



## RESEARCH ARTICLE

10.1029/2019WR026511

# A New Unsupervised Learning Method to Assess Clusters of Temporal Distribution of Rainfall and Their Coherence with Flood Types

Henning Oppel<sup>1,2</sup> and Svenja Fischer<sup>2</sup> <sup>1</sup>Center for Environmental System Research, Kassel University, Kassel, Germany, <sup>2</sup>Institute of Hydrologic Engineering and Water Management, Ruhr-University Bochum, Bochum, Germany**Key Points:**

- An unsupervised learning technique is applied to assess clusters of temporal distribution of rainfall
- Sensitivity of the method to parameterization and its results were evaluated and compared to other clustering methods
- The type and spatial distribution of flood types were evaluated for coherence with clusters of temporal distribution of rainfall

**Supporting Information:**

- Figure S1
- Table S1
- Text S1
- Table S2

**Correspondence to:**H. Oppel,  
henning.oppel@rub.de**Citation:**Oppel H., & Fischer, S. (2020). A new unsupervised learning method to assess clusters of temporal distribution of rainfall and their coherence with flood types. *Water Resources Research*, 56, e2019WR026511. <https://doi.org/10.1029/2019WR026511>

Received 8 OCT 2019

Accepted 17 APR 2020

Accepted article online 23 APR 2020

©2020. The Authors.

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

**Abstract** Several factors have an impact on the generation of floods, for example, antecedent moisture conditions and the shape of the catchment. A very important factor is the event rainfall, especially its temporal distribution. However, the categorization of temporal distributions is riddled with uncertainty, due to a priori assumptions on distribution types. Here, we propose a new clustering approach based on unsupervised learning, using dimensionless mass curves to describe the temporal distributions. The purpose of the proposed method is the identification of reoccurring temporal distributions of precipitation events. Additionally, the correlation of the resulting clusters of temporal distribution with rainfall-induced flood types is investigated. The application to several catchments in Germany showed the existence of spatial patterns of six different clusters for the temporal distribution and a significant coherence with the flood types. It was found that the temporal distribution of rainfall intensities shifts from early peaked to more uniformly distributed when shifting the flood type from short floods with high peaks to long-duration floods, often with several peaks.

## 1. Introduction

Floods belong to the most disastrous hazards in the world and lead to several billion dollars of damage every year (Hattermann et al., 2014). Understanding their generation and development is paramount to avoid future damages. However, floods are very diverse in their origin and generating processes. These processes control their timing, magnitude, duration, and shape. A possibility to increase the understanding of flood generating processes and the coincidences between them during a flood event is the flood typology (Blöschl, 2006). A flood typology can help to improve flood statistics, regionalization, and detection of changes (Hirschboek et al., 2000; Merz & Blöschl, 2008). Many different flood typologies exist; they depend on very different scales, time periods, and data. Tarasova et al. (2019) identified three different approaches of flood typologies: the hydroclimatic, the hydrological, and the hydrograph-based flood typology. While hydroclimatic approaches make use of weather systems and define flood types accordingly (e.g., tropical cyclone, orographic rainfall, etc.), hydrological typologies include hydrometeorological variables as well as the catchment state. In contrast, hydrograph-based typologies focus on the shape of the flood wave and its characteristics. While the perspectives differ, all approaches agree that precipitation plays a crucial role for defining flood types (Merz & Blöschl, 2003; Nied et al., 2014; Sikorska et al., 2015). For this purpose, thresholds based on the event precipitation have been developed in order to define flood types, based, for example, on the precipitation duration, intensity, or amount (Keller et al., 2018; Nied et al., 2017). However, only few studies connected the temporal distribution of rainfall and the hydrological processes shaping the resulting floods. Yet the impact of the temporal distribution on the shape of the hydrograph as well as the impact on peak and volume of a flood event is well known (Fischer et al., 2019). To close this gap, we propose a method that automatically clusters precipitation events according to their temporal distribution. Furthermore, we link these results to flood types. For design floods, different distributions of rainfall are already considered to account for different hydrograph shapes, emphasizing the importance of such an approach (Ball, 1994; Pilgrim & Cordery, 1975).

The key question for design flood estimation and the search for a link between flood type and precipitation event is what types of temporal distributions can be found frequently within precipitation event data?

Huff (1967) introduced the concept of quartile storms. In his work, he transformed rainfall records to dimensionless mass curves, normalizing the event duration and precipitation amount. These curves were grouped by the occurrence of the highest precipitation intensities. For example, a temporal distribution with the highest amount of precipitation falling in the first 25% of the storm duration was considered a first-quartile storm. Further studies of temporal distributions of precipitation events in different regions adopted this concept of quartile storms (de Araújo & González Piedra, 2009; Dolšák et al., 2016; Vandenberghe et al., 2010; Yin et al., 2016). Other studies applied a grouping of temporal distributions comparable to Huff (1967) and Pilgrim and Cordery (1975). They assumed three basic types of distributions: equally distributed, early peaked, and late peaked (Dunkerley, 2012; Schulte et al., 2013; Wu et al., 2006). Wu et al. (2006) and Schulte et al. (2013) introduced a further differentiation of these basic types to create a finer grouping. However, all these studies assumed that the types could be described with a single-peaked event, though temporal distributions with multiple peaks occur and can be used as an additional type of distribution (Wang et al., 2018).

The variety in these studies reveals that the definition of precipitation types is impacted by uncertainty. The classification of temporal distributions based on assumed classes can be misleading, because the a priori estimation might oversee patterns within the data that are present yet unexpected and remain thus simply unexplored. To recognize these patterns, unsupervised learning can be applied (Alpaydin, 2016). Different types of clustering methods that search for naturally occurring groups within data sets are available. Partitioning methods rely on nearest neighbor analysis and split a data set into  $k$  clusters. The number of clusters  $k$  is iterated and in each step evaluated with criteria like the silhouette coefficient (Rousseeuw, 1987). These methods are particularly useful for clusters that are sharply separable, that is, with a low amount of noise. Hannah et al. (2000) and Ternynck et al. (2016) applied this type of technique for hydrograph clustering, Keller et al. (2018) to flood event characteristics. A drawback of partitioning methods is that they are only suitable for spherical clusters (Han & Kamber, 2010; Runkler, 2015). Density methods have been developed to identify clusters with an arbitrary shape in  $n$ -dimensional feature space. Methods like density-based spatial clustering of applications with noise (DBSCAN; Ester et al., 1996) and ordering points to identify the clustering structure (OPTICS; Ankerst et al., 1999) view clusters as dense regions in the feature space that are separated by regions of low density. Data points within these low-density regions are treated as noise. These methods do not require an a priori assumption of  $k$ , the number of clusters. Hierarchical cluster methods follow an agglomerative approach. At initialization, each data point is treated as an individual cluster. With each step, closely located clusters are merged into a single cluster. Following this procedure, the number of clusters is reduced step-wise until all data points are merged within one cluster. The choice of a metric to define the distance between the clusters as well as the level to stop the agglomeration are the critical parameters of a hierarchical clustering. The sensitivity of these parameters induces their often inferior results compared to other techniques (Gelbard et al., 2007; Hancer & Karaboga, 2017). The balanced iterative reducing and clustering using hierarchies (BIRCH) algorithm (Zhang et al., 1996) was designed to overcome the difficulties of agglomerative clustering. BIRCH reduced the clusters to their statistical characteristics, like cluster center and diameter of the cluster. Based on a distance and member thresholds, the algorithm builds a tree structure of clusters and subclusters. Yet BIRCH mainly performs a dimensionality reduction, just as self-organizing maps (Hall & Minns, 1999; Pelletier et al., 2009) and requires a subsequent application of another clustering technique with an a priori estimate of  $k$  to lower its sensitivity to parameterization.

Since, in general, the number and shape of temporal distribution types of rainfall is unknown, we propose to apply unsupervised learning to identify naturally occurring clusters of temporal distributions. Therefore, we clustered dimensionless mass curves from a large data set of rainfall-runoff events observed at multiple catchments. The clusters were subsequently linked to the flood type of the respective flood events. However, the design of this study excluded a majority of the approaches that were previously mentioned: iterative partitioning methods were excluded due to the high dimensionality of the feature space (dimensionless mass curves with 10 or more sampling points) (Houle et al., 2010), hierarchical methods due to their inferior results and high sensitivity (Gelbard et al., 2007). Density methods were considered as a valid choice. However, density methods require certainty about the feature space and are very parameter sensitive. Hence, we propose a different approach for the identification of frequent patterns. Our algorithm is based on the similarity of temporal distributions, taking into account the divergence of the curves to the entire data set. By increasing the level of abstraction, using similarity rather than the feature space for pattern recognition, we intend to lower the uncertainty of the feature space and the utilized distance metric. The similarity measure used in a two-step clustering technique identifies the number of clusters and initial cluster centers in

the first step with an overlapping agglomerative approach. These estimates are used in the second step for an exclusive classification of the temporal distributions. The proposed method is introduced in detail in section 3 together with the applied flood typology. Results of the clustering procedure are given in section 4. A sensitivity analysis and a comparison to other cluster approaches are included in section 5.

