Environmental Research Letters

LETTER

The flood cooking book: ingredients and regional flavors of floods across Germany

Ralf Merz, Larisa Tarasova and Stefano Basso

Department Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale), Germany

E-mail: ralf.merz@ufz.de

Keywords: flood events, flood processes, soil moisture dynamics, random forests, natural hazard

Supplementary material for this article is available online

Abstract

River flooding is a major natural hazard worldwide, whose prediction is impaired by limited understanding of the interplay of processes triggering floods within large regions. In this study we use machine learning techniques such as decision trees and random forests to pinpoint spatio-temporal features of precipitation and catchment wetness states which led to floods among 177 267 rainfall-runoff events observed in 373 German river basins. In mountainous catchments with high annual precipitation rates and shallow soils, event rainfall characteristics primarily control flood occurrence, while wetness conditions and the spatial interplay between rainfall and catchment soil moisture drive flood occurrence even more than event rainfall volume in drier basins. The existence of a snow cover also enhances flood occurrence. The identified ingredients and regional flavors shed new light on the spatial dynamics of hydro-meteorological processes leading to floods and foster regional adaptation of flood management strategies and early warning systems.

1. Introduction

Every year floods cause severe damages to infrastructures and economies worldwide (UNISDR 2015), often leading to fatalities (Petrucci et al 2019). In fact, despite floods having been the focus of a substantial part of hydrological research in the past decades, engineers, public authorities and the private sector still have limited capabilities to predict this type of environmental hazard. In addition, flood risk is expected to increase globally as a result of economic growth, urbanization and climate change (Arnell and Gosling 2016, Winsemius et al 2016, Dottori et al 2018), making advances in this field crucial to support policy and decision makers in adopting effective measures to reduce the impacts of flooding.

The occurrence, duration, extent and severity of river floods are controlled by a variety of hydro-meteorological processes. Because of this heterogeneity, their generation mechanisms at the scale of whole river basins are not well understood yet (Hall et al 2014). As a consequence, these mechanisms are usually ignored in statistical (Merz and Blöschl 2008) and comparative analyses (Salinas et al 2013), resulting in uncertain predictions of flood characteristics and their possible changes (Kundzewicz et al 2014). Progresses in understanding and accounting for physical processes leading to floods are key to improve their prediction and interpret their transformations (Merz et al 2014, Blöschl et al 2017, Berghuijs et al 2019, Barth et al 2019), especially under environmental change.

Existing studies on flood processes comprise examinations of single extreme events (e.g. Marchi et al 2010, Grams et al 2014, Martius et al 2018, Bronstert et al 2018), analyses of meteorological drivers of flooding (e.g. Smith et al 2011, Delgado et al 2012, Hofstäder et al 2016, Schlef et al 2019), and studies on the role of antecedent catchment states (Vormoor et al 2016, Bennett et al 2018, Wasko et al 2020). Works using seasonality as a surrogate of flood-triggerng processes (Villarini 2016, Berghuijs et al 2017, Blöschl et al 2017) also exist, as well as investigations which consider flood generating mechanisms to estimate statistics (Struthers and Sivapalan 2007, Merz and Blöschl 2008, Basso et al 2016), extent (Kemter et al 2020) and trends (Mangini et al 2018, Blöschl et al 2019, Neri et al 2019) of floods.

In recent years, process-based classifications of flood events have gained momentum (e.g. Merz
2. Materials and methods

2.1. Data

The study uses observed daily streamflow data from 373 German catchments ranging from 30 to 23,719 km² (median = 463 km²), daily gridded rainfall data (on a 1 km grid) from the REGNIE (Regionalisiert Niederschlagshöhen) data set (Rauthe et al. 2013) provided by the German Weather Service, air temperatures interpolated using external drift kriging (Zink et al. 2017), and snowmelt and soil moisture (SM) time series (on a 4 km grid) simulated by the mesoscale hydrologic model (mHM) (Zink et al. 2017).

Catchments with streamflow time series significantly impacted by reservoirs or river engineering works are not included in this study (Tarasova et al. 2018a). Large basins with a significant fraction of their areas located outside Germany and where river routing processes might affect flood characteristics are also excluded from this analysis, as well as lowland catchments in north-eastern Germany, where runoff behavior is strongly controlled by lakes and groundwater dynamics.

