ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE

10.1029/2020WR028300

Key Points:

- Flood generating processes are mostly influenced by climate attributes
- Which attributes are influential varies between different flood processes and climate regions
- Mix of flood generating processes can be predicted for ungauged catchments with high accuracy

Supporting Information:

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

Correspondence to:

L. Stein,
lina.stein@bristol.ac.uk

Citation:

Stein, L., Clark, M. P., Knoben, W. J. M., Pianosi, F., & Woods, R. A. (2021). How do climate and catchment attributes influence flood generating processes? A large-sample study for 671 catchments across the contiguous USA. *Water Resources Research*, *57*, e2020WR028300. https://doi.org/10.1029/2020WR028300

Received 14 JUL 2020
Accepted 4 FEB 2021

How Do Climate and Catchment Attributes Influence Flood Generating Processes? A Large-Sample Study for 671 Catchments Across the Contiguous USA

L. Stein[1], M. P. Clark[2], W. J. M. Knoben[2], F. Pianosi[1], and R. A. Woods[1]

[1]Department of Civil Engineering, University of Bristol, Bristol, UK, [2]Coldwater Laboratory, University of Saskatchewan Centre for Hydrology, Canmore, AB, Canada

Abstract Hydrometeorological flood generating processes (excess rain, short rain, long rain, snowmelt, and rain-on-snow) underpin our understanding of flood behavior. Knowledge about flood generating processes improves hydrological models, flood frequency analysis, estimation of climate change impact on floods, etc. Yet, not much is known about how climate and catchment attributes influence the spatial distribution of flood generating processes. This study aims to offer a comprehensive and structured approach to close this knowledge gap. We employ a large sample approach (671 catchments across the contiguous United States) and evaluate how catchment attributes and climate attributes influence the distribution of flood processes. We use two complementary approaches: A statistics-based approach which compares attribute frequency distributions of different flood processes; and a random forest model in combination with an interpretable machine learning approach (accumulated local effects [ALE]). The ALE method has not been used often in hydrology, and it overcomes a significant obstacle in many statistical methods, the confounding effect of correlated catchment attributes. As expected, we find climate attributes (fraction of snow, aridity, precipitation seasonality, and mean precipitation) to be most influential on flood process distribution. However, the influence of catchment attributes varies both with flood generating process and climate type. We also find flood processes can be predicted for ungauged catchments with relatively high accuracy (R^2 between 0.45 and 0.9). The implication of these findings is flood processes should be considered for future climate change impact studies, as the effect of changes in climate on flood characteristics varies between flood processes.

1. Introduction

Flood processes influence flood behavior (Fischer et al., 2016; Gaál et al., 2012; Keller et al., 2018; Merz & Blöschl, 2005; Tarasova et al., 2019). Thus, the need to classify these processes has long been recognized and several studies have developed flood classification approaches (e.g., Berghuijs et al., 2016, 2019; Blöschl et al., 2017; Diezig & Weingartner, 2007; Merz & Blöschl, 2003; Sikorska et al., 2015; Stein et al., 2019; Tarasova et al., 2020). However, very few of those studies evaluate how catchment and climate attributes influence flood generating processes (Merz & Blöschl, 2003; Stein et al., 2019).

Being able to estimate which flood generating processes can be expected in a catchment is relevant for many applications. For hydrological model development it is important to know which process representations must be included (Clark et al., 2016); for model evaluation it can help to evaluate model outputs in the sense of getting the right results for the right reasons (Kirchner, 2006). Moreover, knowing which catchment attributes are relevant for processes in various areas might improve the choice of donor catchments for flood predictions in ungauged catchments through regionalization (Rosbjerg et al., 2013). Furthermore, climate change can drive changes in flood process, which may affect flood magnitude (Blöschl et al., 2017, 2019). Knowing the temporal and spatial distribution of processes can potentially inform or explain changes in flood characteristics.