After the definition of the types of temporal distributions with unsupervised learning, their coherence with the flood types is assessed. In order to clarify the connection of the temporal distribution of event precipitation with the flood type, predominantly designated coherences of distribution types and flood types are evaluated. The connection is analyzed statistically with the Wilcoxon test. Additionally, we assessed the spatial arrangement of temporal distributions. The case study region and the data used in this study are presented in section 2. The assessment of the coherence between temporal distribution of rainfall and flood types is presented in section 4. Conclusions are given in section 5. The main research questions addressed are as follows:

1. How can the temporal distribution of rainfall be clustered without a priori assumptions on the number of clusters?
2. How are the resulting clusters of the temporal distribution of rainfall connected to flood types?

## 2. Data Basis

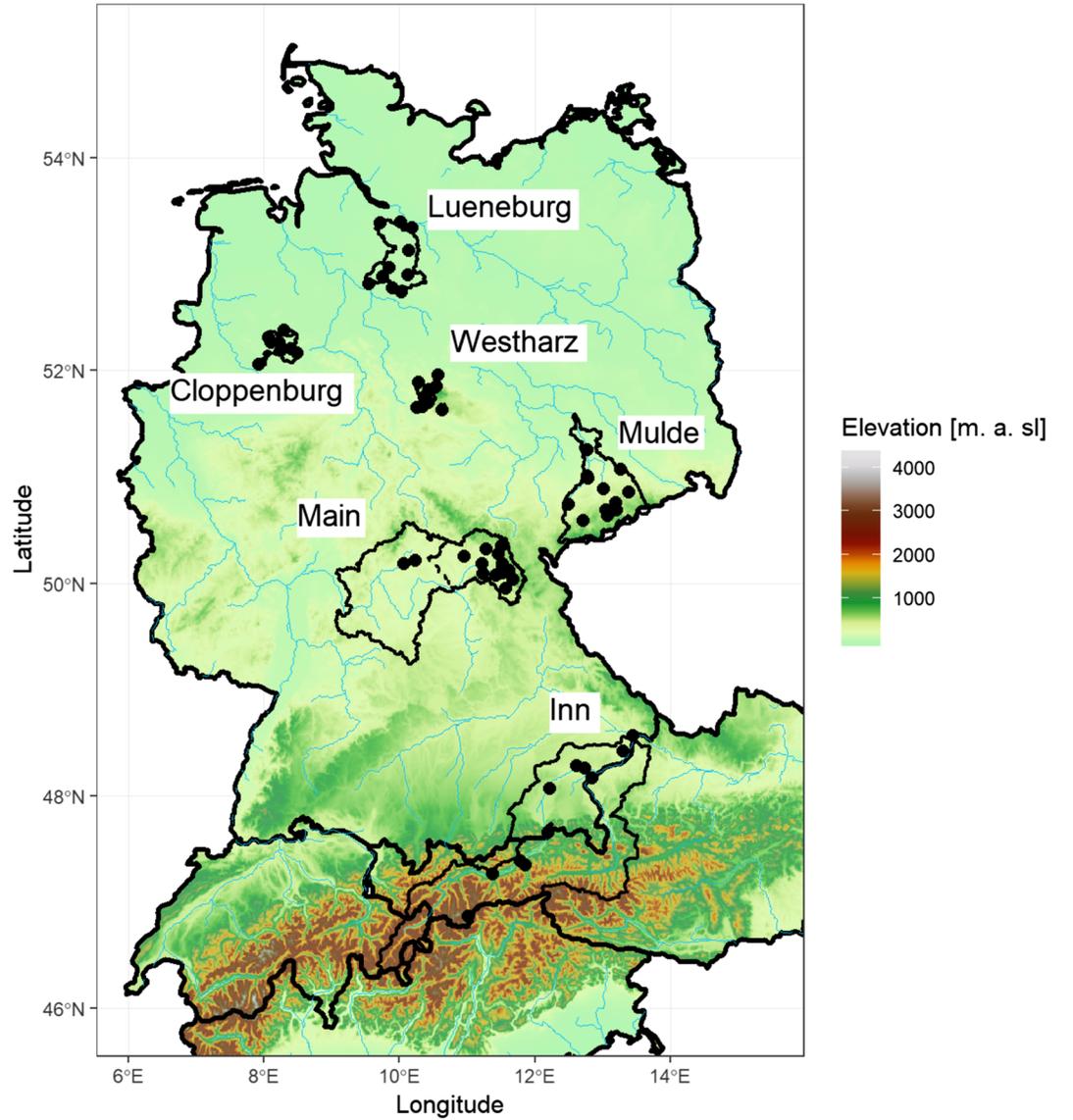
To consider the broad spectrum of flood types and the impact of catchment topography, diverse catchments were investigated. Six different basins were considered, covering different parts of Germany: Cloppenburg and Lueneburg in the North, Mulde in the East, Westtharz and Main in the Center, and the Inn basin at the very South, covering also parts of Austria and Switzerland. With this selection, different climatic and topographic characteristics were covered, for example, alpine, mountainous, flatland, and heather. An overview of the location of the basins is given in Figure 1, and a table with catchment characteristics is given in the supplement. In total, we considered 66 gauges with catchment sizes varying between 17 and 25,941 km<sup>2</sup> and observation lengths between 32 and 105 years. Discharge was available in daily resolution (means) as well as monthly resolution (instantaneous peaks). Daily precipitation was obtained by the data set of the German Weather Service (available at [www.dwd.de](http://www.dwd.de)) respectively from eHYD ([www.ehyd.gv.at](http://www.ehyd.gv.at)) for the Austrian parts of the Inn and transformed to areal precipitation by the Thiessen method with a 40-km buffer radius to include precipitation stations for each catchment. A list of all stations used in this study is given in the supplement. Although in mountainous catchments Kriging is often preferred to Thiessen polygons, we decided to apply the Thiessen method here in order to preserve the natural variability of the observed precipitation between the stations. The use of a geostatistical approach like Kriging would have resulted in a loss of variance in the temporal distributions. Though height differences within the catchments led to biased precipitation volume estimates, no height correction or other interpolation method has been applied. Since we use normalized precipitation, volume errors have been accepted. Instead, the temporal distribution is more important, and hence, Thiessen polygons are preferred for areal precipitation. The daily resolution was chosen due to its long-term and wide-area coverage. Nevertheless, this decision leads to the assumption that some hydrographs could not be fully understood, especially those from small catchments. Since we concentrated on flood types (see section 3.3) and dimensionless mass curves (section 3.1), we considered this disadvantage as an acceptable limitation compared to the added value of the large and extensive data basis of this work. This data basis is necessary to obtain reliable information on correlation and attribution of the clusters of temporal rainfall and the flood types. Nevertheless, we have to point out the discrepancy between the temporal resolution of the input data and the (assumed) reaction times of the smaller catchments of this study.

For daily snowmelt data, the Hydrologiska Byrans Vattenavdelning (HBV)-light model by Seibert and Vis (2012), based on the degree-day method was applied. For this data basis, 4,647 flood events have been identified with respective event precipitation, using the method of Fischer et al. (2019), implemented in the FloodR package (see also section 3.3).

## 3. Methods

### 3.1. Characterization of Temporal Rainfall Patterns

A common procedure to extract the temporal distribution of rainfall events from rainfall intensity data is the use of dimensionless mass curves, established by Huff (1967). Precipitation records are commonly



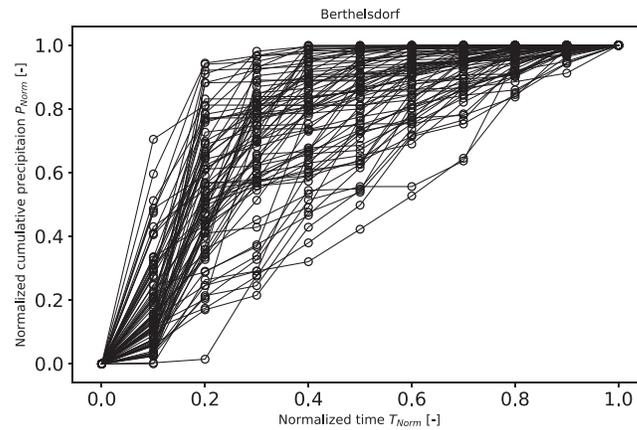
**Figure 1.** Geographical position of gauges and corresponding basins in Germany that were used in this study.

available as sums of precipitation between equidistant time steps, in this case 1 day, that is, a succession of rainfall intensities. Each event  $i$  is described by two vectors, first the precipitation intensities  $P_i = (p_{1,i}, p_{2,i}, \dots, p_{Dur_i,i})$  and second the according time stamps  $T_i = (0, 1, \dots, Dur_i)$ . Note that the time stamps always begin with 0 at the beginning of the event and end with  $Dur_i$ , the duration of the event  $i$ . To calculate the dimensionless mass curves, each element  $P_{i,j}$  and  $T_{i,j}$ ,  $j = 1, \dots, Dur_i$ , of both vectors was normalized as follows:

$$T_{Norm;i,j} = T_{i,j} / Dur_i, \quad (1)$$

$$P_{Norm;i,j} = \frac{\sum_{k=0}^j P_{i,k}}{\sum_{k=0}^{Dur_i} P_{i,k}}. \quad (2)$$

Both new vectors contain cumulative percentages of the event duration (in case of  $T_{Norm;i}$ ) or the event volume (in case of  $P_{Norm;i}$ ) and are increasing monotonically in  $[0, 1]$ . To reduce the feature space of the clustering problem, a number of  $s$  sampling points is defined. The sampling points are equally distributed within  $[0, 1]$  and represent normalized time stamps. The corresponding normalized precipitation values



**Figure 2.** Temporal distribution of precipitation events at gauge Berthelsdorf, Mulde basin. Distribution of precipitation intensities is expressed as cumulative percentages of precipitation volume and time.

were extracted, that is, interpolated from  $P_{Norm,i}$ . The dimensionless mass curves are shown with 10 equally distributed sampling points, see Figure 2 for example dimensionless mass curves of all precipitation events causing floods at gauge Berthelsdorf in the Mulde basin.