A map of Germany showing the spatial distribution of the study catchments along with key climate and landscape characteristics is given in figure S1 (available online at http://stacks.iop.org/ERL/15/114024/mmedia). A detailed description of the study area is available in (Tarasova et al. 2018a).

Various catchment descriptors representative of regional variations in landscape (e.g. topography, geomorphology, land use, soil characteristics, hydro-geological information) and climate (e.g. mean annual precipitation) characteristics (table S1) are used to illustrate general climatological and landscape attributes corresponding to spatial patterns of flood ingredients.

2.2. Flood definition and hydro-meteorological indicators

Goal of this study is to identify specific spatio-temporal interactions between precipitation and wetness state leading to floods in river catchments. In a first step, single rainfall-runoff events are derived from continuous hydroclimatic time series using the procedure illustrated in (Tarasova et al. 2018b). In essence, runoff time series are separated into quick and base flow by means of a simple smoothing algorithm (Institute of Hydrology 1980) and event precipitation is assigned to each separated event by a lag time approach (Mei and Anagnostou 2015). Finally, multi-peak events and overlapping single-peak events are iteratively distinguished by comparing events statistics (Tarasova et al. 2018b). The approach results in a collection of 177,267 runoff events occurred between 1979 and 2002 in 373 German catchments.

In this study, we assume that at least one flood event per year occurs in every catchment. Hence, we identify the highest streamflow of each year to build the annual maximum flood series (AMS) of the catchment and tag each streamflow peak larger than the minimum value of the AMS as flood event. This results in 35,565 events (i.e. ~20% of all runoff events) labeled as flood events.

For each rainfall-runoff event several indicators which capture the spatio-temporal dynamics of event precipitation, catchment wetness state and their interactions are calculated (table 1). Note that indicators directly derived from the streamflow time series, such as event runoff volume or event runoff coefficients are not used as potential flood ingredients.

Some indicators are spatio-temporal averages, such as event rainfall volume divided by mean annual precipitation, (R.vol) and antecedent soil moisture relative to its climatology (Zink et al. 2017). We also consider the portion of catchment that is wet (SM.wet) (Tarasova et al. 2020) to account for runoff generating areas.

However, Seo et al. (2012) and Mei et al. (2014) stated that information on the space-time interaction of hydro-meteorological processes is needed to correctly mimic magnitude and timing of flood events. Therefore, unlike many previous studies on regional flood processes, we additionally use spatially and temporally distributed event characteristics to better capture the spatio-temporal interactions of rainfall and
soil moisture within the catchment (Tarasova et al. 2020). For example, we derive the overlap between precipitation and wetter areas within the catchment (PSM.overlap.mean), to distinguish between rain falling on wetter soils, which tends to produce fast runoff (Tarasova et al. 2018a) and rain falling on drier soils, which tends to be stored in the catchment. This is known to be a key driver of runoff generation (McGlynn et al. 2003). Similarly, we use the spatial covariance of precipitation and snow cover (PSWE.cov) to account for rain on snow events, well reported in the literature as drivers of large scale floods (Musselman et al. 2018).

2.3. Predicting flood occurrence and disentangling flood ingredients

In this study we use machine learning and data mining tools, which have been recently applied to predict the occurrence of flood/non-flood events (Schmidt et al. 2020) in alternative to deterministic rainfall-runoff models (Vormoor et al. 2016, Musselman et al. 2018). Using machine learning allows for consistently analysing the drivers of runoff generation for a large number of events in different climate and landscape settings. Specifically, we use RF and DT approaches. The basic idea of tree-based data mining (Breiman 2001) is to recursively split the data space (in our case event characteristics) into sub-spaces according to the behavior of a response variable (i.e. flood/non-flood labels). The result is a DT where at each node events are divided into two branches according to wether a selected event characteristic is smaller or larger than a threshold. At the end of each branch, events are classified as ‘flood events’ or ‘non-flood events’. RFs (Cutler et al. 2007) are a variant of tree-based models consisting of a large number of DTs that operate as an ensemble. RFs apply bootstrap sampling to compose the training subsamples of the single trees. The average of all the individual DT estimates (or the majority vote in case of categorical classes) is finally taken as prediction. Due to combining many DTs into a single model, predictions tend to be more accurate and less sensitive to the specific training data. Whereas DTs are easily interpretable, because classification rules are given, the classification strategy of RFs is often seen as a black box (Schmidt et al. 2020).

