the categorization based on the proposed multilayer framework. Different from Sikorska et al. (2015), we analyzed the uncertainty at every characterization layer separately, since our framework is envisaged in a way that overcompensation through other thresholds is not possible (i.e., once attributed to a group within a characterization layer, an event cannot change its attribution). In this way we were able to quantify robustness of each layer of the proposed framework.
2. *Uncertainty of indicator choice*: we chose alternative indicators for each characterization layer that potentially allows for similar stratification of events and that were previously used for causative classification of river flood events (Tarasova et al., 2019). In particular, we compared ratio-based indicators with their counterparts based on absolute values (e.g., ratio of snowmelt and rainfall is substituted by their absolute values; Table 1). Covariance-based indicators were instead compared with indicators based on the portion of overlap (e.g., instead of normalized spatial covariance of precipitation volume and antecedent soil moisture, the portion of overlapping precipitation fields and wet areas are used). In the same fashion, indicators that were based on the spatial coefficient of variation are tested (e.g., instead of the spatial coefficient of variation of precipitation volumes we used extent of the precipitation field relative to the catchment area; Table 1). For indicators of catchment-averaged wetness state, we used the antecedent precipitation and snowmelt index (Kohler & Linsley, 1951), which is often used as a substitute of soil moisture (Keller et al., 2017). To avoid possible biases in the application of these absolute indicators in a wide set of catchments, their thresholds were defined as catchment-specific percentiles (Stein et al., 2019). First, we compared the results of categorization using these alternative indicators and the original indicators. In the following step, we performed analysis of their *threshold uncertainty* using a Monte Carlo experiment as described above for the original indicators.

3.4. Validation of the Derived Event Typology: Similarity of Hydrograph Dynamics of Event Types

The validity of the hypothesized runoff generation processes (Table 2) cannot be verified at catchment scale (i.e., we cannot observe these processes at this scale). However, if a set of indicators used for event characterization is able to capture distinct runoff generation processes, similarly classified events should exhibit similar hydrological response (Sikorska et al., 2015). The proposed framework is essentially runoff-free (i.e., runoff data are used only for identifying beginning and end of events and in the regional calibration of the hydrological models which provide soil moisture and snow cover data sets). Therefore, we can validate the derived event typology (section 3.2; Figure 2b) by testing if the identified event types group events with distinct hydrological response, which is comprehensively described by event hydrographs (Blöschl et al., 2013).

To quantify and compare dynamics of event hydrographs within each study catchment, we used cross recurrence plots (CRPs) and recurrence quantification analysis (RQA; Marwan et al., 2007). CRPs aim to describe the degree of similarity/dissimilarity between (nonlinear and nonmonotonic) time series of complex environmental systems and their recurring patterns. They were successfully applied in several disciplines of environmental sciences (e.g., Aceves-Fernandez et al., 2011; Eroglu et al., 2016).

The CRP similarity quantification method is based on the time delay embedded phase space representation of time series, in our case event discharge hydrographs (Wendi et al., 2019). Essentially, every point of an

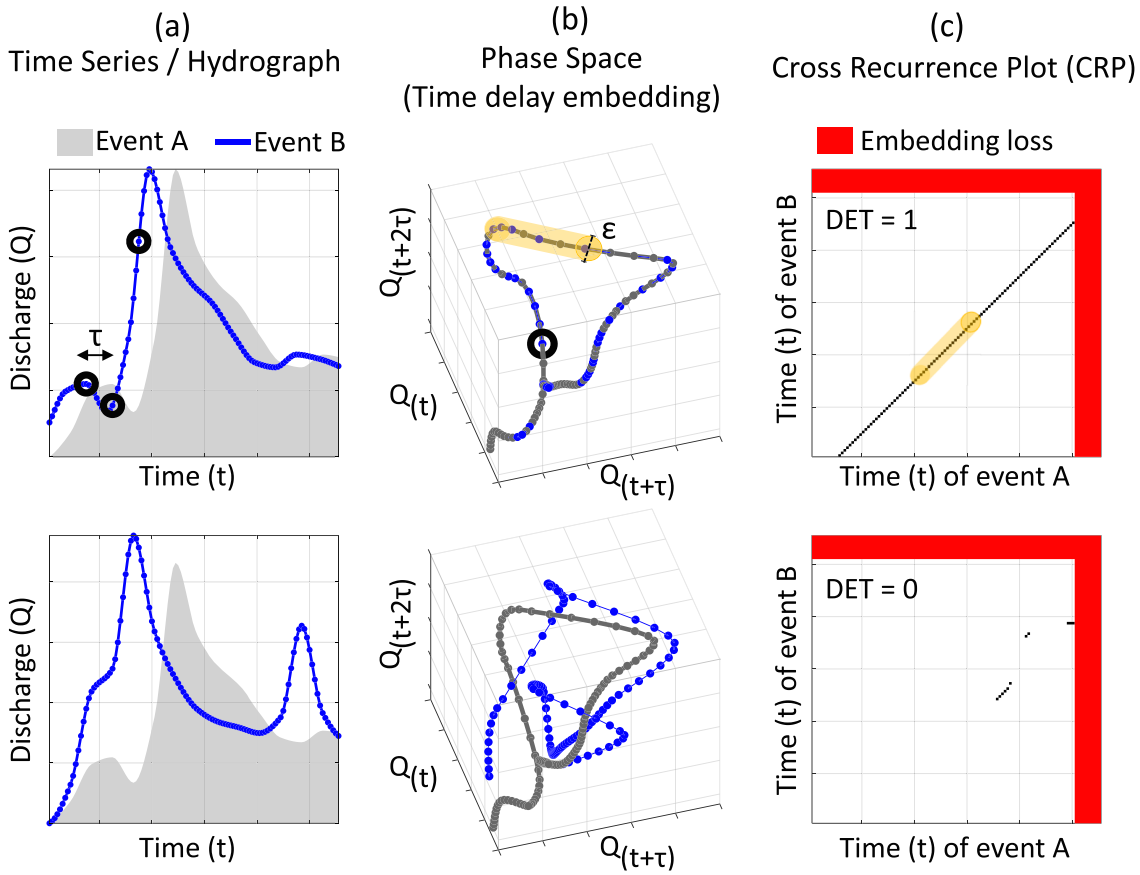

Figure 3. Cross recurrence plot (CRP) illustration for quantifying similarity of runoff dynamics between (a) event hydrographs (events A and B). Top panels illustrate how two identical hydrographs with different time of occurrence (blue line and shaded area in top panel a) are compared in their three-dimensional phase space where the axis correspond to the original hydrograph values separated by the time delay τ and 2τ (top panel b) and CRP (top panel c). The bottom panels illustrate the same comparison, but for two hydrographs with different runoff dynamics. Any identical or similar piece of the phase space trajectory which falls within the distance range/radius (ε; displayed in yellow in the top panel) is converted into a diagonal line in the CRP. These diagonal lines thus summarize the similarity of runoff dynamics between the compared hydrographs. Two perfectly identical runoff dynamics (top) have a DET of 1, while dissimilar dynamics yield low DET values (in our example $DET = 0$). Portion of CRP shaded in red indicate the information loss resulted from the time delay embedding method.

event hydrograph (i.e., streamflow at every time step $Q(t)$, Figure 3a) is represented in the m-dimensional phase space (e.g., a three-dimensional phase space on Figure 3b) by plotting the original values of event hydrograph separated by the time delay τ and connecting these points as a trajectory (Figure 3b). Time delay embedding provides an opportunity to account for the multidimensional relationships between different points in time within the events. This means that temporal evolution of discharge including effect of antecedent conditions is accounted for (Wendi et al., 2019).

Streamflow hydrographs are compared to each other based on the Euclidean distance of their phase space trajectories (Figure 3b). In phase space domain, two event hydrographs that have similar temporal dynamics of streamflow have very close trajectories (Figure 3b, top panel). Instead the trajectories of two hydrographs with disparate streamflow dynamics are far away from each other in the phase space domain (Figure 3b, bottom panel). CRP is finally used to visualize the aforementioned similar trajectories corresponding to their time of similarity occurrence (Figure 3c). These are trajectories within the user-defined distance range/radius (ε), and when the condition is met it is indicated as value 1 or otherwise 0 on a CRP. A CRP is a two-dimensional binary matrix with x and y-axes representing the time of two compared events. When two identical event hydrographs are compared (even if they are time shifted), a single diagonal line appears on the CRP and indicates the time when two trajectories are similar (Figure 3c, top panel). Instead, when two dissimilar hydrographs are compared the line on the CRP plot becomes nonexistent and/or not diagonal (Figure 3c, bottom panel).

The resulting patterns within a CRP (Figure 3c) can be quantified by RQA, which is essentially used as a similarity measure between two event hydrographs. The chosen measure, called determinism (*DET*; see Text S1, Figure 3c), intrinsically varies between 0 and 1 and defines a range from low to high (or identical) similarity in the dynamics of the two times series (Marwan, 2010). It is worth to note that CRP and RQA can effectively compare pairs of event hydrographs regardless of their possibly different duration (Wendi et al., 2019).

For detailed information on construction of CRP and performing RQA, we refer the reader to the work of Wendi et al. (2019). The parameterization of CRP used in this study is described in supplementary information Text S3.