Based on existing literature, we can formulate several hypotheses regarding which climate and catchment attributes we expect to influence the mix of flood generating processes. In the following section, we will describe the studies that inform the hypotheses which are summarized in Table 1. Specifically, we will detail influencing factors for the five hydrometeorologic processes described by Stein et al. (2019): excess rainfall,

Table 1
Literature Informed Hypotheses on Which Catchment Attributes Influence Flood Generating Processes

Attribute	Study location	Positive influence	Negative influence	Neutral/ no influence	Reference
Elevation	DE, DE, US, US	Rainfall/snowmelt			Sui and Koehler (2001); Freudiger et al. (2014); Musselman et al. (2018); Li et al. (2019)
Area	AT	Long rain/excess rain	Short rain	Snowmelt	Merz and Blöschl (2003)
	CH		Short rain		Weingartner et al. (2003)
Slope	US	Snowmelt			Chang et al. (2014)
	UK/US/SE	Short rain			Tetzlaff et al. (2009)
	AT, CH	Short rain			Gaál et al. (2012); Weingartner et al. (2003)
	US			Short rain	Pitlick (1994)
Round catchment shape	-	Short rain			Ward (1978); Murthy (2002)
	CZ			Short rain	David and Davidova (2014)
Precipitation intensity	-	Short rain	Excess rain		Rosbjerg et al. (2013)
	US	Short rain			Pitlick (1994)
	US	Rainfall/snowmelt			Musselman et al. (2018)
Precipitation peak winter	UK	Excess rain			Institute of Hydrology (IoH) (1999)
Precipitation peak summer	DE	Rainfall/snowmelt			Sui and Koehler (2001)
Mean annual precipitation	AT	Excess rain	Short rain		Merz and Blöschl (2009)
	US	Rainfall/snowmelt			Li et al. (2019)
Aridity	Global	Short rain, long rain	Excess rain		Stein et al. (2019)
	-	Short rain			French and Miller (2011)
Soil storage	AT	Excess rain, long rain	Short rain		Merz and Blöschl (2003)
	US		Snowmelt		Chang et al. (2014)
Vegetation	US	Excess rain			Lull and Reinhart (1972)
	US, US		Short rain		Osterkamp and Friedman (2000); Shafer et al. (2007)
	US			Short rain	Pitlick (1994)
	-, US		Snowmelt		D. Miller (1964); Storck et al. (2002)
	US, US, US		Rainfall/snowmelt		Marks et al. (1998); Marks et al. (2001); D. Miller (1964)
Subsurface storage	-	Excess rain, long rain	Short rain		Rosbjerg et al. (2013)

Note. Positive (negative) influence here means that an increase in the attribute leads to an increase (decrease) in the specified process. Study location indicates where the study took place, with a dash (-) indicating that location was not specified.

which is saturation excess flow, short intense rainfall within 24 h, long rainfall over several days on previously dry ground, snowmelt, and rain-on-snow.

1.1. Flood Hypotheses—What Do We Expect?

Which processes in a catchment generate flood events depends on two factors: the availability of the flood producing input, and how the catchment stores and transmits water. Here, we briefly outline some of the possible effects of climate and catchment attributes on flood generation. While we look at a single catchment or climate attribute (or grouped attributes), each attribute usually does not exist independently in space. Climate, topography, soils, vegetation, and geology are tightly interwoven and lead to a myriad of highly correlated attributes. However, describing all possible interactions between attributes would be out of the scope for this paper. Interaction between aridity and snow fraction with all other attributes are included as they have been shown to be influential for flood process distribution (Berghuijs et al., 2016; Stein et al., 2019).