A flowchart of the study is given in figure S2.

To identify regional flood ingredients we first train a RF for each catchment using the potential flood ingredients of table 1 as predictor variables. 60% of all events in the catchment are randomly selected to build up the RF (training sample), while the remaining 40% of them (test sample) are used to validate the prediction performance. The prediction performance is here measured in terms of the Matthews correlation coefficient (MCC) (equation S3) (Lever et al. 2016). MCC ranges from 1 (when the classification is always wrong) to 0 (when it is no better than random) to 1 (when it is always correct). Although no single metric can summarize all possible strengths and weaknesses of a classifier, MCC has proven to outperform other classification metrics in highly imbalanced data (Giuseppe, Boughorbel 2017, Chicco and Jurman 2020). This feature motivates our choice of MCC as classification metric. The RFs thus obtained embody relevant flood ingredients for each catchment.

| Indicator       | Description                                                                 |
|-----------------|------------------------------------------------------------------------------|
| R.vol           | volume of catchment-averaged event rainfall/mean annual precipitation [-]    |
| M.vol           | melt volume of [mm]/mean annual precipitation [-]                           |
| Pmaxl           | maximum daily rate (i.e. intensity) of catchment-averaged event precipitation/50% quantile of event rainfall volume [-] |
| Pts             | ratio of maximum daily precipitation rate and total precipitation volume [-]  |
| Ptcv            | temporal coefficient of variation of daily precipitation rates [-]           |
| Pextent         | portion of catchment affected by precipitation event [-]                    |
| Pspcv           | spatial coefficient of variation of precipitation volumes [-]               |
| SWE             | catchment-averaged antecedent snow water equivalent/mean annual precipitation [-] |
| SWE.extent.mean | mean (in time) extent of snow covered areas during precipitation event [-]   |
| SWE.extent.cv   | temporal coefficient of variation of extent of snow covered areas during precipitation event [-] |
| PSWE.cov        | covariance (in space) of event-averaged snow water equivalent and rainfall event volume/(mean annual precipitation)^2 [-] |
| SM              | antecedent catchment-averaged soil moisture [-]                             |
| SM.spcv         | spatial coefficient of variation of antecedent soil moisture [-]            |
| SM.wet          | portion of catchment that is attributed to the wet antecedent soil moisture state [-] |
| Psteadiness.mean| mean of spatial covariance between consecutive days within precipitation event [-] |
| Psteadiness.cv  | temporal coefficient of variation of spatial covariance between consecutive days within precipitation event [-] |
| PSM.overlap.mean| mean (in time) portion of overlap between areas attributed to wet antecedent soil moisture state and precipitation during event [-] |
| PSM.overlap.cv  | temporal coefficient of variation of overlap between areas attributed to wet antecedent soil moisture state and precipitation during event [-] |
To identify regional differences in the ingredients of flood events, catchments are then pooled into homogeneous groups. To do so, we assume that two catchments have similar flood ingredients if the RF locally derived in one catchment can reliably predict flood/non-flood occurrence in the other catchment. Hence, the RF of each catchment is used to predict flood/non-flood occurrence in all other catchments and MCC is used to measure the predictive performance. Hierarchical clustering techniques (Gordon 1987) which maximize pairwise MCC are used to derive clusters with similar flood ingredients. By compiling all training events within a region together, we finally build RFs to predict flood/non-flood occurrence in each cluster. The number of clusters is selected to maximize the median MCC for predicting flood/non-flood occurrence in the test sample of all catchments by using regional RFs. To get a hint on the most important factors distinguishing flood from non-flood events the distribution of the feature (or variable) importance (measured as mean decrease of the Gini index) (Breimann, 2001) is used. The feature importance score rates how important each variable is for building up the decision forest.