To use DET as a similarity metric of hydrograph dynamics of event types, we calculated DET for each event pair in every catchment separately. Once events were classified, we calculated the average hydrograph similarity DET_{intra} among event pairs belonging to the same type (e.g., *Rain-on-snow*) in each individual study catchment. Similarly, we calculated the similarity DET_{inter} between event pairs belonging to different types (e.g., *Rain-on-snow* and *Rain-on-ice*). In each catchment, DET_{intra} shows how similar events of the same type are while DET_{inter} shows how similar are events classified into different types. If $DET_{intra} > DET_{inter}$, streamflow dynamics of events that are classified into the same type are more similar than those of events that are classified into different types. Fulfillment of this condition would support our assumption that the different hypothesized runoff generation processes (Figure 2b and Table 2) determine distinct streamflow dynamics.

3.5. Exemplary Application of the Derived Event Typology in Germany

The event typology derived in section 3.2 was applied to classify runoff events for the set of German catchments (see section 2). Here our goal was to analyze spatial patterns of event type occurrence, their seasonality in Germany, and to obtain typical runoff characteristics for each event type.

To identify regions with similar event type occurrence, we clustered all catchments in the data set according to their distribution of event type occurrences. A simple *k*-mean clustering approach (Hartigan & Wong, 1979) was used. The number of clusters was determined using silhouette index (Rousseeuw, 1987).

We analyzed the seasonality of the dominant event types by defining the most frequent type within the whole study period in a particular season (i.e., for summer [June–August], autumn [September–November], winter [December–February], and spring [March–May]) and catchment. Additionally, we evaluated interannual variability of event type occurrence for each season using a variability measure for categorical variables (Kader & Perry, 2007):

$$u_2 = 1 - \sum_{i=1}^{j} \left(\frac{k_i}{n}\right)^2, \tag{1}$$

where u_2 is called coefficient of unalikeability (Perry & Kader, 2005), which is a measure of categorical variability for a variable that has j number of categories (i.e., number of possible event types), n observations (i.e., total number of events in a certain season for the whole observation period), and k_i number of objects (i.e., events) within a category (i.e., event type) i. Essentially, the coefficient of unalikeability defines how often the category of observations varies and can be interpreted as a portion of possible comparisons (i.e., event pairs) which are unalike (Kader & Perry, 2007). When one event type dominates in a certain season, the value of u_2 is close to 0. When several event types are equally dominant for a certain season, u_2 increases towards the maximum value of 1.

For each classified event, we derived typical event runoff characteristics that describe event runoff response. Here we considered event runoff coefficient (dimensionless), time scale (days), rise time (dimensionless), peak discharge (mm/day), and total volume of runoff event (mm). These runoff event characteristics are fundamentally different from the indicators of the proposed framework and can be used to shed light on event runoff response (see, e.g., Tarasova, Basso, Zink, et al., 2018) among different event types. The event runoff coefficient is a volumetric ratio between event runoff and event precipitation. It characterizes which portion of precipitation is stored and which portion instead contributes to event runoff. The event time scale is a ratio between event volume and peak discharge (Gaál et al., 2012). It is related to both the shape and duration of the event hydrograph. Shorter event time scales are expected when fast runoff generation processes (i.e.,

overland flow and fast subsurface stormflow) dominate, while longer time scales indicate the dominance of slow runoff components. The event rise time characterizes the duration (in days) from the beginning of the event till the day when peak discharge is observed (Bell & Om Kar, 1969) normalized by the overall duration of the event in days. Therefore, it indicates how fast the peak discharge is reached. The peak discharge and total runoff event volume characterize the magnitude of the event. Since these characteristics depend on size, landscape, and climatic settings of each catchment (Tarasova, Basso, Poncelet, et al., 2018), we scaled them to obtain zero mean and unit variance for their samples in each catchment and thus make them comparable across settings.

4. Results and Discussion

4.1. Uncertainty Analysis and Robustness Check of the Proposed Framework

We performed an uncertainty analysis and robustness check for four test river basins (75 catchments in total) to evaluate possible regional inhomogeneity in the uncertainty of the results of the proposed characterization framework (Figure 4). We quantified the uncertainty as the portion of discrepancy (i.e., portion of events that is differently categorized) for each catchment among cases that use different input data, thresholds, or indicators.

The choice of the precipitation product (i.e., the method used to interpolate data from rainfall gauges; Figure 4, *productP*) greatly affects categorization of events in terms of the space–time organization, while their temporal organization and the nature of inducing events are only slightly affected. These results are somehow expected, as interpolation methods largely decide the spatial organization of precipitation (i.e., the data used in the space–time organization layer). However, given the large density of rain gauges in Germany, different interpolation methods are likely to provide similar results in terms of catchment-averaged precipitation rates (i.e., data used in the inducing event and temporal organization layers, Table 1). Increasing availability of radar data, which are believed to accurately detect the spatial structure of rainfall events (Rabiei & Haberlandt, 2015), might improve consistency also for what concerns the space–time organization of events. Conversely, the choice of the temperature product (i.e., interpolation model; Figure 4, *productT*) and the spatial resolution of input data (Figure 4, *resolution*) do not affect classification results in the examined catchments.

Parametric uncertainty has minor effects on event characterization (Figure 4, *parameters*). In case of wetness state characterization, the visible stability of the performance is an outcome of the choice made to define the classification threshold specifically for each model simulation. We set the threshold equal to the simulated soil moisture value that corresponds to the point of maximum curvature of the catchment-specific function describing the relation between soil moisture simulated by a specific model parameter set and observed event runoff coefficients (see section 3.1). Except for the spatial organization of pre-event catchment states, the uncertainty linked to different hydrological models is comparable with the effect of parametric uncertainty of the initially selected model (Figure 4, *parameters* and *model*).

There is also no evidence of regional differences in the framework, save for the temporal organization layer. Indicators of temporal organization in the Sieg and Naab River basins seem to be sensitive to variations of their thresholds. A similar pattern of discrepancy is observed for the original and alternative indicators, signaling that the problem probably does not lie in the indicators themselves. However, we could not identify the exact reason of this discrepancy.

Covariance- and ratio-based indicators proved to be less sensitive to the choice of specific thresholds than indicators based on coefficient of variations (Figure 4, *thresholds*). Generally, the choice of indicators plays a crucial role for the characterization results, except for the temporal organization of inducing events (Figure 4, *alternative indicators*). According to the Monte Carlo experiment, the tested alternative indicators (which are based on absolute values and use catchment-specific percentiles of these values as thresholds) are less stable than the original indicators (Figure 4, *alternative indicators' thresholds*). This means that selecting quartiles as thresholds, even if it supports regional homogeneity of classification rules for indicators based on absolute values (Stein et al., 2019), makes characterization results very sensitive to the choice of thresholds. In fact, this behavior also emerges for the original (i.e., dimensionless) indicators in the space–time organization layer (*Local* vs. *Extensive*), where a quartile of the spatial coefficient of variation was used as threshold