1.1.1. Climate and Weather

The flood generating forcing conditions depend on both climate and weather. The spatial and temporal distribution of precipitation and temperature influence snowpack accumulation. Locations with winter precipitation and winter temperatures continuously below zero can accumulate a snowpack that will not melt until the spring or summer. In these regions rain-on-snow events occur early in the season (September) or in spring/early summer when precipitation changes to rain again. In catchments with winter temperatures fluctuating around freezing, rain-on-snow events can also occur during the winter (McCabe et al., 2007; Musselman et al., 2018). Southern Germany for example often experiences floods in late December caused by rain-on-snow (Sui & Koehler, 2001). In glacierized catchments glacier melt can be a substantial contribution to streamflow in the summer (Merz & Blöschl, 2003; Sikorska et al., 2015).

For floods generated by short rain, long rain or excess rainfall, the availability of forcing condition is dependent on rainfall and evapotranspiration distribution. As the name implies short rainfall floods occur after short intense periods of rainfall that exceed infiltration capacity or quickly saturate the catchment (Merz & Blöschl, 2003). In arid regions, this flood type occurs often. Convective thunderstorms that exceed infiltration capacity are a common precipitation input (French & Miller, 2011). At the same time, a lack of snow and high evaporation leading to dry soils can make other flood processes less likely. One distinction between excess rainfall and long rainfall is based on antecedent conditions (Stein et al., 2019). Antecedent conditions are influenced by the seasonality of precipitation and evapotranspiration. If the precipitation maximum is in the cold season, then precipitation and evapotranspiration are out of phase. This means that the precipitation maximum falls during a time when the drying of the soil is at a minimum, leading to saturated conditions. This increases the chances of excess rainfall floods. With in phase seasonality, precipitation maximum and evapotranspiration maximum align, leading to drier conditions. Heavy multiday rainfall is needed to saturate catchment storage before runoff can increase.

In catchments with steady forcing conditions, for example under continuously saturated conditions with no snow, flood generating process is independent of catchment attributes because of the five possible processes only one can occur. Catchment attributes might influence the runoff coefficient or flood magnitude but not flood generating process. However, this is rare. Most catchments experience multiple flood generating processes (Merz & Blöschl, 2003; Sikorska et al., 2015; Stein et al., 2019; Tarasova et al., 2020). In catchments with variable forcing conditions, catchment storage and transmission behavior will heavily influence which process generates runoff high enough for a flood event. For example, Viglione and Blöschl (2009) find that the largest flood magnitudes are reached when storm duration is similar to catchment response time. A catchment with a short response time and therefore attributes that can indicate a short response time (slope, shape, and area) would likely have higher magnitude floods after short rain events, thus making high magnitude short rain floods more likely than long rain floods. However, a short time of concentration can also be reached through prior saturation of the catchment (Acreman & Holden, 2013; Ward, 1978). If this is the case, it will mean the catchment is more prone to excess rainfall floods.

Snowmelt and the interaction of rainfall and snowmelt is dependent on snowpack conditions. These depend both on climate conditions as well as weather conditions during snowpack accumulation and melting season. Rainfall retention capacity of a catchment varies depending on the snowpack conditions (Singh et al., 1998; Würzer et al., 2016). This influences the reaction of the snowpack to rainfall, thus increasing or decreasing the chance of a rain-on-snow flood. This kind of flood is strongly dependent both on antecedent condition of the snowpack and the rainfall producing weather system (Marks et al., 1998, 2001; Musselman et al., 2018; Sui & Koehler, 2001; Würzer et al., 2016). Li et al. (2019) note for the United States that the rainfall amount received by a region is relevant for the frequency of rainfall/snowmelt events. Windward slopes receive more rainfall than leeward slopes due to rain shadow effect. Consequently, these regions have a higher occurrence rate of rainfall/snowmelt events.