The choice of the number of sampling points is crucial for the representation of the temporal distribution. An increasing number of sampling points increases the accuracy of the represented temporal distribution but increases the dimensionality of the feature space. With an increasing number of dimensions, the complexity of the clustering problem increases. Another aspect to consider is that the number of sampling points defines the minimum length of the original precipitation event. All events with less observations than the chosen number of sampling points have to be excluded from the study. The reason is that the extraction of more sampling points than available data points adds artificial variance to the dimensionless mass curve that could bias the outcome of the cluster analysis.

Alternatively, a distribution function could have been fitted to each empirical distribution. The use of theoretical distribution functions results in smooth temporal distributions. Intensity bursts or rainfall pauses might be erased from the temporal distribution. The parsimony of this approach can thus rather be a drawback. Hence, the empirical mass curves were used in this study. The number of sampling points has been determined with an iterative analysis. The number has been altered from four to 20 sampling points. For each number of points, the variance of the distribution in comparison to the variance of the curve using all available points has been determined, as well as the number of remaining events in the data set. Ten sampling points offered the best solution for the trade-off between variance reproduction and remaining data.

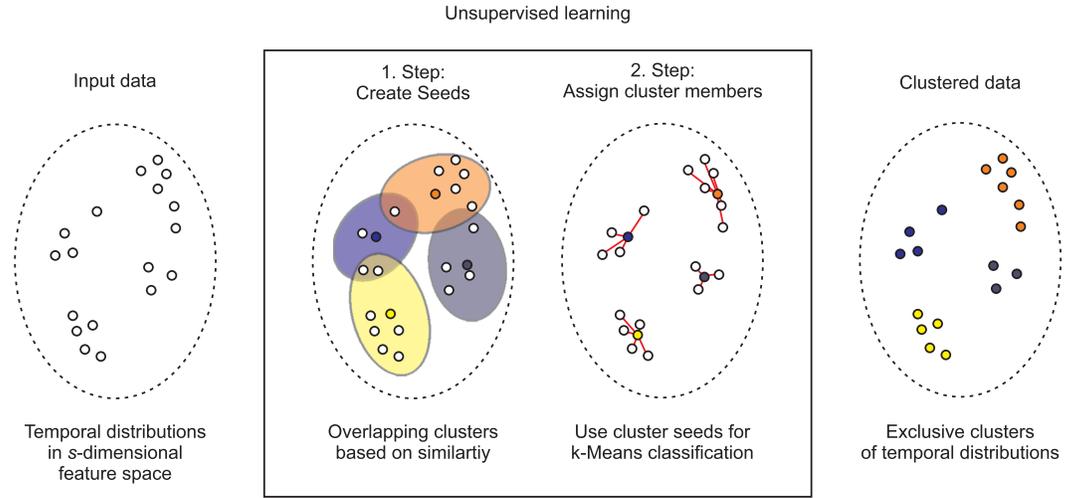
### 3.2. Curve Clustering

A two-step clustering technique was developed and has been applied to the data set of dimensionless mass curves. The schematic procedure is shown in Figure 3. In the first step, the learning step, the number of cluster seeds  $k$ , and initial seeds for their cluster centers were identified within the input data. The criterion for the identification of the seeds was curve similarity. These seeds were subsequently used for  $k$ -means classification, returning an exclusive classification of the given input data.

The input data were the  $n$ -dimensionless mass curves of our data set, each with  $s$  sampling points. To assess the similarity of these curves, the Euclidean distance between each curve  $\delta(x, y)$  was calculated as follows:

$$\delta(x, y) = \sqrt{\sum_{i=0}^s (x_i - y_i)^2} \quad (3)$$

with  $x_i$  and  $y_i$  being the  $i$ th sampling points of two dimensionless mass curves  $x$  and  $y$ , respectively. After applying equation (3) to all  $n$  curves, the  $n \times n$  distance matrix  $D$  was obtained. Each row of matrix  $D$  displays the dissimilarity of one event to all other events of the data set. Because distance, as calculated with equation



**Figure 3.** Schematic depiction of the unsupervised learning strategy, a two-step approach to cluster the given dimensionless mass curves. In the first step, the algorithm defines the number of clusters and cluster centers that were used subsequently for classification of the data with the  $k$ -means algorithm.

(3), is the complement to similarity, we calculated the linear correlation  $\rho(p, q)$  between the rows of matrix  $D$ :

$$\rho(p, q) = \frac{Cov(p, q)}{\sigma(p) \cdot \sigma(q)} \quad (4)$$

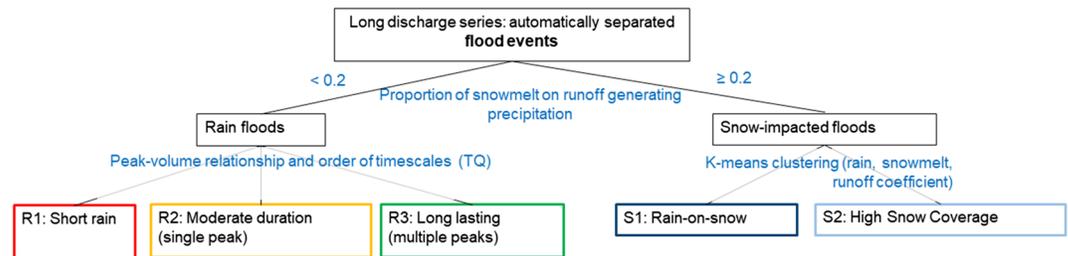
with  $\sigma(p)$  being the variance of row  $p$  of matrix  $D$  and  $Cov(p, q)$  being the covariance between the rows  $q$  and  $p$ . After application of equation (4) to all pairs of rows of matrix  $D$ , a symmetrical  $n \times n$  correlation matrix  $C$  was obtained. Each element of matrix  $C$  represented the correlation of two dimensionless mass curves, based on their divergence to the remaining data set. We defined that two events that differed similarly from the remaining data set could be considered as similar. Hence, the value of  $\rho$  was interpreted as a similarity measure between the mass curves.

Based on this definition and on the correlation matrix  $C$  we performed an agglomerative, overlapping clustering. A threshold  $\xi_1$  was introduced to define the level of agglomeration, taking values within  $[0, 1]$ . For each row  $p$  of  $C$ , all elements that were larger or equal to threshold  $\xi_1$  were determined. These sets of curves were combined in vector  $K$ :

$$K = \begin{bmatrix} \{i | 0 < i < n; c_{1,i} > \xi_1\} \\ \vdots \\ \{i | 0 < i < n; c_{p,i} > \xi_1\} \\ \vdots \\ \{i | 0 < i < n; c_{n,i} > \xi_1\} \end{bmatrix}, \quad (5)$$

where  $c_{p,i}$  is the entry in row  $p$  and column  $i$  of matrix  $C$  with  $p$  and  $i$  taking values within  $[0, n]$ .

Each element of vector  $K$  is a set of similar dimensionless mass curves. Note that vector  $K$  has  $n$  elements at this point. To reduce the number of clusters, we eliminated subsets within vector  $K$  and agglomerated clusters by their centroid distance. Clusters with a centroid distance smaller than a threshold  $\xi_2$  were merged. Threshold  $\xi_2$  was defined as the  $q$ th quantile of matrix  $D$  and was set to 5% in this study. Note that other  $q$  values within the range of  $[1, 15]$  were tested. The sensitivity of both parameters,  $\xi_1$  and  $\xi_2$ , will be further discussed in section 4.1. After the reduction of vector  $K$ , the vector contained a number of  $k$  clusters. The clustering of the curves was overlapping at this point of the procedure, that is, curves were assigned to multiple clusters. In order to obtain an exclusive classification, we discarded all cluster assignments and used the  $k$  cluster centers as seeds for a (supervised)  $k$ -means classification, based on the distance calculated with equation (4). For implementation details of the  $k$ -means we refer to Pedregosa et al. (2011).