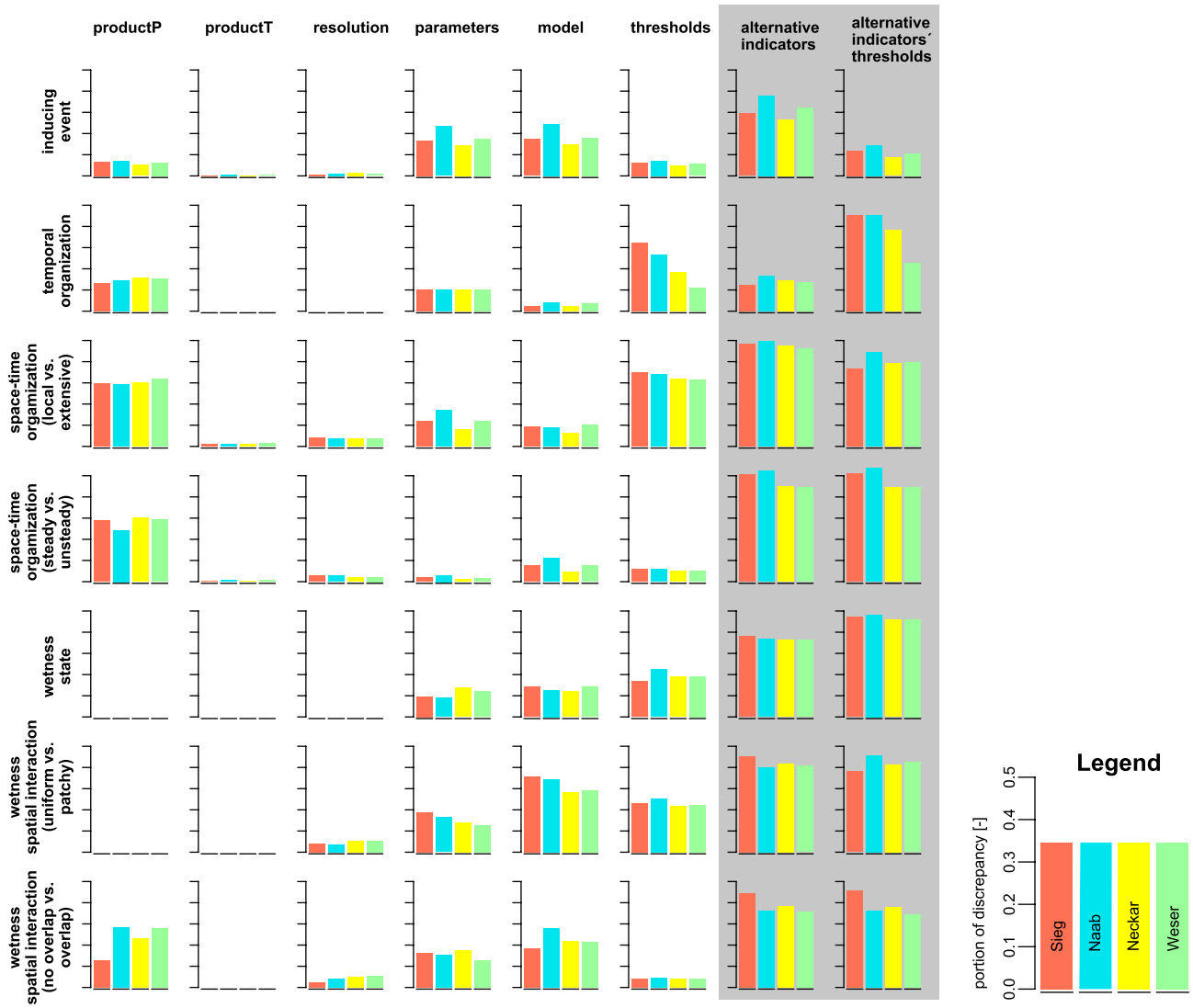

Figure 4. Quantification of input uncertainty and robustness of the characterization framework for four test river basins (75 catchments in total): the Sieg, Naab, Neckar, and Weser rivers. Categorization discrepancy (i.e., portion of all events that is differently categorized) was calculated for each layer of the framework (i.e., inducing event, temporal organization, spatial organization, space–time organization, wetness state, and spatial interaction; see rows of the figure). The portion of discrepancy may vary from 0 to 1. High discrepancy indicates that categorization results are highly sensitive to the specific choice of the tested feature (i.e., input data, thresholds, or indicators). Different features were examined (see columns of the figure): Different precipitation (*productP*) and temperature (*productT*) products, original spatial resolution of products upscaled by a factor of 2 (*resolution*), set of 100 different soil moisture and snow water equivalent realizations simulated by 100 equifinal parameter sets of the mHM model (*parameters*), simulations of soil moisture, and snow water equivalent from the alternative SALTO model instead of the original mHM run (*model*). We performed the robustness test for thresholds (*thresholds*) by means of a Monte Carlo experiment, where values of thresholds were randomly sampled from distributions assigned to the respective indicators (see Table 1). We averaged the categorization discrepancy over 10,000 Monte Carlo runs. The gray background highlights categorization discrepancy when alternative indicators listed in Table 1 (*alternative indicators*) were used instead of the original indicators. The robustness test for their thresholds (*alternative indicators thresholds*) shows average categorization discrepancy of 10,000 Monte Carlo runs.

to avoid biases due to threshold selection in catchments of different sizes. This resulted in the larger threshold sensitivity compared to other dimensionless indicators, which do not use thresholds based on quartiles (Figure 4, *thresholds*).

4.2. Validation of the Derived Event Typology: Similarity of Hydrograph Dynamics of Event Types

The results of the RQA averaged for all catchments in the data set show distinct differences in the event hydrographs generated during wet and dry conditions (Figure 5a, red pixels in the lower right corner).

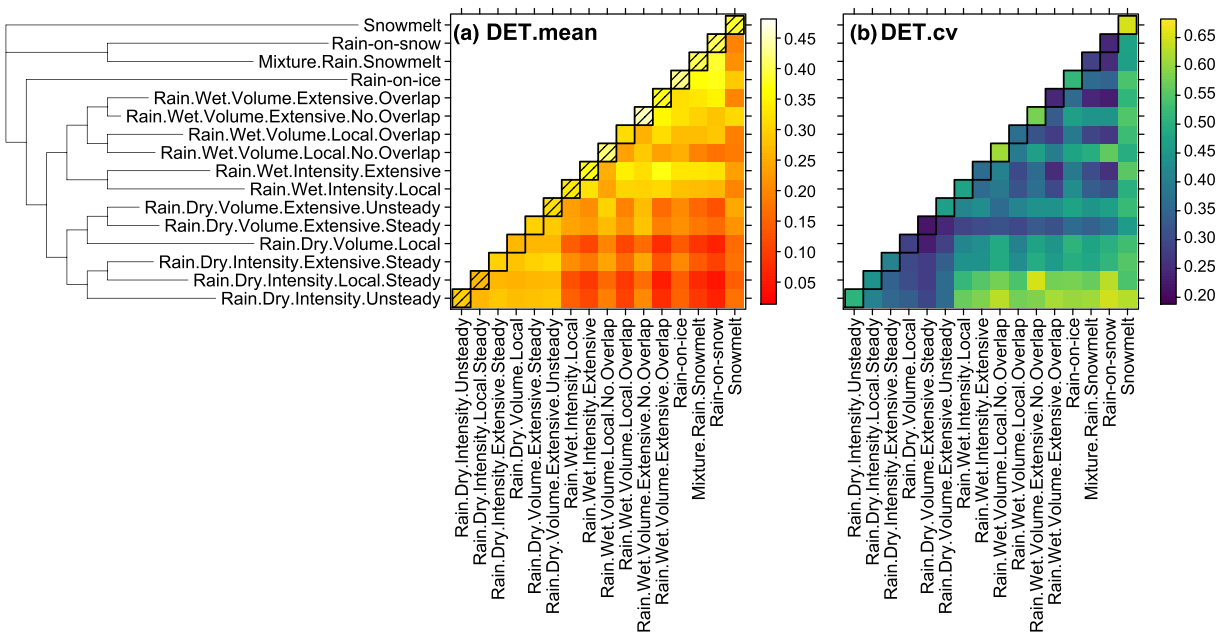

Figure 5. Matrix of intra-type and inter-type value of determinism (*DET*, see section 3.4, Text S3, and Wendi et al. [2019]), which measures the similarity of hydrograph dynamics of event pair (*DET* = 0 corresponds to a low similarity in the dynamics): (a) mean and (b) coefficient of variation of the average values of DET_{inter} and DET_{intra} for all catchments. The diagonal shows DET_{intra} (black outline), and horizontal and vertical bars show DET_{inter}. Events of a certain type exhibit more similar runoff dynamics compared to other event types if the value on the diagonal is higher than the values of the corresponding horizontal and vertical bars (these conditions are indicated by hatched squares). The matrix of coefficient of variations shows the variability of DET_{intra} (diagonal values, black outline) and DET_{inter} (values on horizontal and vertical bars) in the set of study catchments. The hierarchical classification tree corresponds to the decision tree in Figure 2b.

Events that occurred during dry conditions tend to have more diverse intra-type dynamics than events occurred during wet conditions (i.e., wet events have lower DET_{intra} values on the diagonal of the matrix than the dry ones, Figure 5a). This can be explained by the fact that during wet cases a larger portion of catchment is contributing to runoff generation than during dry cases, thus resulting in more homogeneous runoff generation conditions.

Generally, *Local* rainfall events show much more variability in runoff dynamics (i.e., on average lower values of DET, Figure 5a) compared to their *Extensive* counterparts, indicating that the exact location of rainfall in the catchment for *Local* events might be the dominant control of a particular event runoff dynamic in both dry and wet cases.

Inter- and intra-type similarities are comparable for events that are generated during dry conditions, indicating that events generated during dry conditions have somewhat similar dynamics which are not affected by the space–time properties of rainfall events. This phenomenon is consistently observed in most of the catchments (i.e., low coefficient of variation of DET values in the lower left corner, Figure 5b). On the contrary, a clear difference among intensity- and volume-dominated events can be observed for events generated during wet conditions (i.e., large differences between DET_{intra} and DET_{inter} values for *Rain.Wet.Intensity* and *Rain.Wet.Volume* than compared to their dry counterparts, Figure 5a).

Interestingly, very distinct characteristics also distinguish runoff events that are generated when there is high or low spatial interaction between rainfall and wet areas (i.e., most of rainfall occurs in the parts of catchment that is respectively wet or dry) as the DET_{inter} between *Overlap* and *No Overlap* events is rather low (i.e., dark orange colors, Figure 5a), while the corresponding DET_{intra} of these types are very high (i.e., light yellow colors on the diagonals, Figure 5a). *Rain-on-ice* events have a very specific dynamics that is to some extent similar to *Rain-on-snow* and *Mixture of Rainfall and Snowmelt* events. Pure *Snowmelt* events have even more specific hydrograph dynamics that does not match any other event type, as evidenced by the presence of light yellow colors (i.e., high DET) only in the diagonal of Figure 5a. However, in some catchments, the variability of runoff dynamics within *Snowmelt* event type can be higher than on average

for the whole data set (Figure 5b). Interestingly, there is more resemblance between *Rain.Wet.Volume.Extensive.Overlap* and events that involve snowmelt (i.e., *Rain-on-snow* and *Mixture of Rain and Snowmelt*) compared to other rainfall-induced events (i.e., yellow colors of DET_{inter} between these types, Figure 5a), indicating that the runoff generation conditions in these cases (despite the differences in the nature of the inducing event itself) are controlled by similar space–time dynamics of inducing event and wetness state of the catchment. This appears to be consistent for most of the analyzed catchments in the data set (i.e., low values of coefficient of variation of DET_{inter} for these three event types, Figure 5b).