1.1.2. Slope

The influence of catchment average slope varies between different flood processes. Chang et al. (2014) find that steep catchments transport meltwater more quickly to the stream, especially in combination with thin soils. In regard to rainfall induced floods, slope can influence transit time, with steeper catchments transporting water more quickly to the outlet (Tetzlaff et al., 2009). Longer transit times in flatter terrain only occurred in areas with permeable soils. Slope can also be an influential attribute as a proxy for soil thickness, where steeper slopes may also have thinner soils and therefore less storage and quicker transmission (Pitlick, 1994). Quicker transmission leads to more erosion, which leads to more efficient drainage systems, which leads to more short rain floods (Gaál et al., 2012; Weingartner et al., 2003). Despite these findings, in a global flood frequency study A. Smith et al. (2015) did not find slope to be a good predictor for the shape of the flood frequency curve. Similarly, Pitlick (1994) demonstrates that flood magnitude does not vary with catchment slope in their study region. The effect of slope on flood generating processes is therefore still under debate.

1.1.3. Area

It is plausible that catchment area may affect the flood process type in some environments. While absolute flood magnitude increases with area (Murthy, 2002; A. Smith et al., 2015), specific magnitude (normalized by area) sees a decrease with area (Eaton et al., 2002; Padi et al., 2011). These relationships differ between different climates (Padi et al., 2011) and flood generating processes (Eaton et al., 2002; Merz & Blöschl, 2003). Rainfall variability in arid regions means area is not a good predictor of flood magnitude (A. Smith et al., 2015; Tooth, 2000).

Area affects time of concentration, with smaller catchments having a shorter time of concentration. Small catchments can be covered in its entirety by high intensity convective storms. A larger catchment might only be partially covered by a convective storm, leading to rainfall amounts too small to cause a flood (Merz & Blöschl, 2003; Weingartner et al., 2003). Therefore, in humid areas short rainfall flood events might be more common in smaller catchments due to these two effects. Merz and Blöschl (2003) found only minor change in specific flood magnitude of snowmelt floods with catchment area.

1.1.4. Shape

Catchment shape influences flood peak shape (Murthy, 2002; Ward, 1978). In a round catchment with simultaneous input everywhere the flood waves from different parts of the catchment will concentrate quickly. This effect will be strongest when storm duration is the same as catchment time of concentration (Blöschl et al., 2013; Viglione & Blöschl, 2009). There are exceptions though. Elongated catchments can receive very high peak flow if a storm cells moves along the catchment toward the outlet. The flood wave from upstream will overlap with runoff generation downstream causing high peak flow (Murthy, 2002). David and Davidova (2014) did not find a relationship between catchment shape and flood magnitude.

1.1.5. Soils

Soil development is influenced by a range of factors. Soil properties like storage and infiltration capacity and soil depth, among others, are therefore closely tied to the climate, geology, topography, flora, and fauna of a catchment. Those properties in turn influence runoff and storage behavior within the catchment. A high storage capacity in combination with a high infiltration capacity requires larger input volumes before runoff occurs (Merz & Blöschl, 2003). Once storage capacity is exceeded floods can reach larger magnitudes which is visible as a step-change in the flood frequency curve (Rogger et al., 2012). In arid and semiarid regions both types of runoff generation, saturation excess and infiltration excess, can occur. In areas with low infiltration capacity for example due to crusting or bare rock surfaces, infiltration excess is prevalent (Cantón et al., 2002; Ries et al., 2017; Sohrt et al., 2014). Wood et al. (1990) found that soil properties are most relevant for floods of small magnitude while rainfall properties are more relevant for larger magnitude floods. For the same catchment, varying temporal distribution of rainfall volume or intensity can elicit a different reaction or runoff process (Cantón et al., 2002; Ries et al., 2017).

1.1.6. Open-Water Storage

Wetlands, lakes, and reservoirs contribute to the storage capacity of a catchment. A wetland's effect on downstream flood characteristics depends largely on the saturation state. Once saturated most rainfall contributes immediately to runoff (Acreman & Holden, 2013; Bullock & Acreman, 2003; McCartney et al., 1998). Catchments with larger storage capacity would therefore be more likely to flood after saturated conditions (excess rainfall floods). If a reservoir adds storage capacity to the catchment will depend on the local water resources management plan.