As the feature importance scores of RFs only provide the relative importance between flood ingredients (but no indications about how they combine), we also fit DTs to all trainings events in each cluster. In fact, although DTs tend to have a weaker predictive performance than RFs, the resulting trees explicitly represent the classification rules, which can then be hyrologically interpreted. As the objective of this study is not the optimization of prediction performances, but a deeper comprehension of which spatio-temporal combinations of rainfall and soil moisture lead to flooding, we use RFs and DTs in parallel to take advantage of their individual strengths.

3. Results and discussion

Figure 1(a) displays the predictive performance of individual RFs for each catchment in terms of MCC. The average MCC for all basins is 0.70, while its upper and lower quartiles are 0.63 and 0.75. Figure 1(b) shows the clustering of catchments based on the similarity of their flood ingredients, together with the predictive performance of the regional RFs (i.e. one RF per cluster). The predictive performance obtained by using regional RFs to predict flood/non-flood occurrence in each catchment is on average larger than by using local RFs. In fact, the mean MCC for all catchments increases from 0.70 to 0.77 when using regional RFs instead of local ones. Similarly, the lower and upper quartiles increase from 0.63 to 0.70 and from 0.75 to 0.82. This counterintuitive result is explained by the existence of two opposing factors affecting predictive performance when moving from local to regional RFs. As flood occurrence certainly depends on the local hydrological conditions, which can be better represented by local than regional RFs, one would expect a decrease in predictive performance when regional RFs are used. However, floods are by definition rare events and hence the sample size of local flood events always tends to be small. 

![Figure 1](image-url)

Figure 1. (a) Matthews correlation coefficient (MCC) (shade of the dots) between observed and predicted occurrence of flood/non-flood events for the test sample in each basin, obtained by using individual Random Forests for each catchment. Dots represent the basins' outlets; (b) Clustering of regions with similar flood ingredients (colored clusters) and MCC (shade of the dots) between observed and predicted occurrence of flood/non-flood events for the test sample in each basin, obtained by using one Random Forest per cluster (i.e. using the regional RFs).
Although pooling catchments together may neglect local factors impacting flood characteristics, the resulting sample size is larger and more information on flood characteristics is available. Hence, assembling all events within a region enables to better train RFs and thus enhance prediction performance. Our results indicate that the second factor is dominating. These findings are in line with the general idea of pooling catchments to better estimate local flood characteristics (e.g. Salinas et al 2013).

The resulting clusters (figure 1(b)) of similar event characteristics responsible for the occurrence of floods mostly match the spatial distribution of key climate and landscape descriptors (CDs) in Germany. Figure 2(a) shows boxplots of selected normalized CDs for catchments in each cluster. These descriptors are solely used in the study to illustrate climatic and landscape attributes of the clusters identified from event characteristics. The importance of the single event characteristics for the occurrence of floods in each cluster is displayed in figure 2(b).

Cluster 1 mainly comprises mountainous catchments in the pre-alpine region and the higher parts of the mid-elevation mountain ranges of the Black and Bavarian Forests and the Rheinish Massif, which face the highest mean annual rainfall rates of about 1130 mm yr\(^{-1}\). These catchments are widely forested and characterized by shallow soil depths, small topographic wetness indices (which indicate catchments with only a few flat areas near the streams) and small groundwater yield. Cluster 5, on the contrary, includes catchments in the transition zone between
central Germany and the northern lowlands. Catchments in this cluster tend to be significantly drier with mean annual rainfall rates of 660 mm yr$^{-1}$. The catchments are flatter than in cluster 1, with deeper soils and larger groundwater yields. Generally, aridity index (ratio between long-term potential evapotranspiration and rainfall), topographic wetness index, soil depth and groundwater yield increase from cluster 1 to 5, while mean annual precipitation, drainage density, catchment altitude, amount of forested areas and slope tend to decrease from cluster 1 to 5.

Figure 2(b) displays the normalized importance of features used for splitting in the regional RF of each cluster. The figure reveals that rainfall volume (R.vol) is the most important feature to distinguish between floods and non-flood events in clusters 1 to 3. In the wettest cluster 1 the maximum intensity of event rainfall (P.maxI) is the second most important feature, which even precedes soil moisture. This may be explained by high annual rainfall and low evaporation rates, leading to relatively wet catchments with high chances of event-fed saturation which exhibit runoff peaks largely dependent on event rainfall characteristics (Tarasova et al.2018a).