5. Exemplary Application of Event Typology in Germany

In this section we applied the event typology (Figure 2b and Table 2) derived for the set of German catchments for investigating the spatial distribution of event type occurrence, the seasonal variations of the dominant event types in Germany, and their runoff characteristics. These are exemplary applications, which are reported and discussed here to give a taste of the possible uses of such a typology.

5.1. Spatial Patterns of Event Types

To analyze spatial patterns of event types in Germany, all catchments in the data set were clustered according to the frequency of occurrence of their event types. We identified six clusters (Figure 6a) with distinct distributions of event types (Figure 6b). The resulting spatial clustering of catchments resembles the one obtained by using event runoff characteristics (i.e., mean, variability, and seasonality of event runoff coefficient, time scale, and rise time) that is in turn comparable to the spatial variability of catchment and climatic descriptors (see Tarasova, Basso, Poncelet, et al., 2018, Figure 5).

The first cluster mostly consists of mountainous catchments of the Alpine Foreland, Black Forest, and Rhenish Massif (Figure 6a). Here almost 25% of events are either *Rain-on-snow* or *Mixture of Rainfall and Snowmelt* (Figure 6b). A big portion of rainfall events occurs during wet conditions and only less than 25% of events occurs during dry conditions. This can be explained by the fact that this cluster largely corresponds to the group of catchments with small subsurface storage and in-phase seasonality of soil moisture and rainfall (i.e., substantial rainfall amount falls during the cold season when evaporation is low and soils are wet) (Tarasova, Basso, Poncelet, et al., 2018). This explains the high portion of rainfall events on wet conditions (Figure 6, green and blue colors). Around 5% of all events are volume-dominated rainfall events in wet conditions, when most of rainfall occurs on less saturated soil (i.e., events categorized as *No Overlap*). This indicates that the corresponding runoff events might be the result of event-fed saturation in these catchments.

Cluster 2 combines wet lowland catchments of the western North German Plain, mid-range catchments of the western Central Uplands, and the South German Scarplands. Here the role of events involving snow, especially *Rain-on-snow* events, is lower than in the previous cluster as the elevations are limited. More than 40% of rainfall event occurs during wet conditions due to in-phase seasonality of rainfall and soil moisture (Tarasova, Basso, Poncelet, et al., 2018). Volume-dominated rainfall events clearly prevail during both dry and wet conditions. Event-fed saturation during wet conditions when large portions of catchments are active and contribute to runoff generation is probably a rare situation in these catchments as only a small portion of events is categorized as *No Overlap* events. However, during dry conditions when the distribution of active runoff generation zones is very heterogeneous, local event-fed saturation is the most frequent runoff generation mechanism (i.e., there is abundance of *Rain.Dry.Volume.Local* and *Rain.Dry.Volume.Extensive.Steady* event types).

Cluster 3 encompasses mountainous catchments of the eastern Central Uplands (i.e., the Harz, Ore Mountains and Thuringian Forest) and the Alpine Forelands. Similarly, to Cluster 1, almost 25% of events are associated with *Rain-on-snow* and *Mixtures of Rainfall and Snowmelt*. However, rainfall events are distributed almost equally between dry and wet conditions indicating large variability in the extent of area that actively contributes to streamflow generation in these catchments compared to Cluster 1. Almost 25% of all events are generated by means of localized event-fed saturation during dry conditions when the connectivity among active runoff generation areas is poor (*Rain.Dry.Volume.Local* and *Rain.Dry.Volume.Extensive.Steady* events). During wet conditions, both extensive and local runoff generations are almost exclusively controlled by the spatial interaction between soil moisture state and rainfall amounts (i.e.,

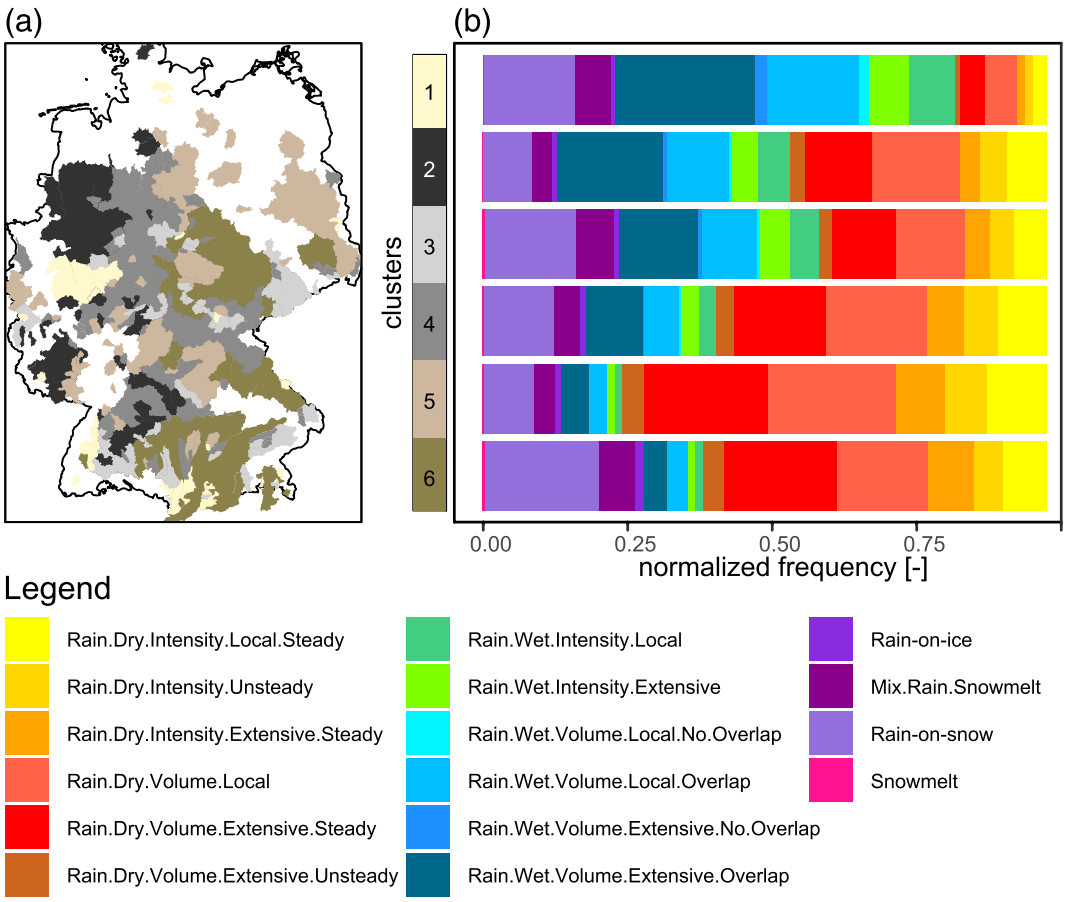

Figure 6. Regional pattern of event type frequency in Germany: (a) spatial distribution of six homogeneous clusters in terms of frequency of event types and (b) frequency of event types per cluster for all events.

Rain.Wet.Volume.Local.Overlap and *Rain.Wet.Volume.Extensive.Overlap* prevail among events occurring in a wet catchment state).

The fourth cluster combines the catchments of the central parts of the North German Plain, the hilly catchments of the Central Uplands and South German Scarplands and the lower Alpine Foreland. A larger portion of rainfall events occurs during dry conditions indicating that generally only small portions of catchments contribute to streamflow generation. As in all other regions in Germany, during wet conditions (that are usually in the cold period of the year) volume-dominated events prevail. However, in this cluster both local and extensive volume-dominated rainfall events are only generated when there is strong interaction with soil moisture (i.e., pre-event saturation is necessary for runoff generation).

The fifth cluster consists of the dry lowland catchments in the east of the North German Plain and the South German Scarplands. Here only a small portion of events belongs to *Rain-on-snow* or *Mixture of Rainfall and Snowmelt* types. Only few rainfall events occur during wet conditions, as seasonality of soil moisture and rainfall is not synchronized. Around 25% of all events are intensity-dominated rainfall events with prevailing local runoff generation and possible infiltration excess. Most of volume-dominated rainfall events during dry conditions have local character (i.e., large portion of *Rain.Dry.Volume.Local*) possibly indicating very spatially heterogeneous runoff generation patterns that largely depend on the intensity of inter-event evaporation controlling pre-event soil moisture state.