**Figure 4.** Decision tree for the flood typology.

### 3.3. Flood Typology

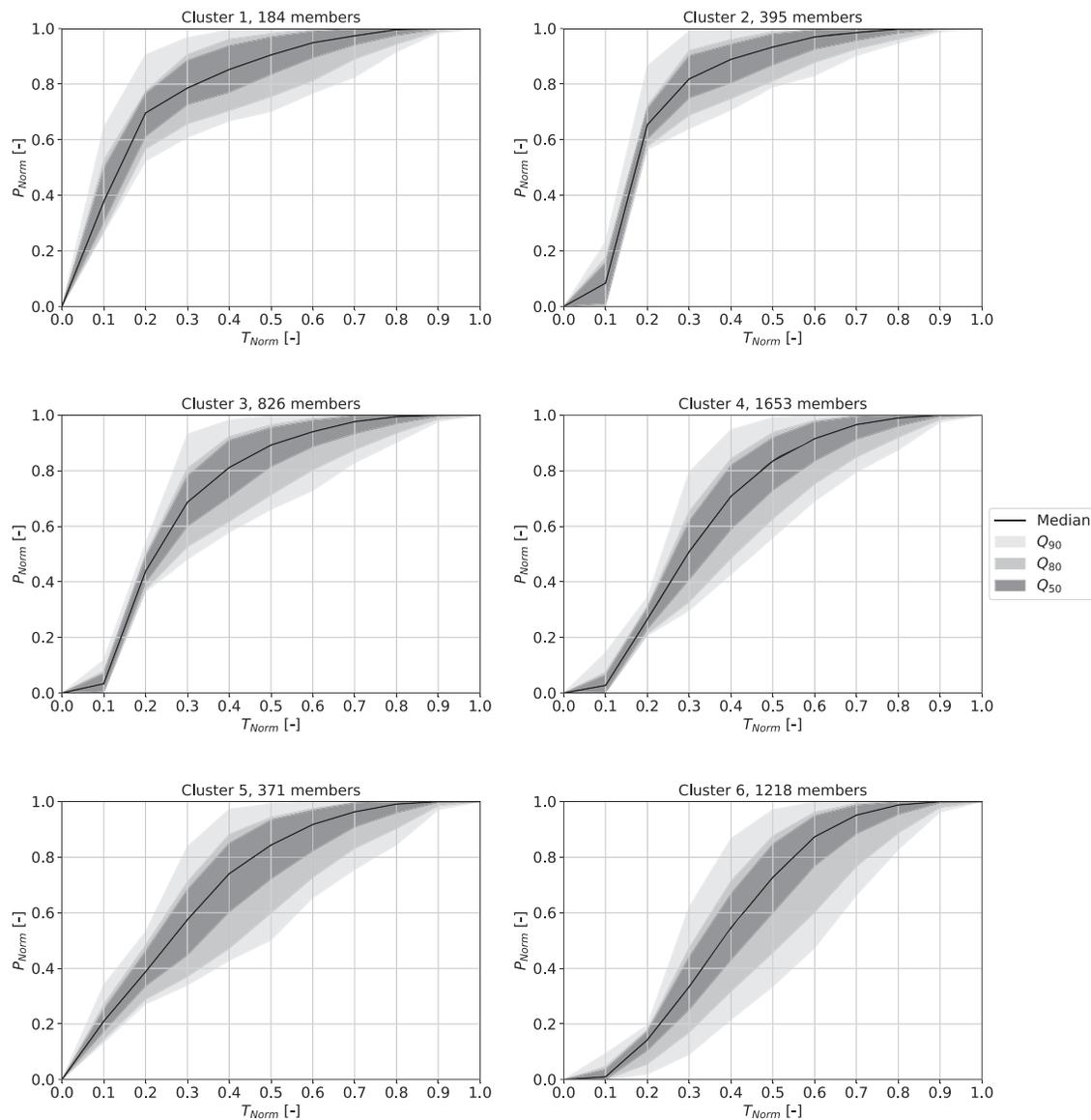
We applied the flood typology proposed by Fischer et al. (2019). It is hydrograph-based (Tarasova et al., 2019). The typology is based on a decision tree, leading to five different flood types (Figure 4): short rain flood (R1), moderate duration rain flood (R2), long-lasting rain flood (R3), rain-on-snow flood (S1), and floods caused by snow melt with a high snow coverage (S2). The categorization is based on few characteristics (snow melt, precipitation, peak, and volume), uses a daily time step, and is thus suitable for long data series where time series of, for example, soil moisture are not available. This made it preferable for our purpose. For this study, only flood events of the first three types (R1, R2, and R3) were considered, since these are the rainfall-induced flood events. The three rainfall-induced flood types show significant differences in their hydrograph shape and flood characteristics. While R1 events tend to be very short with high peaks and small volumes, R2 events show moderate peaks and moderate volume, and R3 events are long lasting with very high volume and often small but multiple peaks. For these flood types, it was found that the hydrological processes differed in several ways but were distinguishable by the rainfall and the runoff coefficient (Fischer et al., 2019). In the alpine basin of the Inn river and in parts of the Mulde and Westharz basins, snow-impacted events contribute to floods. For example, in the Inn basin, about 20% and in the Mulde basin 30% of all observed flood events belong to flood types S1 or S2. However, since our goal is to detect coherence between flood type and temporal rainfall distributions, these events were excluded here and only the rainfall-induced types (R1, R2, and R3) are considered in the following analysis.

Please note that the flood typology is not affected by the temporal resolution of the data. Since the timescale is applied based on peak and volume (both are the same for hourly and daily data) and only events with duration longer than the chosen sample points are considered, the same flood types result for hourly and daily data.

## 4. Results and Discussions

Precipitation events from all catchments used in this study were transformed to dimensionless mass curves and were given as input data to the unsupervised learning routine. Depending on the parametrization, three to 15 clusters have been returned by the algorithm. Six clusters were obtained with a parametrization of  $\xi_1$  set to 0.7 and the  $q$  value for  $\xi_2$  set to 5%. All clusters are shown in Figure 5 as cluster centers and confidence intervals of the assigned curves. The parameterization has been chosen mainly to allow a certain degree of variance between the clusters. Furthermore, it accounts for findings of preceding studies of temporal distributions (e.g., Pelletier et al., 2009; Schulte et al., 2013). However, the choice of the parameters was subjective and required discussion. The effect of parametrization on the clustering result will be discussed in the following section 5.

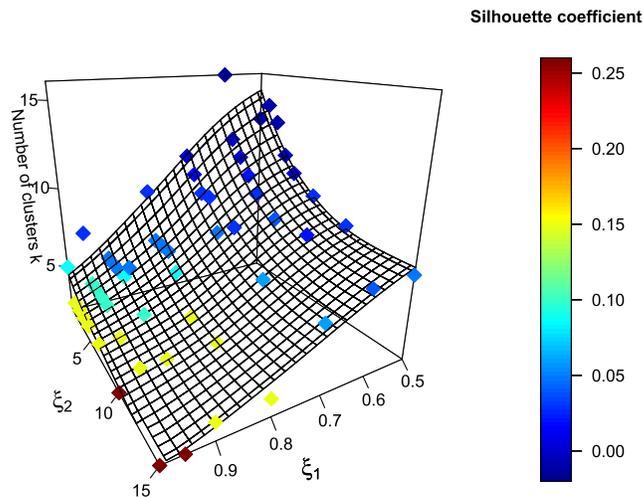
Rainfall events assigned to Cluster 1 have an intensity burst at the beginning of the event. Within the first 10% of the event duration, 30%–60% of the precipitation sum fall. After 20% of the event duration, rainfall intensities decrease sharply. Events in Cluster 2 also start with an intensity burst at the beginning of the event, but have a start-up phase that lasts 10% of the event duration. This phase is followed by a short intensity burst that diminishes after another 10% of event duration. Within this short time period, approximately 60% of the total precipitation amount fall. A comparable course of the precipitation event characterizes Cluster 3. Again, a 10% duration start-up period is followed by an intensity burst, but the intensities are lower than in Cluster 2 and attenuate step-wise and over a longer period. Clusters 4 and 5 characterize events that have an intensity emphasis on the center of the event or are nearly equally distributed. Cluster 4 has the common



**Figure 5.** Clusters of temporal distributions obtained by unsupervised learning. Cluster centers are shown as median (black line) and in confidence intervals for 50%, 80%, and 90% (gray scales). Cluster frequencies are given in each panel header.

start-up phase of 10% of event duration and shows nearly constant intensities from 10% to 40% event duration and 40% to 80% precipitation amount. The numbers are similar for Cluster 5 but without the start-up phase. These events showed nearly constant intensities from the beginning. Events with a late intensity peak were merged in Cluster 6. The start-up takes about 20% of the event duration and only 10%–20% of precipitation amount fall during this phase. It is followed by an increase of the intensities that stays constant to 60% of the event duration. The tail of decreasing precipitation intensities of these late peaked events is caused by the event separation applied in this case study. We extracted the full event precipitation up to the time where intensity falls below 5 mm per day for the first time rather than the precipitation leading to the main peak of the event. Hence, smaller precipitation amounts following the main intensity bursts are within the event records, causing the tails. The precipitation amounts after the peak of the event are especially important for understanding flood events of long duration with multiple peaks, as considered in type R3.