In the driest clusters 4 and 5 wet antecedent SM states and spatial overlapping of precipitation and soil moisture (PSM.overlap) tend to be more important than rainfall volume. Event runoff coefficients are usually low in dry catchments, but can significantly increase if soils are wet (Merz et al.2009, Tarasova et al.2018a). Hence, overlapping of rainfall and wet soils is a key ingredient for the generation of floods in these catchments. The emerging importance of PSM.overlap for the occurrence of flood/non-flood events underpins findings of previous studies conducted at plot and hillslope scales (Mcgley and Mcdonnell 2003, Freyberg et al.2014), which show that event runoff is mainly produced in limited areas of the catchment. Our results also suggest that the other event characteristics have minor importance for flood occurrence.

To shed light on the composition of flood recipes in each region, we fit DTs to all training events in a region and analyzed the classification rules provided. As expected, the overall predictive performance in terms of mean MCC decreases to 0.70 by using regional DTs compared to regional RFs. This value is comparable to the performance of local RFs. Figure 3 displays the DTs for predicting flood/non-flood occurrence in each cluster together with sketches of archetypical catchments, where the main regional climate and landscape descriptors of each cluster are visualized to support transferring the findings of this study to other regions in the world. The flood ingredients unveiled by the DT of each cluster agree with those highlighted by the feature importance of the RF classification. Event rainfall volume (R.vol) is the first splitting criterion in clusters 1 to 3, showing the importance of event rainfall for flood occurrence. While in the wettest cluster 1 event rainfall volume (R.vol) and maximum precipitation intensity (P.maxI) are sufficient to label 70% of all events as ‘non-floods’, information on catchment wetness state always appears as splitting criterion in the two top branches in clusters 2 and 3. In these clusters floods only occur as a result of the interaction of rainfall volume and catchments wetness state.

In line with the feature importance of the RF approach, catchment-averaged soil moisture and the spatial overlap of precipitation and wet soil (PSM.overlap) tend to play a more important role than event rainfall volume for the generation of floods in clusters 4 and 5, which comprise drier catchments. PSM.overlap and SM are respectively used in these clusters as the top branch splitting criterion.

In clusters 1 to 4 also melt runoff (M.vol) and snow water equivalent (SWE) appear as splitting factor. The rather low threshold values for melt runoff in these clusters suggest that even if it does not substantially contribute to runoff volume, the existence of a snow cover is relevant because it might give rise to the occurrence of rain-on-snow events. This finding further stresses the importance of this type of flood as suggested by previous studies (McGabe et al.2007, Freudiger et al.2014, Vormoor et al.2016, Berghuijs et al.2016, Krug et al.2020).

4. Summary and conclusions

This study aims at deciphering spatio-temporal interactions between precipitation and wetness states of catchments that trigger floods. By using machine learning approaches to facilitate the analysis of 177 267 rainfall-runoff events, we consistently identify event characteristics which lead to flood runoff in a large number of catchments. Although tree based approaches used in this study do not necessarily disentangle causal relationships between precipitation and catchment wetness states and the occurrence of flood events, the resulting trees reveal a regional variability of flood ingredients that can be interpreted in the context of physiographic and climatic differences among catchments.