The sixth cluster covers a large portion of the Alpine Forelands, as well as Loess Belt catchments that drain mid-range Mountains of the Central Uplands (e.g., the Harz, Thuringian Forest). For this cluster, *Rain-on-snow* and *Mixtures of Rainfall and Snowmelt* events play an important role as their portion is the largest compared to other clusters. The portion of rainfall events during wet conditions is rather small

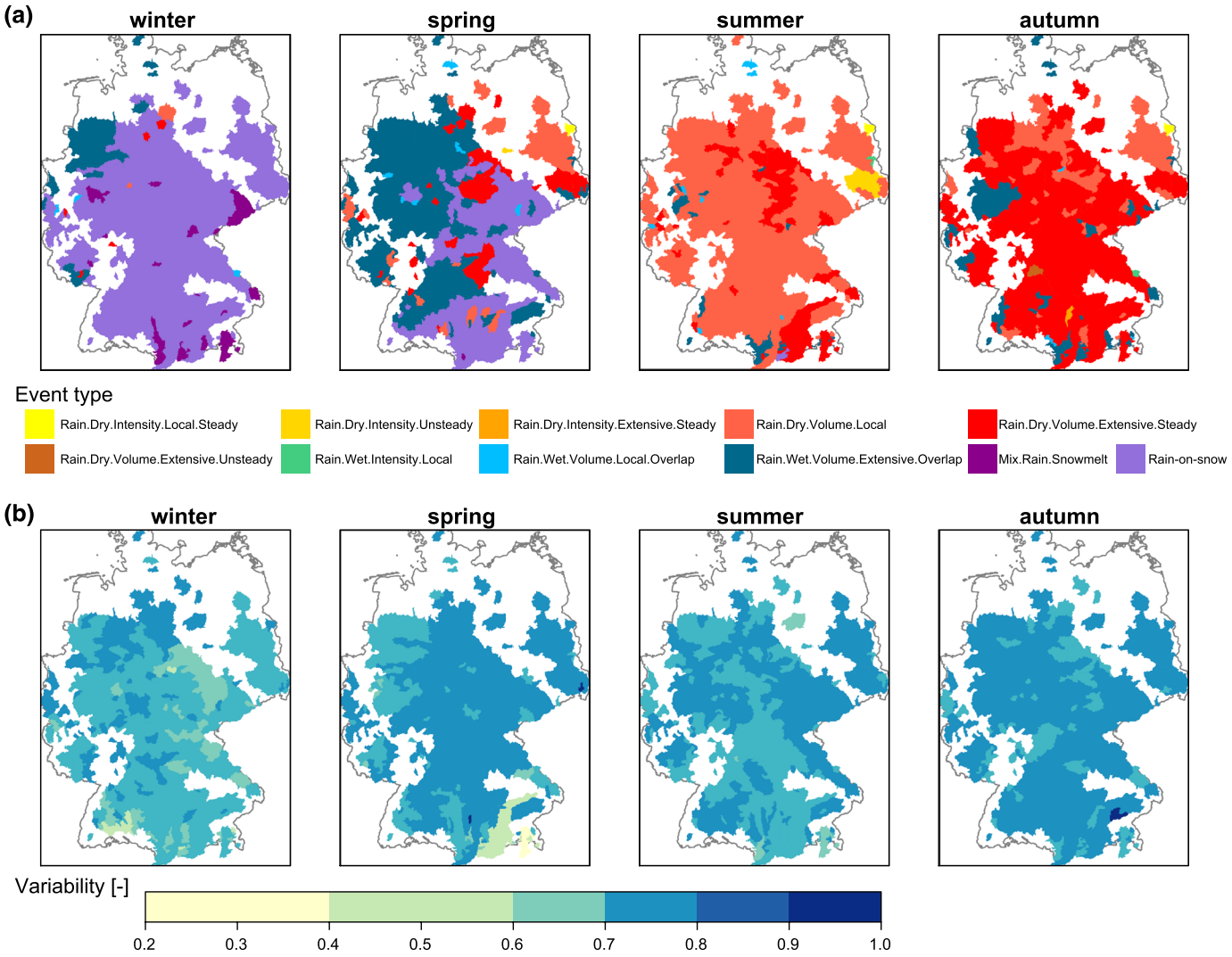

Figure 7. Seasonality of event types in Germany: (a) dominant type (i.e., most frequent type in the season) and (b) variability of event type occurrence for each season defined as coefficient of unalikeability (see section 3.5 and Kader and Perry [2007]). This metric defines the portion of all event pairs in the season that are unalike. When one event type dominates in a certain season, the value of the coefficient of unalikeability is close to 0. When several event types are equally dominant for a certain season, the value of the coefficient increases towards the maximum value of 1.

because rainfall and soil moisture are not seasonally synchronized. Steady rainfall events during dry conditions are very frequent, potentially indicating the occurrence of orographic slow-moving storms, which are mainly volume-dominated but can occasionally have an intensity-dominated structure. Almost 20% of all rainfall events are intensity-dominated indicating that infiltration excess is a possible runoff generation mechanism in these catchments.

The above-mentioned evidences of similarity among regions with homogeneous event type frequency and archetypical catchment behavior in terms of event runoff response (as reported in Tarasova, Basso, Poncelet, et al., 2018) indicate that on one hand the similarity of event runoff response might be controlled by the occurrence of similar event types, and on the other hand that catchments with similar physiographic and climatic conditions are likely to have similar distribution of event type frequency.

5.2. Seasonality of Dominant Event Types

Event type occurrence naturally varies throughout the year (Figure 7). In large parts of Germany, *Rain-on-snow* and *Mixture of Rainfall and Snowmelt* are the two most typical event types in winter

(Figure 7a). However, in the western part of the North German Plain, where snowfalls seldom occur, extensive steady rainfall on wet conditions is the most common winter event type. In spring, the situation changes, and extensive steady rainfalls on wet conditions dominate the whole Western part of Germany, indicating the dominance of Western cyclones and resulting synoptic precipitation (Hofstätter et al., 2016). Similar rainfall events also prevail in the Eastern part of Germany, but the catchment wetness state is generally lower (Figure 7a). The Southern part of Germany (i.e., the Alpine Forelands and Bavarian Forest) and the eastern part of the Central Uplands are still dominated by *Rain-on-snow* events in spring (Figure 7a), as snowpack is still present in areas with higher elevation. In summer, the character of precipitation changes, leading to dominance of local volume-dominated rainfall events on dry conditions. The frequent occurrence of intensity-dominated events in the Eastern part of Germany indicates the importance of convective precipitation for runoff event generation in warm periods of the year. In autumn, increased precipitation in mountainous catchments of the Alpine Foreland, Black Forest, and Rhenish Massif leads to abundance of rainfall events on wet conditions. Lower catchments still remain dry, but rainfall events mostly have extensive coverage. The transformation from dry to wet state in most of the catchments and the change from local to extensive rainfall events in autumn lead to increased variability of event type occurrence (Figure 7b). Similarly, the variability is high in spring when event types change from events which involve snow to rainfall-induced events. The variability of event type occurrence is lower in summer when rainfall events on dry conditions dominate. Except for the western part of the North German Plain, where air temperature is fluctuating around critical values for snowfall, the lowest variability is observed in winter when the bulk of events results from processes which involve snow (Figure 7b).

5.3. Runoff Characteristics of Event Types

Here we analyzed five runoff event characteristics (i.e., event runoff coefficient, time scale, rise time, peak discharge, and total event volume; see section 3.5) of the derived event types. These characteristics describe short-term runoff dynamics, which might shed light on the nature of precipitation partitioning and storage in the catchments and indicate principal mechanisms of water release during events (Tarasova, Basso, Zink, et al., 2018). In addition to the set of all events, subsamples of larger (i.e., runoff events with peak discharge higher than the third quartile) and smaller (i.e., runoff events with peak discharge smaller than the third quartile) events are also considered in this analysis with the purpose to investigate possible differences among runoff generation mechanisms of small and large events.

Generally, the majority of all runoff events in Germany and especially of small events results from rainfall events occurring during dry soil moisture conditions (Figures 8a and 8c; warm colors). Rainfall during wet conditions (indicated by green and blue palette), and especially *Rain-on-snow* and *Mixtures of Rainfall and Snowmelt* (violet colors), are much more common among larger events (Figure 8b). Only a small portion of events was classified as *Rain-on-ice* and pure *Snowmelt* event types (Figure 8a), indicating that these are rare runoff generation processes in Germany. Nonetheless, a portion of larger events are generated by these processes (Figure 8b) indicating that rare processes might still be responsible for generation of large flood events (e.g., Stein et al., 2019).

A clear increase of event runoff coefficients is observed from dry rainfall events to wet rainfall events (Figure 8) indicating that a substantial portion of rainfall is stored during dry conditions. The highest runoff coefficients are observed for *Rain-on-ice* and events which involve snow (i.e., *Mixtures of Rainfall and Snowmelt*, *Rain-on-snow* and *Snowmelt*). These event types are often characterized by very wet conditions (Table 2). Moreover, frozen soils and snow cover might also hinder infiltration of rain water, which then immediately contributes to runoff. When events of all sizes are considered, a very similar pattern is revealed for peak discharges and runoff event volumes (Figure 8a). Interestingly, when only larger events are considered, smaller runoff volumes are typical for intensity-dominated rainfall events, but there are little differences between event runoff volumes of extensive rainfall events during wet conditions (dark blue colors) and all events which involve snow. This can be explained by the fact that these events are indeed similar in terms of space–time organization of the precipitation events and wetness state of the catchments (Figure 5a).