The frequencies of the clusters in the data set are given additionally in Figure 5. It is striking that clusters that are linked to the highest precipitation intensities (Clusters 1 and 2) were among the clusters with the lowest



**Figure 6.** Impact of parameterization ( $\xi_1$ ,  $\xi_2$ ) on the number of clusters  $k$  defined and the silhouette coefficient in the feature space. A hyperplane is shown as wire frame fitted to the sample points of the parameter space.

frequency (184 or 395 out of 4,629 events). The most frequent temporal distribution has a slight emphasis on the event center with a short start-up phase (Cluster 4).

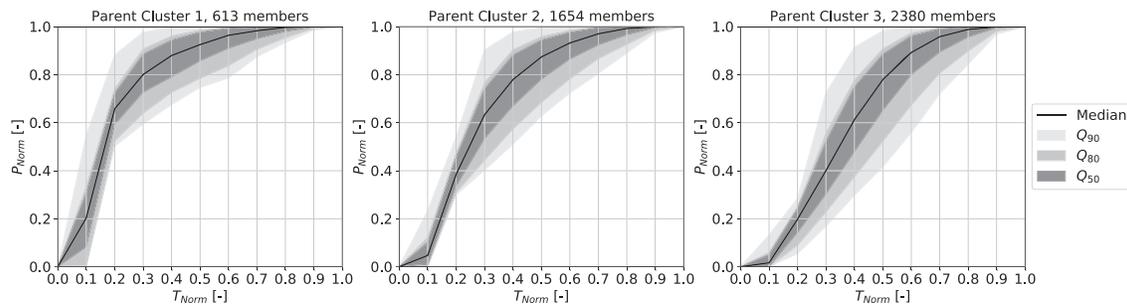
To assess the performance of the proposed unsupervised learning technique, some critical points need to be discussed. First, the effect of the parameters  $\xi_1$  and  $\xi_2$  on the cluster results has to be evaluated. A sensitivity analysis was performed (section 4.1). Second, it has to be discussed if the sensitivity and results of the proposed method are comparable to other clustering methods. Results of other methods and a comparison to the our results are presented in section 4.2. However, the quality of a clustering technique lies within its ability to carry out meaningful subdivisions of the data independently from the parameterization. Hence, the discussions are concluded in section 4.3 with an assessment of the coherence between our clusters of temporal distributions of rainfall and flood types.

#### 4.1. Parameter Sensitivity

The proposed method involved two parameters. Parameter  $\xi_1$  defines a threshold for the linear correlation indicating the similarity between two events. It was used to build up the connections between events that were used to find the cluster seeds. An increase of  $\xi_1$  was expected to decrease the number of clusters  $k$ . A second threshold  $\xi_2$  defines cluster seeds that are too closely located in the feature space. The value of  $\xi_2$  was determined as the  $q$ th quantile of the distance matrix  $D$ . An increase of  $q$ , that is, of parameter  $\xi_2$ , was expected to decrease  $k$ . A sensitivity analysis varying one factor at a time has been carried out. The number of clusters  $k$  in dependence to the chosen parameterization is shown in Figure 6. We used the data points ( $\xi_1$ ,  $\xi_2$  and  $k$ ) to interpolate a hyperplane, generalizing the effect of the parameters on  $k$ , see the fitted hyperplane is shown in Figure 6.

From Figure 6 we concluded that our expectations have been met. Parameter  $\xi_1$  did increase the number of clusters  $k$  found in the data. Also, an increase of  $\xi_2$  resulted in a lower number of clusters  $k$  found in the data. Colors in Figure 6 represent the values of the silhouette coefficient (Rousseeuw, 1987), calculated in the 10-dimensional feature space of the dimensionless mass curves, using equation (3) as evaluation metric. It is visible that an increasing number of clusters is associated with a decrease of the coefficient. However, the range of the silhouette coefficients was comparably small regarding the complete range of  $S$ ,  $[-1, 1]$ . Keller et al. (2018) obtained similar results for their classification and attributed the small range of  $S$  to the decrease of explanatory power of distance metrics in higher dimensions of the feature space (Houle et al., 2010). A problem becomes visible in this case, too; although differences within the silhouette coefficients are visible, they are not significant compared to the uncertainty of the distance metric and the dimension of the feature space.

Like all other unsupervised learning techniques, the proposed method requires parameterization. The choice of the parameters affects the number of clusters delineated by the algorithm. As Figure 6 showed,



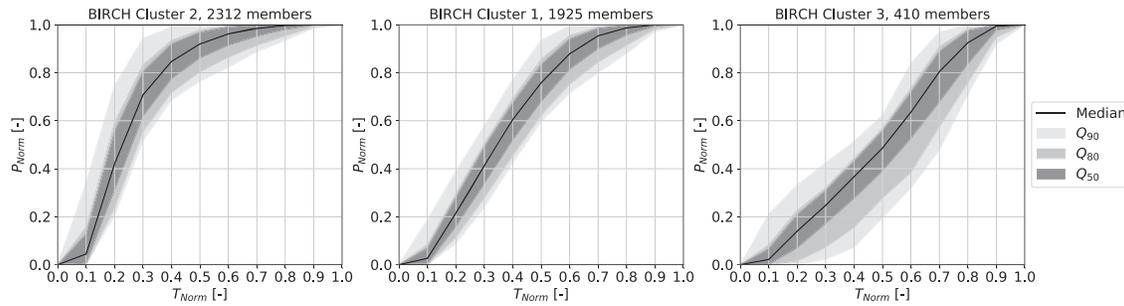
**Figure 7.** Parent clusters of temporal distributions obtained by unsupervised learning with  $\xi_1 = 95$  and  $\xi_2 = 1\%$ . Cluster centers are shown as median (black line) and in confidence intervals for 50%, 80%, and 90% (gray scales). Cluster frequencies are given in each panel header.

three to 15 clusters were defined within the parameter boundaries. In order to set this result into perspective, a comparison to other clustering techniques will be presented in the following section 4.2. To assess if certain levels of clustering detail, created with different parametrizations, are interconnected, two cluster results were compared. The lowest number of clusters defined by the algorithm was compared to the results presented in section 4. Figure 7 presents the minimum number of clusters  $k = 3$  obtained with  $\xi_1 = 0.95$  and  $\xi_2 = 1\%$ . These clusters were labeled as *parent clusters* because they could be connected to the clusters shown in Figure 5. Parent Cluster 1 can be interpreted as the early peaked temporal distribution, merging Clusters 1 and 2 from Figure 5. Center-peaked distributions were merged in Parent Cluster 2. In a finer resolution, this parent cluster splits into Clusters 3 and 5, as shown before. Finally, Parent Cluster 3 merged late peaked events and includes temporal distributions from Clusters 4 and 6. This cluster nexus is reflected in the number of clusters and parent cluster members with the exception that approximately 500 curves migrated from Parent Cluster 2 to Cluster 4 in the finer resolution. These observations showed that the parameterization of the algorithm affects the degree of clustering detail. Depending on parameters  $\xi_1$  and  $\xi_2$ , the algorithm either returns coarse clusters, merging the basic shapes of temporal distributions, or subclusters of these basic shapes with a greater variety of shapes.

#### 4.2. Comparison to Other Clustering Methods

To evaluate the performance of the proposed method, we repeated the clustering of the dimensionless mass curves with established methods. We chose methods that met our basic requirement, that is, to perform the clustering without a priori estimation of  $k$ . This reduced our options to two basic concepts: first, density-based methods such as DBSCAN and OPTICS. DBSCAN (Ester et al., 1996) separates clusters by their density in the  $s$ -dimensional space. However, DBSCAN is known to be very parameter sensitive, especially in high-dimensional feature space (Han & Kamber, 2010). In our application this proved to be true. DBSCAN returned one to eight clusters but treated 98% of the data as noise. OPTICS (Ankerst et al., 1999) was designed to overcome the parameter sensitivity of DBSCAN but also failed in our application (again more than 96% of the data was treated as noise). The dependencies between the  $s$  dimensions of the feature space might explain the inferior results of the density-based methods.