Our results show that regional RFs derived for clusters of catchments which are grouped together according to the similarity of their flood ingredients improve prediction of flood/non-flood occurrences compared to individual RFs derived for each catchment. Floods are intrinsically rare events and hence local flood samples might suffer from their limited sizes. The increase of information on flood generating processes resulting from the use of larger samples obtained by pooling basins together prevails over the benefit of having local RFs which better represent local flood characteristics. The clustering of catchments with similar flood ingredients show clear spatial patterns, which follow general climate
Figure 3. Archetypical catchment structure and regional decision tree to predict flood occurrence for each cluster with similar flood ingredients. Numbers in end leaves of decision trees correspond to percentage of all events falling in each end leave. Flood ingredients used in regional decision trees are the ratios between (i) volumes of catchment-averaged event rainfall and mean annual precipitation (R.vol), (ii) melt volume and mean annual precipitation volume (M.vol), and (iii) maximum daily rate (i.e. intensity) of catchment-averaged event precipitation and median of event rainfall volume (P.maxI), (iv) the mean of spatial covariance between consecutive days within a precipitation event (P.steadiness.mean), (v) the ratio between catchment-averaged antecedent snow water equivalent (SWE) and mean annual precipitation, (vi) antecedent catchment-averaged soil moisture (SM) and (vii) mean (in time) portion of overlap between areas with wet antecedent soil moisture state and precipitation during event (PSM.overlap.mean). Typical catchment properties are presented based on selected catchment descriptors (table S1).
and landscape characteristics. In mountainous catchments facing large rainfall rates and wet soils all over the year, the amount of event rainfall primarily determines flood occurrence, while catchment wetness states seem to be less important for distinguishing flood from non-flood events. However, for the drier regions the interplay of rainfall with catchment soil moisture gains importance. In drier catchments small rainfall events falling on wet soils can also result in flood runoff. In the second driest region of our study, the overlap of rainfall and wet soils tends to be more important than the event rainfall volume itself. In the driest catchments, wetness conditions tend to assume the driving role. The existence of a snow cover seems instead to enhance flood occurrence.

The identified regional flavors of flood ingredients may help to explain the spatially variable performance of flood forecasting models and flood monitoring systems, eventually strengthening their reliability. In fact, recognizing catchments where the interplay between precipitation and antecedent wetness dominates flood occurrence may indicate regions where enhanced accuracy in the representation of soil moisture dynamics and spatial rainfall patterns is required to establish effective flood hazard early warning systems. Future work will focus on investigating relevant ingredients for the occurrence of ordinary versus extreme floods and on modifications of flood recipes over time due to land use and climate change.

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 research project Propensity of rivers to extreme floods: climate-landscape controls and early detection (PREDICTED, Grant Number 421396820), and of the Helmholtz Centre for Environmental Research—UFZ is gratefully acknowledged. For providing the discharge data for Germany, we acknowledge the Factory Inspectorate (LUWG), Office of 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 (Big, 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/).

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Ralf Merz https://orcid.org/0000-0002-3698-2339
Larisa Tarasova https://orcid.org/0000-0002-7675-0731
Stefano Basso https://orcid.org/0000-0003-3551-9155