Differences in peak discharges among event types for larger events almost vanish compared to the case when all events are considered. However, volume, intensity, and duration of the corresponding precipitation events (Figure 9b) indicate distinct differences among event types even for large events.

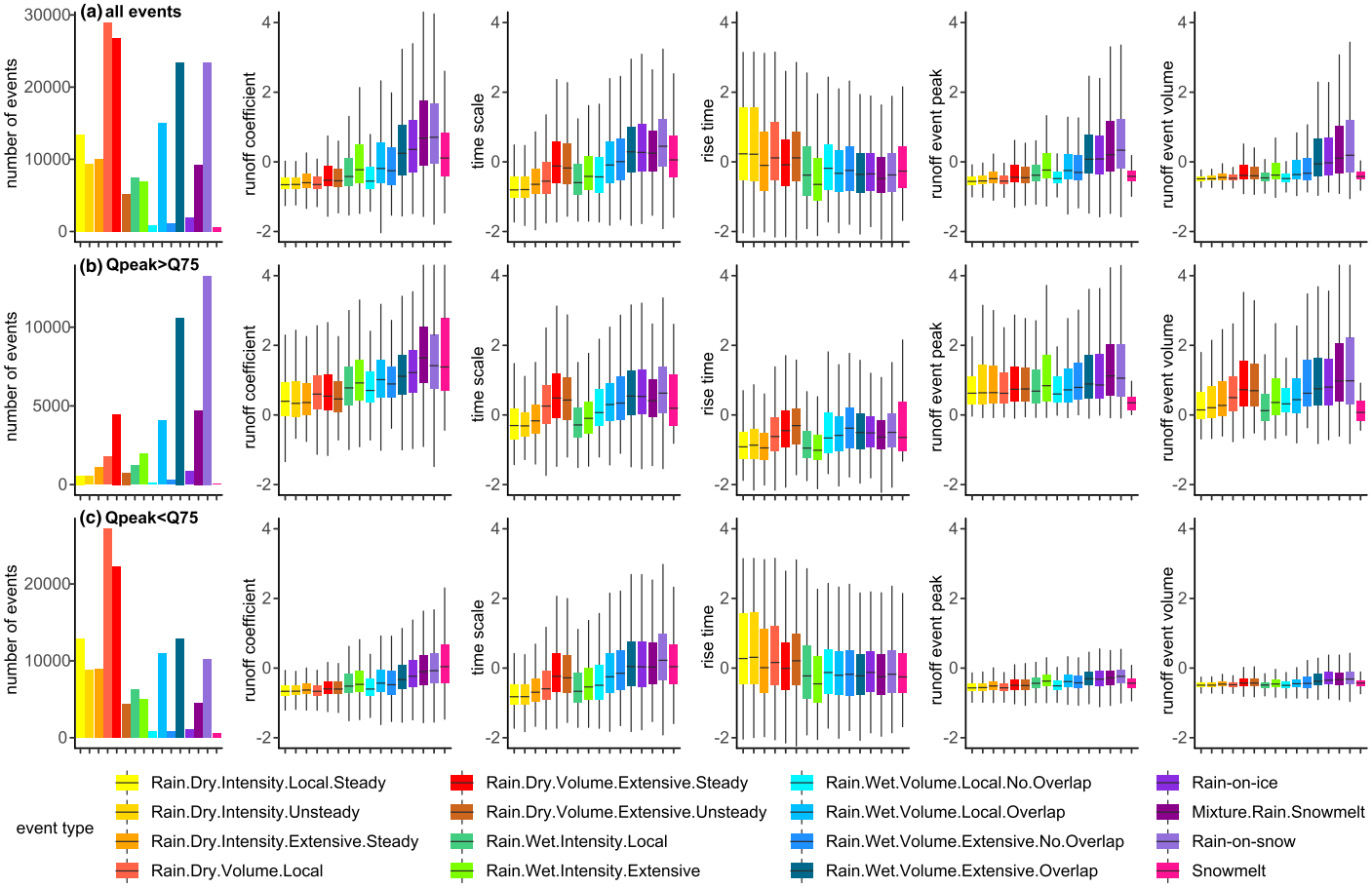

Figure 8. Hydrological response of event types in terms of their runoff event characteristics (i.e., event runoff coefficient, time scale, rise time, peak discharge, and volume): (a) all events; (b) events with $Q_{peak} > Q75$; and (c) events with $Q_{peak} < Q75$. Runoff event characteristics are standardized by scaling all events to zero mean (by subtracting the mean of the sample from the original values) and unit variance (by dividing them by the standard deviation) for each catchment to make them comparable. Legend shows color coding of event types that is also valid for Figure 9.

Intensity-dominated rainfall events on dry conditions are characterized by higher precipitation intensities than their counterparts with wet conditions. Similarly, volume-dominated rainfall events on dry conditions have higher precipitation event volume than their counterparts with wet conditions (Figure 9b), but show little difference in the event runoff volumes (Figure 8b), likely because the storage available in dry conditions is of the order of the difference between rainfall amounts of the two states.

The event time scale of intensity-dominated rainfall events (yellow and green palette) is smaller than the one of their volume-dominated counterparts (red and blue palette; Figure 8). These differences are especially distinct for larger events (Figure 8b). Event time scale of *Rain-on-snow* and *Mixtures of Rainfall and Snowmelt* is usually longer than for rainfall events (Figures 8a and 8c), but this tendency is less pronounced for larger events (Figure 8b).

When events of different magnitudes are considered together (Figure 8a), no clear differences in event rise time are detectable (Figure 8a). However, for larger magnitudes, intensity-dominated rainfall events seem to have shorter rise time than volume-dominated ones (Figure 8b), showing that when runoff is generated by intensity-dominated events fast flow paths can be activated.

Distinct runoff characteristics of event types are valuable features for constructing event type-based design hydrographs (e.g., Brunner et al., 2017). The above-mentioned disparate runoff and precipitation characteristics of event types provide useful information on how runoff characteristics might change in the future given the expected change in precipitation properties (e.g., if intense rainfall events become more frequent in the future (Kendon et al., 2014).

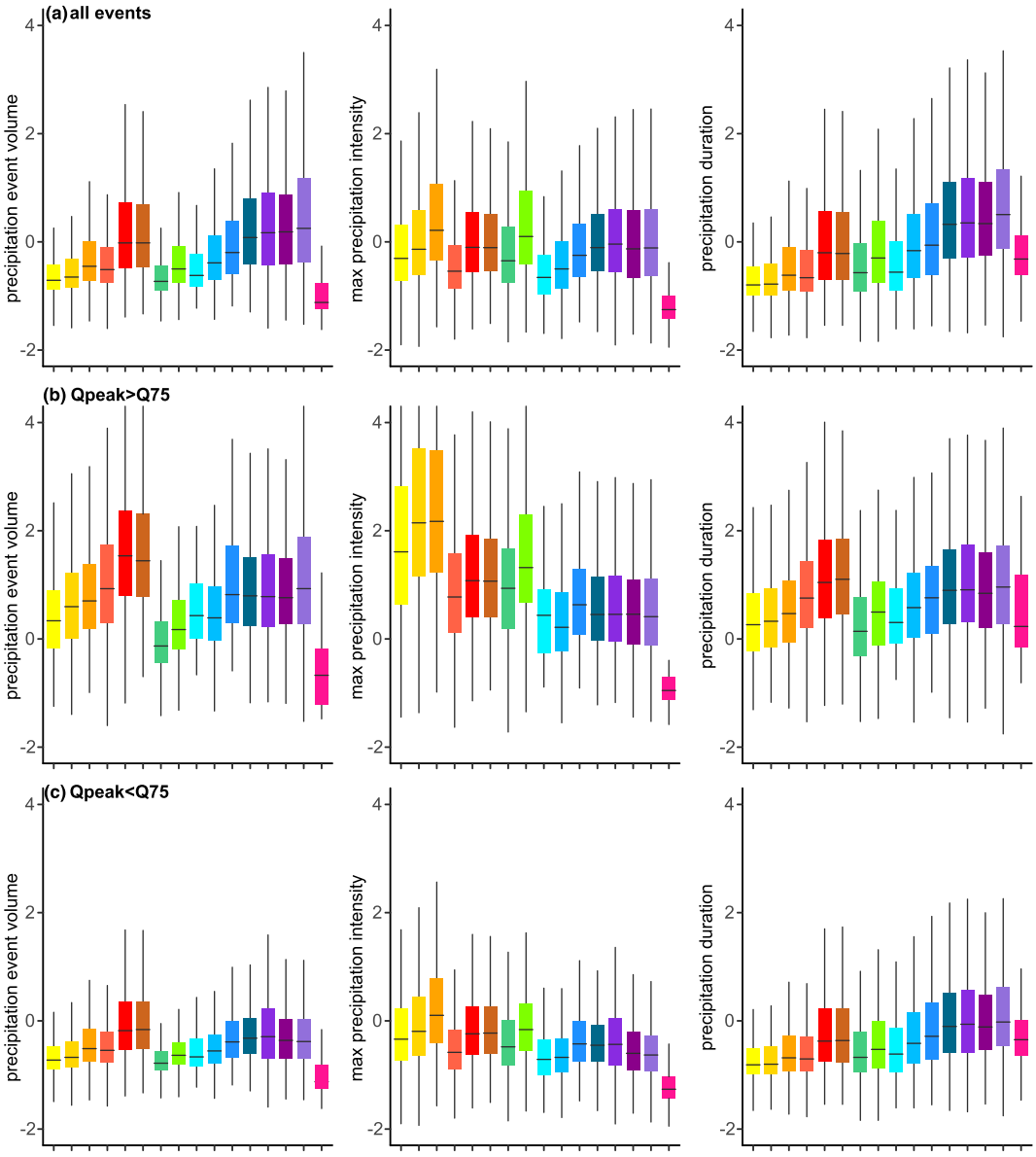

Figure 9. Precipitation properties of event types that were not used for event characterization (i.e., maximum precipitation intensity, total volume of precipitation event, and duration of precipitation event): (a) all events; (b) events with $Q_{peak} > Q75$; and (c) events with $Q_{peak} < Q75$. Maximum precipitation intensity, total volume of rainfall, and duration of precipitation events are standardized by scaling all events to zero mean (by subtracting the mean of the sample from the original values) and unit variance (by dividing them by the standard deviation) for each catchment to make them comparable. The color-coding of this figure corresponds to the color key in Figure 8.