The second algorithm applied was BIRCH (Zhang et al., 1996) but without postprocessing with a supervised classification algorithm. BIRCH performs clustering and dimension reduction by hierarchical ordering of the data. Samples are drawn based on the Euclidean diameter around possible clusters. If the diameter of a cluster is too large, as defined by threshold  $T$ , a new cluster is introduced. If the number of parallel clusters exceeds the number of  $B$ , the second parameter, a new branch is introduced. Each cluster of the first level is then split into a maximum of  $B$  child clusters on a second level of the cluster feature tree. Note that BIRCH does not store the individual samples but so called *cluster features* that statistically describe the individual clusters. Without postprocessing, the lowest level child clusters were treated as results of BIRCH. In the application to the dimensionless mass curves, BIRCH proved to be very sensitive to parameter  $T$  but insensitive to  $B$ . Threshold  $T$  in this case was defined within  $[0, 1]$ , but values above  $T > 0.4$  resulted in all curves merged into a single cluster. With  $T = 0.4$ , we obtained three clusters and silhouette coefficient in the feature space of the dimensionless mass curves of 0.32. These clusters are shown in Figure 8. A lower boundary of  $T = 0.2$  returned 47 clusters and a silhouette coefficient of 0.15. Between these boundaries, the number of clusters decreased in an almost linear way. The results showed that BIRCH was significantly more sensitive to parameterization. Parameter  $T$  has been altered on a small range but had significant impact



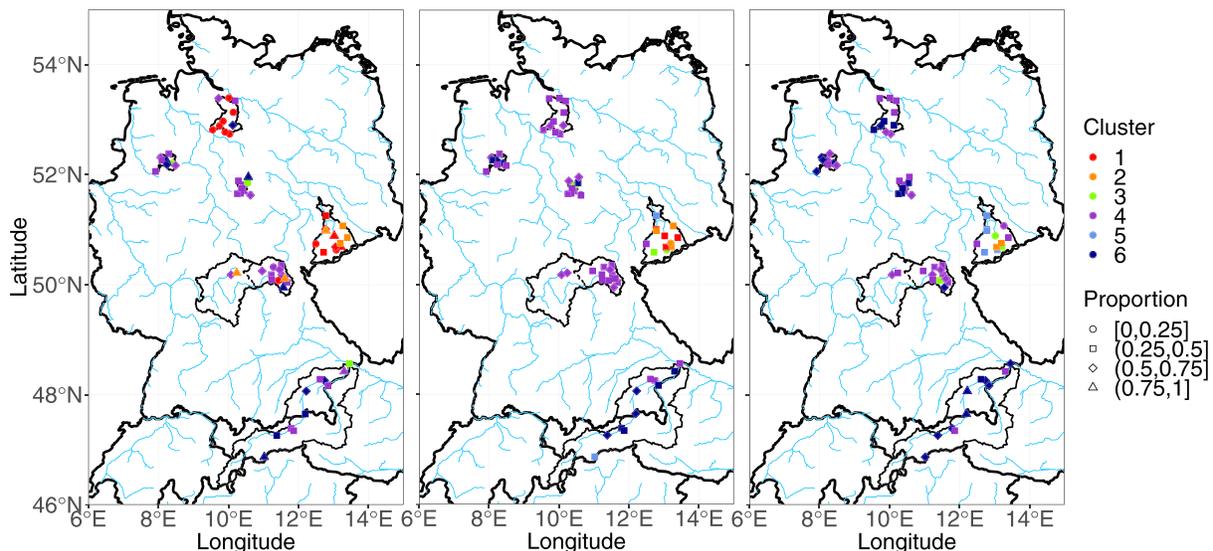
**Figure 8.** Clusters of temporal distributions obtained by balanced iterative reducing and clustering using hierarchies (BIRCH) with  $T = 0.4$ . Cluster centers are shown as median (black line) and in confidence intervals for 50%, 80%, and 90% (gray scales). Cluster frequencies are given in each panel header.

on the cluster results. The number of clusters defined increased stronger than for the clustering method introduced in this study. Moreover, the number of clusters defined by the proposed methods never dropped below three. This allowed to define *parent* clusters. The comparison of the parent clusters (Figure 7) and the BIRCH clusters (Figure 8) showed another difference between the algorithms; while the proposed algorithm is more sensitive to the shape of the curves, BIRCH is more sensitive to the position of the curves in the feature space. This difference is a result of the different interpretation of similarity. While BIRCH relies on cluster densities, the proposed new algorithm relies on correlation.

This analysis showed that unsupervised learning requires parameterization. Hence, uncertainty due to user choices is present in all approaches. Yet we could show that the proposed algorithm is less sensitive to parameterization than other algorithms. Moreover, it proved its superiority to density-based methods in the complex feature space of temporal distributions. We could also show that, in contrast to other algorithms, applicable to temporal distributions, the proposed method is focused on curve shape rather than positions in the (uncertain) feature space.

### 4.3. Coherence between Precipitation Cluster and Flood Type

The six clusters, obtained with the  $\xi_1 = 0.7$  and  $\xi_2 = 5$  parametrization, showed quite different shapes and temporal processes. Hence, the different types of temporal distributions should have led to different flood types. Naturally, catchment characteristics, for example, topography and initial states such as soil moisture, play a crucial role in the definition of the flood types. One would expect short rain floods (R1) mainly be caused by rainfall events with large intensities at the beginning of the event, whereas long-duration rainfall floods (R3) should be characterized by moderate intensity during the whole event. To assess the validity of



**Figure 9.** Predominantly designated precipitation cluster for each catchment for the flood type R1 (left), R2 (center), and R3 (right) considering six clusters. The symbol indicates the proportion of the precipitation cluster on the events of the respective flood type.

these assumptions, we determined the predominantly designated cluster, that is, the most frequent cluster of temporal rainfall distributions at each gauge of the data set for each flood type. Results are shown in Figure 9 and demonstrate that the connection of flood and temporal distribution type are spatially dependent.

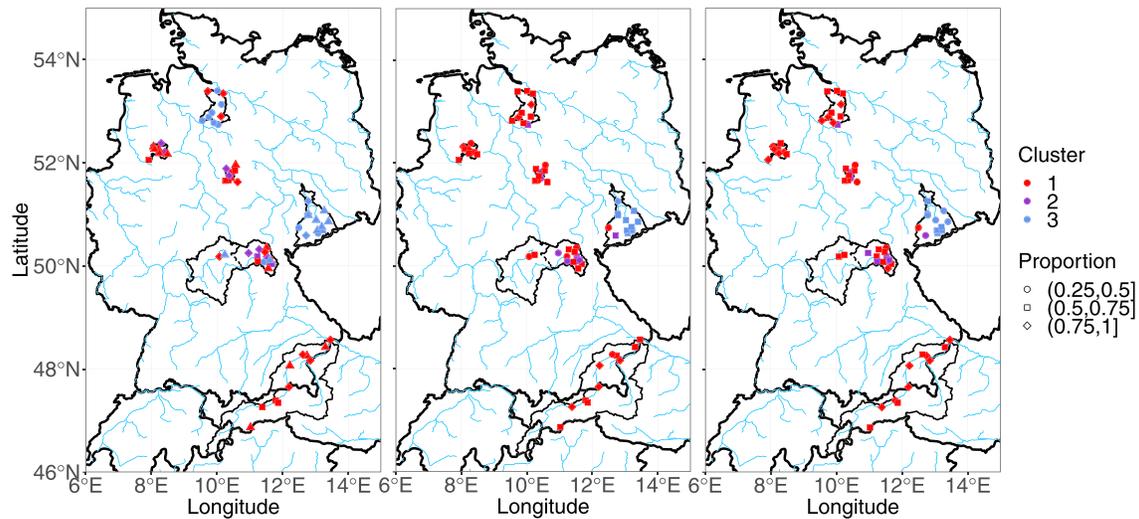
In Figure 9, distinct spatial differences can be observed for all flood types. The Mulde basin in the eastern part of Germany differed noticeably from the remaining basins. Whereas the short rain floods (R1) in the Mulde basin were mainly caused by precipitation with high intensity at the beginning (Clusters 1 and 2), for the Cloppenburg and Westharz basins Cluster 4 dominated this flood type. On the contrary, the Inn basin in the Alps varied in the dominating cluster for R1 flood events. While the upstream gauges in the Alps were dominated by Cluster 6 precipitation, the downstream gauges were more affected by Clusters 4 precipitation and Cluster 3 precipitation when closest to the outlet. For the northern Lueneburg basin, for most gauges, no dominating precipitation type could be identified for R1 events, although Cluster 1 has the largest proportion. For the Main basin in central Germany, also Cluster 4 is predominantly designated for R1 floods.

For flood events of type R2 (moderate peak and volume), in general a shift of the precipitation intensity peak can be observed (i.e., increasing cluster number). Especially the Inn basin showed a high dominance of Clusters 5 and 6, while the downstream gauges were predominantly designated as Cluster 4 precipitation events, and even the outlet gauge shifts from Cluster 3 to 4. The Main basin as well as the Lueneburg basins showed a homogeneous dominance of Cluster 4 precipitation events, similar to the results for flood type R1. The gauges in the Cloppenburg basin had the highest proportion of Cluster 4 and 6 precipitation events, whereas for R1 floods Cluster 1 had been predominantly designated. The Mulde basin, in contrast to the R1 flood events, was not dominated by one cluster. But a tendency toward Cluster 2 and 3 was visible. In conclusion, for the Inn and Lueneburg basins and most parts of the Mulde basin an increase of the predominantly designated cluster could be observed, where this was not the case for the Cloppenburg and Westharz basins and only for few catchments in the Main basin.