References

Arnell N W and Gosling S N 2016 The impacts of climate change on river flood risk at the global scale Clim. Change 134 387–401
Barth N A, Villarini G and White K 2019 Accounting for mixed populations in flood frequency analysis: a bulletin 17C perspective J. Hydrol. Eng. 24 1–12
Basso S, Schirmer M and Botter G 2016 A physically based analytical model of flood frequency curves Geophys. Res. Lett. 43 9070–6
Bennett B, Leonard M, Deng Y and Westra S 2018 An empirical investigation into the effect of antecedent precipitation on flood volume J. Hydrol. 567 435–45
Berghuijs W R, Harrigan S, Molnar P, Slater L J and Kirchner J W 2019 The relative importance of different flood-generating mechanisms across Europe Water Resour. Res. 55 4582–93
Berghuijs W R, Woods R A, Hutton C J and Sivapalan M 2016 Dominant flood generating mechanisms across the United States Geophys. Res. Lett. 43 4382–90
Berghuijs W, Aalbers E, Larsen J, Trancoso R and Woods R 2017 Recent changes in extreme floods across multiple continents Environ. Res. Lett. 12 114035
Blöschl G, Hall J, Parajka J, Perdiguero A P, Merz B, Arheimer B and Chirico G B 2017 Changing climate shifts of European floods Science 357 388–90
Blöschl G, Hall J, Viglione A, Perdiguero A P, Parajka J, Merz B and Živković N 2019 Changing climate both increases and decreases European river floods Nature 573 108–11
Boughorbel S B 2017 Optimal classifier for imbalanced data using Matthews correlation coefficient metric PLOS One 12 e0117767
Breiman L 2001 Random Forests Mach. Learn. 45 5–32
Bronstert A et al 2018 Forensic hydro-meteorological analysis of an extreme flash flood: the 2016–05–29 event in Braunsbach, SW Germany Sci. Total Environ. 630 977–99
Chicco D and Jurnag G 2020 The advantages of the Matthew’s correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation BMC Genomics 21
Cutler D R, Edwards T C Jr., Beard K H, Cutler, A, Hess K T, Gibson J and Lawler J 2007 Random forests for classification in ecology Ecology 88 2783–92
Delgado J M, Merz B and Apel H 2012 A climate–flood link for the lower mekong river Hydro. Earth Syst. Sci. 16 1533–41
Dottori F et al 2018 Increased human and economic losses from river flooding with anthropogenic warming Nat. Clim. Change 8 781–6
Freudiger D, Kohn I, Stahl K and Weiler M 2014 Large-scale analysis of changing frequencies of rain-on-snow events with flood-generation potential Hydro. Earth Syst. Sci. 18 2695–709
Freyberg J, Radny D, Gall H E and Schirmer M 2014 Implications of hydrologic connectivity between hillslopes and riparian zones on streamflow composition J. Contamin. Hydro. 69 62–74
Giuseppe J 2012 A comparison of MCC and CEN error measures in multi-class prediction PloS One 7 e41882
Gordon A D 1987 A review of hierarchical classification J. R. Stat. Soc. Ser. A 150 119–57
Grans C M, Binder H, Pfahl S, Piaget N and Wernli H 2014 Atmospheric processes triggering the central European floods in June 2013 Nat. Hazards Earth Syst. Sci. 14 1691–702
Hall J et al 2014 Understanding flood regime changes in Europe: a state-of-the-art assessment Hydro. Earth Syst. Sci. 18 2735–72
Hofstätter M, Chimani B, Lexer A and Blöschl G 2016 A new classification scheme of European cyclone tracks with relevance to precipitation Water Resour. Res. 52 7086–104
Institute of Hydrology 1980 Low flow studies (Rep 1) (Wallingford, UK: Institute of Hydrology)
Keller L, Rössler O, Martius O and Weingartner R 2017 Delineation of flood generating processes and their hydrological response Hydro. Process. 32 228–40
Kemter M, Merz B, Marwan N, Vorogushyn S and Blöschl G 2020 Joint trends in flood magnitudes and spatial extents across Europe Geophys. Res. Lett. 46 e2020GL087464
Krug A, Primo C, Fischer S, Schumann A and Ahrens B 2020 On the temporal variability of widespread rain-on-snow floods Meteorol. Z. 29 1–13
Kundzewicz Z W et al 2014 Flood risk and climate change: global and regional perspectives Hydrol. Sci. J. 59 1–28
Lever J, Krzywinski M and Altman N 2016 Classification evaluation Nat. Methods 13 541–2
Mangini W et al 2018 Detection of trends in magnitude and frequency of flood peaks across Europe Euro. Hydrol. Sci. J. 63 493–512
Marchi L, Borga M, Precioso E and Gaume E 2010 Characterisation of selected extreme floods in Europeand implications for flood risk management J. Hydrol. 394 118–33
Martius O et al 2018 Reconstruction and simulation of an extreme flood event in the Lago Maggiore catchment in 1868 Nat. Hazards Earth Syst. Sci. 18 2717–39
McCabe G J, Clark M P and Hay L E 2007 Rain-on-snow events in the western United States Bull. Am. Meteorol. Soc. 88 319–28
Mcginn L B and Mcdonnell J J 2003 Quantifying the relative contributions of riparian and hillslope zones to catchment runoff Water Resour. Res. 39 1310