6. Conclusions and Outlook

A new process-based framework for characterizing runoff events and deriving event typologies was proposed. The indicators used in the framework categorize runoff events based on space–time dynamics of observed precipitation and simulated snowmelt and soil moisture and their mutual interactions within river catchments. A rigorous uncertainty analysis shows that the indicators of the proposed characterization framework are robust and regionally consistent. The adoption of dimensionless covariance- and ratio-based indicators reduces classification uncertainty compared to commonly used indicators relying on absolute values of metrics such as rainfall volume, duration, or intensity. This indicates that the proposed characterization framework is a reliable tool that could be used for regional studies on runoff generation processes and for investigating their temporal evolution in the past as well as in the future.

The analysis of CRPs showed that the event typology derived in this study proves able to stratify events with distinct hydrograph dynamics, even though streamflow was not directly used for classification. This indicates that the derived event typology effectively captures first-order controls of event runoff response in a wide variety of catchments. Application of the derived event typology to a country-wide data set from Germany revealed six distinct regions characterized by different dominant event types, which match well with the corresponding regions exhibiting similar runoff response. Unveiled seasonal patterns of event type occurrence show high spatial coherence that agrees with the seasonality of hydroclimatic conditions prevailing across Germany. These are valuable findings for regionalization of hydrological signatures and prediction of the seasonal character of the runoff response in ungauged locations.

The proposed framework, which allows for a consistent characterization and comparison of runoff events of various sizes and recurrence intervals in a wide range of German catchments, is expected to perform similarly well in regions where hydrometeorological data of comparable quality in terms of density of observations, spatial and temporal resolution of the input data is available.

Instead, application of the framework to regions with extremely pronounced seasonality (e.g., snow- or monsoon-dominated river regimes) might not be trivial, since the very definition of runoff event in these areas differs from the one adopted for temperate climates (e.g., a runoff event can last several months making computation of temporal and spatial organization of the corresponding inducing event and wetness state ambiguous using the proposed indicators).

Thanks to the adaptive structure of the proposed characterization framework, the complexity of derived event typologies can be adapted to specific applications or adjusted in case of known problems in the reliability or availability of input data. The characterization framework and the derived event typology can be further applied for understanding regional differences of runoff and flood events and detect temporal changes of the dominant flood generation processes. This might help explain previously detected disparate trends in unstratified floods. The same approach could be also applied for disentangling the variability of solutes and particulates exports from catchments during different runoff event conditions (as recently suggested by Karwan et al. (2018) and Knapp et al. (2020)), thus providing additional insights on variability of water quality metrics observed in streams.

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)", the Austrian Science Fund ("Fonds zur Förderung der wissenschaftlichen Forschung", FWF) in terms of subproject I 3174 and the Helmholtz Centre for Environmental Research (UFZ) is gratefully acknowledged. We thank Charles Luce, the Associate Editor, and three anonymous reviewers for their valuable comments that helped to improve the original manuscript. For providing the discharge data for Germany, we are grateful to Bavarian State Office of Environment (LfU), Baden-Wurttemberg Office of Environment, Measurements and Environmental Protection (LUBW), Brandenburg Office of Environment, Health and Consumer Protection (LUGV), Saxony State Office of Environment, Agriculture and Geology (SMUL), Saxony-Anhalt Office of Flood Protection and Water Management (LHW), Thüringen State Office of Environment and Geology (TLUG), Hessian Agency for the Environment and Geology (HLUG), Rhineland Palatinate Office of Environment, Water Management and the Factory Inspectorate (LUWG), Saarland Ministry for Environment and Consumer Protection (MUV), Office for Nature, Environment and Consumer Protection North Rhine-Westphalia (LANUV NRW), Lower Saxony Office for Water Management, Coast Protection and Nature Protection (NLWKN), Water and Shipping Management of the Fed. Rep. (WSV), the European Water Archive (EWA), and the Global Runoff Data Centre (GRDC) prepared by the Federal Institute for Hydrology (BfG, http://www.bafg.de/GRDC). The simulations of the mHM model are available at http://www.ufz.de/index.php?en=41160. Climatic data can be obtained from the German Weather Service (DWD; https://opendata.dwd.de/climate_environment/CDC/). CRP toolbox for MATLAB is available at http://tocsy.pik-potsdam.de/CRPtoolbox/. A Data Set S1 of classified events can be found at Zenodo repository (https://zenodo.org/record/3575024). The manuscript and Supporting Information provide all the information needed to replicate the results.

References

Aceves-Fernandez, M. A., Ramos-Arreguín, J. M., Pedraza-Ortega, J. C., & Tovar-Arriaga, S. (2011). Finding trends of airborne harmful pollutants by using recurrence quantification analysis. *American Journal of Environmental Engineering*, *1*(1), 10–14. https://doi.org/10.5923/j.ajee.20110101.02

Basso, S., Frascati, A., Marani, M., Schirmer, M., & Botter, G. (2015). Climatic and landscape controls on effective discharge. *Geophysical Research Letters*, *42*, 8441–8447. https://doi.org/10.1002/2015GL066014

Bell, F. C., & Om Kar, S. (1969). Characteristic response times in design flood estimation. *Journal of Hydrology*, *8*, 173–196. https://doi.org/10.1016/0022-1694(69)90120-6

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. (2019). The relative importance of different flood—Generating mechanisms across Europe water resources research. *Water Resources Research*, *55*, 4582–4593. https://doi.org/10.1029/2019WR024841

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophysical Research Letters*, *43*, 4382–4390. https://doi.org/10.1002/2016GL068070

Bergström, S. (1995). The HBV model. In V. Singh (Ed.), *Computer models of watershed hydrology*, (pp. 443–476). Highlands Ranch, Colorado: Water Resour. Publ.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., 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

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., & Savenije, H. (Eds) (2013). *Runoff prediction in ungauged basins—Synthesis across processes, places and scales*. Cambridge, England: Cambridge University Press.

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A. C., & Seibert, J. (2017). Flood type specific construction of synthetic design hydrographs. *Water Resources Research*, *53*, 1390–1406. https://doi.org/10.1002/2016WR019535

Diederen, D., Liu, Y., Gouldby, B., Diermanse, F., & Vorogushyn, S. (2019). Stochastic generation of spatially coherent river discharge peaks for continental event-based flood risk assessment. *Natural Hazards and Earth System Sciences*, *19*, 1041–1053.

Diezig, R. & Weingartner, R. (2007). Hochwasserprozesstypen—Schlüssel zur Hochwasserabschätzung, Wasser und Abfall.

Donnelly, C., Greuell, W., Andersson, J., Gerten, D., Pisacane, G., Roudier, P., & Ludwig, F. (2017). Impacts of climate change on European hydrology at 1.5, 2 and 3 degrees mean global warming above preindustrial level. *Climatic Change*, *143*(1-2), 13–26. https://doi.org/10.1007/s10584-017-1971-7

Dunne, T. (1978). Field studies of hillslope flow processes. In M. J. Kirkby (Ed.), *Hillslope hydrology*, (pp. 227–293). Chichester, UK: John Wiley&Sons.

Eroglu, D., McRobie, F. H., Ozken, I., Stemler, T., Wyrwoll, K.-H., Breitenbach, S. F. M., et al. (2016). See–saw relationship of the Holocene East Asian–Australian summer monsoon. *Nature Communications*, *7*(12929), 1–7. https://doi.org/10.1038/ncomms12929

Flerchinger, G. N., Lehrsch, G. A., & McCool, D. K. (2005). Freezing and thawing/processes. In *Encyclopedia of soils in the environment*, (pp. 104–110). Oxford, UK: Elsevier.