Another shift, that is, an increase in the predominantly designated cluster, in parts of the basins was observed for R3 flood events (small but often multiple peaks with large volumes), when compared to R2 events. For almost all basins, at least one catchment showed an increased cluster number. For the Inn basin, now mostly Cluster 6 could be observed. This is also the cluster that had an increased occurrence for many of the catchments in the Lueneburg, Cloppenburg, and Westharz basins. For the Mulde basin, still no dominant cluster could be observed, but the clusters tend more to Clusters 4 and 5.

The proportion of events belonging to this cluster has to be taken into account for the analysis of predominantly designated clusters of precipitation. For most gauges, a proportion of at least 25% was obtained. This can be seen as significant proportion, since the remaining 75% were distributed mostly equally across the remaining clusters and stayed below 20%. The predominantly designated clusters with a proportion of less than 25% cannot be seen as predominant, as the probability to obtain a different cluster is more or less equal. This is especially the case for many of the clusters designated for flood type R1 in the Main and Lueneburg basins. Here, potential shifts have to be taken with care since an almost equally likely cluster could lead to less significant results. However, to indicate a tendency, we included these in our analyses.

In summary, the results showed that the predominantly designated precipitation clusters and hence the temporal distributions of the event rainfall seemed to change with flood types. Generally, floods of short rainfall events (R1) were caused by early peaked precipitation events. With increasing event duration (R2), the emphasis of the precipitation event moved toward the center of the event duration. Floods caused by the longest rainfall events were caused by nearly equally distributed rainfall events or by late peaked rainfall events. Additionally, our analysis revealed spatial differences of designated precipitation clusters. The Mulde basin in the eastern part of Germany in general was much more affected by precipitation with high intensities at the beginning of the event than any other basin considered in this study. With its location in the Ore Mountains, the Mulde basin is dominated by continental climate, with the coldest average temperature in Germany. In the center and south of Germany, where the Inn, Main, Westharz, and Cloppenburg basins are located, a much more transitional climate dominates with oceanic impact. The northern Lueneburg catchment is characterized by oceanic climate. Again, this is the basin with another different cluster structure and no dominating cluster of temporal distribution. The spatial patterns of temporal rainfall distributions hence seem to be connected to climate, but this needs further systematic investigation.



**Figure 10.** Predominantly designated precipitation cluster for each catchment for the flood types R1 (left), R2 (center), and R3 (right), considering three parent clusters. The symbol indicates the proportion of the precipitation cluster on the events of the respective flood type.

To validate the results, the differences observed in the predominantly designated clusters were evaluated statistically. For this purpose, we applied the robust Wilcoxon test (Wilcoxon, 1945). The test evaluated the significance of the visible differences of predominantly designated clusters and flood types. We applied the test to the precipitation clusters of all flood event types pairwise. The comparison of clusters for flood types R1 and R2 showed a significant shift with a  $p$  value of 0.036 and hence lower than the significance level of 5%. Another shift from R2 to R3 also delivered a significant test result for a shift in the mean predominantly designated cluster with  $p = 0.047$ . As expected, this behavior could then also be observed for a shift from R1 to R3 floods with a  $p$  value of 0.0006. Again, the Mulde basin behaved differently from the remaining basins: R1 floods were mainly connected to clusters with high rainfall intensities at the beginning and hence small cluster numbers. This could lead to wrong results if the Wilcoxon test was applied to all basins. It could reduce the  $p$  value artificially. For this reason, we also applied the Wilcoxon test to all basins minus the Mulde basin. In this case, the comparison of clusters for flood type R1 and R2 did not show a significant shift. The  $p$  value of 0.38 was higher than the significance level of 5%. However, the Wilcoxon test applied to the clusters of flood type R1 and R3, delivered a  $p$  value of 0.029, and hence indicates a significant change in the clusters. An application to floods of type R2 and R3 delivered a  $p$  value of 0.056. The comparison of the clusters in the Mulde basin solely showed no significant shift. The  $p$  values for R1–R2 flood events was  $p = 0.646$ , while the  $p$  value for R1–R3 events was  $p = 0.166$ . As this was a single basin, only 12 data points (gauges) could be considered, resulting in a reduced efficiency of the test. This is also the reason why we did not apply the Wilcoxon test to each basin separately.

Overall, the visual results are supported, at least for the shift from R1 to R3 floods, where a significant increase of the predominantly designated cluster was confirmed by the Wilcoxon test. This means that a significant shift of the temporal distribution of rainfall between these flood types took place. The distributions shifted from higher intensities at the beginning to more uniformly distributed intensities and emphasis on the end for rainfall events with longer duration.

Analogous to the comparison above, the coherence of the three parent clusters, obtained with a  $\xi_1 = 0.99$  and  $\xi_2 = 1$  parametrization, with the flood types was analyzed. This tests if a lower number of clusters would still lead to significant coherences between temporal distribution of rainfall and flood types (Figure 10). Again, a significant shift of predominantly designated clusters between flood types R1 and R2 ( $p$  value = 0.009) as well as between R1 and R2 ( $p$  value = 0.02) was observed, while no significant change could be detected between R2 and R3. However, the information obtained on the spatial patterns is less detailed. For example, for the Inn basin the cluster of temporal distribution of rainfall is constant (Cluster 1) for all three flood types, while for many gauges of the Mulde basin Cluster 3 remains. It seems that the use of three clusters smoothens and only large differences between temporal distributions are recognized. However, it is clear

that both parameterizations led to clusters with significant correlation to flood types and hence meaningful results.

Finally, also the results of the BIRCH algorithm are compared between the three flood types. In the supporting information, Figure S1, it becomes clear that the differences between the predominantly cluster for temporal distribution of rainfall are significantly smaller than for the method proposed here. In fact, for some basins no change of cluster can be observed at all. Hence, significant differences are only obtained between flood types R2 and R3 are ( $p$  value = 0.004). The results are similar to those of the three parent clusters, though inverted.

## 5. Conclusions

In this study, temporal distributions of rainfall events were clustered using a new unsupervised learning algorithm, in order to assess the connection of clusters of temporal distribution of rainfall to flood types. While there are common approaches to represent the temporal distribution of rainfall, here we use dimensionless mass curves for the first time, leaving the number of clusters to be chosen unknown a priori. Hence, the first research question addressed in this study was how can temporal distributions of rainfall be grouped without a priori assumptions on the number of clusters?

Although several clustering techniques are available, most of them were not applicable to the problem stated. Either they required an iteration of possible cluster numbers, accompanied by a performance criterion, or were impacted by high parameter sensitivity, or they delivered inferior results. Additionally, common cluster techniques require certainty about the feature space because their results are bound to the position of the curves within the feature space. In case of temporal distributions, the feature space was on the one hand uncertain (number of sampling points), and on the other hand, the shape of the distributions, that is, the succession of the dimensions of the feature space rather than the feature space itself, was the point of interest for the definition of clusters. Therefore, a new method has been proposed in this study. The method was based on the similarity of curve shapes, that is, correlations between two similar events and their divergence from the entire data set. By increasing the level of abstraction, using similarity rather than the feature space for pattern recognition, the uncertainty of the feature space as well as the parameter sensitivity was lowered. A comparison to other clustering techniques showed that the new algorithm was less sensitive to parameterization than all other applied techniques. Moreover, most other clustering techniques failed when applied to temporal distributions of rainfall. Only the BIRCH algorithm showed comparable applicability to the problem given. A comparison of the clusters obtained showed that BIRCH focused on the position of the samples in the feature space, while our new algorithm focused on curve shape.

Within the parameter boundaries of the algorithm, we found three to 15 clusters of temporal distributions. This range of clusters has been proposed by several other studies (Dunkerley, 2012; Pelletier et al., 2009; Schulte et al., 2013; Wu et al., 2006). Yet we found that the clusters defined with our new algorithm were connected, meaning that the minimum number of three clusters could be considered as parent clusters, representing the three basic shapes of temporal distributions. Depending on the parameterization chosen, these parent clusters are divided into subclusters, offering a wider variety of distribution shapes. The analyses carried out to answer the second research question, that is, if clusters of temporal distributions are linked to flood types, showed that the chosen parameterization did not affect the meaningfulness of the delineated clusters. This result underlines the quality of our unsupervised learning method.