Mey I and Anagnostou E N 2015 A hydrograph separation method based on information from rainfall and runoff records J. Hydrol. 523 636–49
Mey I, Anagnostou E N, Stampoulou D, Nikolopoulos E I, Borga M and Vegara H J 2014 Rainfall organization control the flood response of mid-slope hillslopes J. Hydrol. 510 565–77
Merz B J et al 2014 Floods and climate: emerging perspectives for flood risk assessment and management NHESS 14 1921–42
Merz R and Blöschl G 2008 Flood frequency hydrology 2: combining data evidence Water Resour. Res. 44 W08432
Merz R and Blöschl G 2009 A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria Water Resour. Res. 45 W01405
Merz R and Blöschl G 2003 A process typology of regional floods Water Resour. Res. 39 1–20
Musselman K N, Lehner F, Ikeda K, Clark M P, Prein A F, Liu C, Barlage M and Rasmussen R 2018 Projected increases and shifts in rain-on-snow flood risk over western North America Nat. Clim. Change 8 808–12
Neri A, Villarini G, Slater L J and Napolitano F 2019 On the statistical attribution of the frequency of flood events across the U.S. Midwest Adv. Water Resour. 127 225–36
Nied M, Pardowitz T, Nissen K, Ulbrich U, HucP e Y and Merz B 2014 On the relationship between hydro–meteorological patterns and flood types Jour. Hydro. 519 3249–5262
Petrusci et al 2015 Flood fatalities in europe, 1980–2018: variability, features, and lessons to learn Water 11 1682
Rathie M, Steiner H, Riediger U, Mazurkiewicz A and Gratzki A 2013 A central European precipitation climatology—part I: generation and validation of a high-resolution gridded daily data set (HYRAS) Meteorol. Z. 22
Salinas J L, Laaha G, Rogger M, Parajka J, Viglione A, Sapinvalan M and Blöschl G 2013 Comparative assessment of predictions in ungauged basins; part 2: flood and low flow studies Water Resour. Res. 49 12637–52
Schel K E, Moradkhani H and Lal U 2019 Atmospheric circulation patterns associated with extreme united states floods identified via machine learning Sci. Rep. 9 7171
Schmidt L, Hefe F, Attinger S and Kumar R 2020 Challenges in applying machine learning models for hydrological inference: a case study for flooding events across Germany Water Resour. Res.
Seo Y, Schmidt B A and Sapinvalan M 2012 Effect of storm movement on flood peaks: analysis framework based on characteristic timescales Water Resour. Res. 48 1–12
Sharma A, Wasko C and Lettenmaier D P 2018 If precipitation extremes are increasing, why aren’t floods? Water Resour. Res. 54 8354–51
Shen C 2018 A transdisciplinary review of deep learning research and its relevance for water resources scientists Water Resour. Res. 54 8558–93
Sikorska A E, Vivirolí D and Seilbert J 2015 Flood-type classification in mountainous catchments using crisp and fuzzy decision trees Water Resour. Res. 51 7995–79
Smith J A, Villarini G and Baecck M L 2011 Mixture distributions and the climatology of extreme rainfall and flooding in the eastern US J. Hydrometeorol. 12 294–309
Stein L, Pianosi F and Woods R 2019 Event-based classification for global study of river flood generating processes Hydro. Process. 34 1514–29
Struthers I and Sivapalan M 2007 A conceptual investigation of process controls upon flood frequency: role of thresholds Hydro. Earth Syst. Sci. 11 1405–16
Tarasova L et al 2019 Causative classification of river flood events WIREs Water Res. 6 e1353
Tarasova L, Basso S, Wendi D, Viglione A, Kumar R and Merz R 2020 A process-based framework to characterize and classify runoff events – the event typology of German Water Resour. Res. 56
Tarasova L, Basso S, Poncelet C and Merz R 2018a Exploring controls on rainfall-runoff events: 2. regional patterns and spatial controls of event characteristics in Germany Water Resour. Res. 54 7688–7710
Tarasova L, Basso S, Zink M and Merz R 2018b Exploring controls on rainfall-runoff events: 1. time series-based event separation and temporal dynamics of event runoff response in Germany Water Resour. Res. 54 7711-7732
UNISDR 2015 Making development sustainable: the future of disaster risk reduction. centre for research on the...
epidemiology of disasters United Nations Office for Disaster Risk Reduction (UNISDR) Geneva, Switzerland

Villarini G 2016 On the seasonality of flooding across the continental United States Adv. Water Resour. 87 80–91

Vormoor K, Lawrence D, Schlichting L, Wilson D and Wong W K 2016 Evidence for changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway J. Hydrol. 538 33–48

Wasko C, Nathan R and Peel M C 2020 Changes in antecedent soil moisture modulate flood seasonality in a changing climate Water Resour. Res. 56 e2019WR026300

Winsemius H et al 2016 Global drivers of future river flood risk Nat. Clim. Change 6 381–5

Zink M, Kumar R, Cuntz M and Samaniego I 2017 A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty Hydrol. Earth Syst. Sci. 21 1769–90