Freyberg, J., Radny, D., Gall, H. E., & Schirmer, M. (2014). Implications of hydrologic connectivity between hillslopes and riparian zones on streamflow composition. *Journal of Contaminant Hydrology*, *69*, 62–74. https://doi.org/10.1016/j.jconhyd.2014.07.005

Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., & Blöschl, G. (2012). Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resources Research*, *48*, W04511. https://doi.org/10.1029/2011WR011509

Gosling, S. N., Zaherpour, J., Mount, N. J., Hattermann, F. F., Dankers, R., Arheimer, B., et al. (2017). A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1° C, 2° C and 3° C. *Climatic Change*, *141*(3), 577–595. https://doi.org/10.1007/s10584-016-1773-3

Hartigan, J. A., & Wong, M. A. (1979). A K-means clustering algorithm. *Journal of the Royal Statistical Society*, *28*(1), 100–108. https://doi.org/10.2307/2346830

Hirschboeck, K. K. (1987). 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, U.S.A.*, pp. 199–212.

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*, 613–615. https://doi.org/10.1002/2016WR019146

Horton, R. E. (1933). The role of infiltration in the hydrologic cycle. *Transactions of the American Geophysical Union*, *14*, 446–460. https://doi.org/10.1029/TR014i001p00446

House, P. K., & Hirschboeck, K. K. (1997). Hydroclimatological and paleohydrological context of extreme winter flooding in Arizona, 1993. In *Storm-induced geological hazards: Case histories from the 1992-1993 Winter Storm in Sothern California and Arizona*, *Geological Society of America Reviews in Engineering Geology*, (Vol. 11, pp. 1–24). Boulder, Colorado.

Institute of Hydrology (1980). *Low flow studies (Rep.1)*. Wallingford, UK: Institute of Hydrology.

Jencso, K. G., Mcglynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., & Marshall, L. A. (2009). Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale. *Water Resources Research*, *45*, W04428. https://doi.org/10.1029/2008WR007225

Kader, G. D., & Perry, M. (2007). Variability for categorical variables. *Journal of Statistics Education*, *15*. https://doi.org/10.1080/10691898.2007.11889465

Kampf, S. K., & Lefsky, M. A. (2016). Transition of dominant peak flow source from snowmelt to rainfall along the Colorado Front Range: Historical patterns, trends, and lessons from the 2013 Colorado Front Range floods. *Water Resources Research*, *52*, 407–422. https://doi.org/10.1002/2015WR017784

Karwan, D. L., Pizzuto, J. E., Aalto, R., Marquard, J., Harpold, A., Skalak, K., et al. (2018). Direct channel precipitation and storm characteristics influence short-term fallout radionuclide assessment of sediment source. *Water Resources Research*, *54*, 4579–4594. https://doi.org/10.1029/2017WR021684

Keller, L., Rössler, O., Martius, O., & Weingartner, R. (2017). Delineation of flood generating processes and their hydrological response. *Hydrological Processes*, 228–240. https://doi.org/10.1002/hyp.11407

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*. https://doi.org/10.1038/NCLIMATE2258

Knapp, J. L. A., von Freyberg, J., Studer, B., Kiewiet, L., & Kirchner, J. W. (2020). Concentration-discharge relationships vary among hydrological events, reflecting differences in event characteristics. *Hydrology and Earth System Sciences Discussions*. https://doi.org/10.5194/hess-2019-684

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resources Research*, *54*, 5088–5109. https://doi.org/10.1029/2018WR022913

Kohler, M. A., & Linsley, R. K. (1951). Predicting the runoff from storm rainfall, U.S. Weather Bareau Res. Paper 34.

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

Marwan, N. (2010). How to avoid potential pitfalls in recurrence plot based data analysis. *International Journal of Bifurcation and Chaos*, *21*(4), 1003–1017. https://doi.org/10.1142/S0218127411029008

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5–6), 237–329. https://doi.org/10.1016/j.physrep.2006.11.001

McGlynn, B. L., & McDonnell, J. J. (2003). Quantifying the relative contributions of riparian and hillslope zones to catchment runoff. *Water Resources Research*, *39*(11), 1310. https://doi.org/10.1029/2003WR002091

Mei, Y., & Anagnostou, E. N. (2015). A hydrograph separation method based on information from rainfall and runoff records. *Journal of Hydrology*, *523*, 636–649. https://doi.org/10.1016/j.jhydrol.2015.01.083

Mei, Y., Anagnostou, E. N., Stampoulis, D., Nikolopoulos, E. I., Borga, M., & Vegara, H. J. (2014). Rainfall organization control on the flood response of mild-slope basins. *Journal of Hydrology*, *510*, 565–577. https://doi.org/10.1016/j.jhydrol.2013.12.013

Merz, R., & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, *39*(12), 1340. https://doi.org/10.1029/2002WR001952

Merz, R., & Blöschl, G. (2008). Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resources Research*, *44*, W08432. https://doi.org/10.1029/2007WR006744

Merz, R., Tarasova, L., & Basso, S. (2020). Parameter's controls of distributed catchment models—How much information is in conventional catchment descriptors? *Water Resources Research*, *56*, e2019WR026008. https://doi.org/10.1029/2019WR026008

Minaudo, C., Dupas, R., Gascuel-odoux, C., & Roubeix, V. (2019). Seasonal and event-based concentration-discharge relationships to identify catchment controls on nutrient export regimes. *Advances in Water Resources*, *131*. https://doi.org/10.1016/j.advwatres.2019.103379

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*(PD), 3249–3262. https://doi.org/10.1016/j.jhydrol.2014.09.089

Perry, M., & Kader, G. (2005). Variation as unalikeability. *Teaching Statistics*, *27*(2), 58–60. https://doi.org/10.1111/j.1467-9639.2005.00210.x

Rabiei, E., & Haberlandt, U. (2015). Applying bias correction for merging rain gauge and radar data. *Journal of Hydrology*, *522*, 544–557. https://doi.org/10.1016/j.jhydrol.2015.01.020

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*. https://doi.org/10.1127/0941-2948/2013/0436

Rinderer, M., van Meerveld, H. J., & McGlynn, B. L. (2019). From points to patterns: Using groundwater time series clustering to investigate subsurface hydrological connectivity and runoff source area dynamics. *Water Resources Reasearch*, *55*, 5784–5806. https://doi.org/10.1029/2018WR023886

Rogger, M., Pirkl, 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

Rogger, M., Viglione, A., Derx, J., & Blöschl, G. (2013). Quantifying effects of catchments storage thresholds on step changes in the flood frequency curve. *Water Resources Research*, *49*, 6946–6958. https://doi.org/10.1002/wrcr.20553

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

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

Sawicz, K. A., Kelleher, C., Wagener, T., Troch, P., Sivapalan, M., & Carrillo, G. (2014). Characterizing hydrologic change through catchment classification. *Hydrology and Earth System Sciences*, *18*(1), 273–285. https://doi.org/10.5194/hess-18-273-2014

Seo, Y., Schmidt, A. R., & Sivapalan, M. (2012). Effect of storm movement on flood peaks: Analysis framework based on characteristic timescales. *Water Resources Research*, *48*, W05532. https://doi.org/10.1029/2011WR011761

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

Sivakumar, B., & Singh, V. P. (2012). Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. *Hydrology and Earth System Sciences*, *16*(11), 4119–4131. https://doi.org/10.5194/hess-16-4119-2012

Slater, L. J., & Wilby, R. L. (2017). Measuring the changing pulse of rivers. *Science*, *357*(6351), 552. https://doi.org/10.1126/science.aao2441

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

Stein, L., Pianosi, F., & Woods, R. (2019). 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., Poncelet, C., & Merz, R. (2018). Exploring controls on rainfall-runoff events: 2. Regional patterns and spatial controls of event characteristics in Germany. *Water Resources Research*, *54*, 7688–7710. https://doi.org/10.1029/2018WR022588

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

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*, *6*(4), 1–23. https://doi.org/10.1002/wat2.1353

Tromp-Van Meerveld, H. J., & McDonnell, J. J. (2006). Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resources Research*, *42*(2), 1–11. https://doi.org/10.1029/2004WR003800

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*(November 2015), 1–16. https://doi.org/10.1016/j.wace.2016.10.001

van Loon, A. F., & van Lanen, H. A. J. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, *16*, 1915–1946. https://doi.org/10.5194/hess-16-1915-2012

Viglione, A., Chirico, G. B., Komma, J., Woods, R., Borga, M., & Blöschl, G. (2010). Quantifying space-time dynamics of flood event types. *Journal of Hydrology*, *394*(1–2), 213–229. https://doi.org/10.1016/j.jhydrol.2010.05.041

Viglione, A., Chirico, G. B., Woods, R., & Blöschl, G. (2010). Generalised synthesis of space-time variability in flood response: An analytical framework. *Journal of Hydrology*, *394*(1–2), 198–212. https://doi.org/10.1016/j.jhydrol.2010.05.047

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography Compass*, *4*(1), 901–931.

Wendi, D., Merz, B., & Marwan, N. (2019). Assessing hydrograph similarity and rare runoff dynamics by cross recurrence plots. *Water Resources Research*, *55*, 1–23. https://doi.org/10.1029/2018WR024111

Woods, R., & Sivapalan, M. (1999). A synthesis of space-time variability in storm response: Rainfall, runoff generation, and routing. *Water Resources Research*, *35*(8), 2469. https://doi.org/10.1029/1999WR900014

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