For the second research question, the Wilcoxon test was applied in order to compare the clusters to three rainfall-induced flood types. A significant correlation between flood type and cluster of temporal distribution of event rainfall was observed. The temporal distribution of rainfall shifted significantly from higher intensities at the beginning of the event to more uniformly distributed rainfall and even an emphasis on the end of the event, when the flood type changed from short rain flood to long-duration rain flood (often with multiple peaks). These results lead to emphasize that the temporal distribution of rainfall is one of the main processes that have an impact on the flood type. Of course, they are not the only explanatory factor, and other characteristics such as soil moisture and catchment characteristics have to be taken into account. Moreover, the resulting clusters for temporal distribution of rainfall have found to be spatially variable within Germany. Especially the behavior of event precipitation for the eastern Mulde basin with temporal distributions tending to high intensities of rainfall at the beginning of the events was striking, since it differed much from the behavior of the remaining basins. This seems to be connected to climate. It has to be noted that

these results have been obtained for long-lasting rainfall events with daily sums as input data. Especially small subbasins, with a time of concentration lower than the resolution of the used data, might offer different behavior than detected in this study. Hence, the differences we detected and the conclusion we drew were on the interbasin scale, rather than the intrabasin (i.e., subbasin scale). For further detailing within the basins, higher resolution data (at least hourly sums) are required. However, the outlined *performance* of the proposed clustering technique is not affected by this limitation. The use of dimensionless mass curves eliminated the impact of temporal resolution. Moreover, the application of timescales for flood typing is not affected by the resolution of the data since only events with duration of more than one day were considered. However, the low temporal resolution of discharge and precipitation data could be a reason, why no correlation between rainfall cluster and soil characteristics and topography was detected. This has to be investigated in future studies.

Future research should address the applicability of the unsupervised learning technique to other clustering problems. The impact of the dimensionality of the feature space, as well as the variance of the curve shapes on the clustering results needs to be evaluated. Moreover, the application to other problems such as hydrograph grouping or spatial distributions of rainfall events should be tested. So far, only a hydrograph-based typology was used. In the future, also typologies motivated by hydrology and hydroclimate will be compared to the temporal structure of rainfall.

**Acknowledgments**

The financial support of the Deutsche Forschungsgesellschaft (DFG) in terms of the research unit SPATE (FOR 2416) for Svenja Fischer is gratefully acknowledged. All data used in this work is freely available online. The discharge data for the Mulde basin are available by the LfULG Saxony at [www.umwelt.sachsen.de/umwelt/infosysteme/ida/](http://www.umwelt.sachsen.de/umwelt/infosysteme/ida/). For the basins Lueneburg, Cloppenburg, and Westharz, the discharge data are available by the NLWKN of Lower Saxony at [www.wasserdaten.niedersachsen.de/cadenza/](http://www.wasserdaten.niedersachsen.de/cadenza/). For the Main basin and the German parts of the Inn basin, the discharge data are available at [www.lfu.bayern.de/wasser/wasserstand&urlscore;abfluss/](http://www.lfu.bayern.de/wasser/wasserstand&urlscore;abfluss/), while the discharge data for the Austrian catchments of the Inn are available at [ehyd.gv.at/](http://ehyd.gv.at/). All precipitation data as well as temperature used here are available by the German Weather Service (DWD) at the climate data center (CDC; [cdc.dwd.de/portal/](http://cdc.dwd.de/portal/)). The separation as well as the flood typology applied here is published by Fischer et al. (2019), available at [doi.org/10.5281/zenodo.3738149](https://doi.org/10.5281/zenodo.3738149).

**References**

Alpaydin, E. (2016). *Machine learning: The new AI*, The MIT Press Essential Knowledge series. Cambridge: The MIT Press.

Ankerst, M., Breunig, M. M., Kriegel, H.-P., & Sander, J. (1999). Optics: Ordering points to identify the clustering structure. In *Acm Sigmod Record* (Vol. 28, pp. 49–60). New York, NY.

Ball, J. E. (1994). The influence of storm temporal patterns on catchment response. *Journal of Hydrology*, 158(3-4), 285–303.

Blöschl, G. (2006). Hydrologic synthesis: Across processes, places, and scales. *Water Resources Research*, 42, W03S02. <https://doi.org/10.1029/2005WR004319>

de Araújo, J. C., & González Piedra, J. I. (2009). Comparative hydrology: Analysis of a semiarid and a humid tropical watershed. *Hydrological Processes*, 23(8), 1169–1178.

Dolšak, D., Bezak, N., & Šraj, M. (2016). Temporal characteristics of rainfall events under three climate types in Slovenia. *Journal of Hydrology*, 541, 1395–1405.

Dunkerley, D. (2012). Effects of rainfall intensity fluctuations on infiltration and runoff: Rainfall simulation on dryland soils, Fowlers gap, Australia. *Hydrological Processes*, 26(15), 2211–2224.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd* (Vol. 96, pp. 226–231). Portland, Oregon.

Fischer, S., Schumann, A., & Bühler, P. (2019). Timescale-based flood typing to estimate temporal changing flood frequencies. *Hydrological Sciences Journal*, 64(16), 1867–1892.

Gelbard, R., Goldman, O., & Spiegler, I. (2007). Investigating diversity of clustering methods: An empirical comparison. *Data & Knowledge Engineering*, 63(1), 155–166.

Hall, M. J., & Minns, A. W. (1999). The classification of hydrologically homogeneous regions. *Hydrological Sciences Journal*, 44(5), 693–704.

Han, J., & Kamber, M. (2010). *Data mining: Concepts and techniques* (2nd ed.). The Morgan Kaufmann series in data management systems. Amsterdam: Elsevier/Morgan Kaufmann.

Hancer, E., & Karaboga, D. (2017). A comprehensive survey of traditional, merge-split and evolutionary approaches proposed for determination of cluster number. *Swarm and Evolutionary Computation*, 32, 49–67.

Hannah, D. M., Smith, B. P. G., Gurnell, A. M., & McGregor, G. R. (2000). An approach to hydrograph classification. *Hydrological Processes*, 14(2), 317–338.

Hattermann, F. F., Huang, S., Burghoff, O., Willems, W., Österle, H., Büchner, M., & Kundzewicz, Z. (2014). Modelling flood damages under climate change conditions—A case study for Germany. *Natural Hazards and Earth System Sciences*, 14(12), 3151–3168.

Hirschboeck, K. K., Ely, L., & Maddox, R. (2000). Hydroclimatology of meteorologic floods. In Wohl, E. E. (Ed.), *Inland flood hazards: Human, riparian and aquatic communities*. Cambridge: Cambridge University Press, pp. 39–72.

Houle, M. E., Kriegel, H.-P., Kröger, P., Schubert, E., & Zimek, A. (2010). Can shared-neighbor distances defeat the curse of dimensionality? In M. Gertz & B. Ludäscher (Eds.), *Scientific and statistical database management*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 482–500.

Huff, F. A. (1967). Time distribution of rainfall in heavy storms. *Water Resources Research*, 3(4), 1007–1019.

Keller, L., Rössler, O., Martius, O., & Weingartner, R. (2018). Delineation of flood generating processes and their hydrological response. *Hydrological Processes*, 32, 228–240.

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>

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.

Nied, M., Schröter, K., Lüdtke, S., Nguyen, V. D., & Merz, B. (2017). What are the hydro-meteorological controls on flood characteristics? *Journal of Hydrology*, 545, 310–326.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 2825–2830.

Pelletier, G., Anctil, F., & Filion, M. (2009). Characterization of 1-h rainfall temporal patterns using a Kohonen neural network: A Québec city case study. *Canadian Journal of Civil Engineering*, 36(6), 980–990.

- Pilgrim, D. H., & Cordery, I. (1975). Rainfall temporal patterns for design floods. *Journal of the Hydraulics Division*, *101*(1), 81–95.
- 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.
- Runkler, T. A. (2015). *Data mining*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Schulte, M., Oppel, H., & Schumann, A. (2013). Analyse und simulation der zeitlichen verteilung des niederschlages innerhalb von starkregenereignissen. *Korrespondenz Wasserwirtschaft*, *6*(5), 255–261.
- Seibert, J., & Vis, M. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*, 3315–3325.
- 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>
- Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., et al. (2019). Causative classification of river flood events. *WIREs Water*, *6*.
- Ternynck, C., Ben Alaya, M. A., Chebana, F., Dabo-Niang, S., & Ouarda, T. B. (2016). Streamflow hydrograph classification using functional data analysis. *Journal of Hydrometeorology*, *17*(1), 327–344.
- Vandenbergh, S., Verhoest, N. E. C., Buyse, E., & De Baets, B. (2010). A stochastic design rainfall generator based on copulas and mass curves. *Hydrology and Earth System Sciences*, *14*(12), 2429–2442.
- Wang, A., Qu, N., Chen, Y., Li, Q., & Gu, S. (2018). A 60-minute design rainstorm for the urban area of Yangpu District, Shanghai, China. *Water*, *10*(3), 312.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83.
- Wu, S.-J., Yang, J.-C., & Tung, Y.-K. (2006). Identification and stochastic generation of representative rainfall temporal patterns in Hong Kong territory. *Stochastic Environmental Research and Risk Assessment*, *20*(3), 171–183.
- Yin, S., Xie, Y., Nearing, M., Guo, W., & Zhu, Z. (2016). Intra-storm temporal patterns of rainfall in China using Huff curves. *Transactions of the ASABE*, *59*(6), 1619–1632.
- Zhang, T., Ramakrishnan, R., & Livny, M. (1996). BIRCH: An efficient data clustering method for very large databases. In *Acm Sigmod Record* (Vol. 25, ACM, pp. 103–114). New York, NY.