1.1.7. Elevation

Merz and Blöschl (2003) found that in Austria flood processes and time of occurrence changes with elevation. The higher the elevation the later in the spring snowmelt floods occur. Elevation is directly related to temperature and precipitation. For the United States, McCabe et al. (2007) and Li et al. (2019) find rainfall/snowmelt floods to be most prevalent at mid-elevation catchments (1,000–1,500 m).

1.1.8. Geology

The way the topography of a drainage network will form depends on geology as well as climate, vegetation, and soils. Large subsurface storage dampens flood response. This leads to less erosion and more soil development thus again increasing storage (Rosbjerg et al., 2013).

1.1.9. Vegetation

The influence of vegetation, in particular deforestation and reforestation, has been discussed in depth in the literature (Rogger et al., 2017). While some large scale studies find forests to have an effect on magnitude and frequency of flooding (Bradshaw et al., 2007), others disagree (Bruijnzeel, 2004; Calder & Aylward, 2006). In regard to flood processes, snowmelt floods have been shown to be influenced by coniferous trees, which intercept snowfall and increase sublimation rates (Storck et al., 2002). Vegetation decreases quick runoff since it both increases surface roughness as well as soil infiltration capacity (Lull & Reinhart, 1972). Some of the disagreement is due to scale. The effect of land-use on floods is stronger in smaller catchments (Calder & Aylward, 2006) and for smaller flood magnitudes (van Dijk et al., 2009). Vegetation is an important influence on runoff behavior in semiarid and arid regions as it increases infiltration capacity (Osterkamp & Friedman, 2000; Ries et al., 2017; Shafer et al., 2007).

1.2. Aims and Objectives

The majority of studies reviewed above do not in fact evaluate the influence of catchment attributes on flood generating processes. Instead, the majority of previous work only evaluates the impact of catchment attributes on indices that describe floods (e.g., the runoff coefficient, flood magnitude, flood duration, etc.). We can only develop hypotheses for the effect on flood generating processes, as we have done in Table 1. A comprehensive, data-based, comparative study to test the influence of catchment attributes on flood generating processes is still missing. To this end, the aim of this study is to evaluate which assumptions and prior findings are supported when tested on a large sample of catchments across several climates within the contiguous United States. We hypothesize that climate attributes will be very influential on flood generating processes as they have the strongest influence on seasonal availability of flood process deciding input and saturation conditions. Catchment attributes that influence the storage behavior of the catchment will likely have an effect as well (Merz & Blöschl, 2009).

We have previously developed and tested the first global event-based flood classification methodology (Stein et al., 2019). In this study, we use that methodology to classify flood generating processes for a large sample of catchments across several climates (Section 2.2). We then explore the influence of catchment attributes on flood generating processes. Finding influential attributes in correlated data sets is challenging, as the correlation among attributes might obscure findings (Dormann et al., 2013). We therefore use two approaches that complement each other and allow interpretation while taking collinearity into account. The first is a statistics based approach that evaluates the influence of each attribute individually (Section 2.4.1). The second approach uses a random forest model and an interpretable machine learning method (accumulated local effects [ALE], Apley and Zhu (2016), Sections 2.4.2 and 2.4.3) which is unbiased toward correlated predictors (Molnar et al., 2018).

2. Methodology

2.1. Data

We used the publicly available CAMELS data set which combines hydro-climatological data (Newman et al., 2015) with catchment attributes (Addor et al., 2017) for 671 catchments in the contiguous United States. The daily data covers a time period from 1980 to 2014. All catchments have a minimum of 20 years of data included in the analysis (Kjeldsen, 2015). The majority of catchments (84%) include 34 years of data. Newman et al. (2015) selected these catchments specifically to have minimal human influence. The majority are therefore small headwater catchments. For the flood event classification daily observed streamflow data and Daymet meteorological forcing data (precipitation, temperature) included in the CAMELS data set were used. For the soil moisture routine (Stein et al., 2019) available water capacity of the soil is a necessary variable. This is not included in CAMELS. Instead, we used data from the Gridded National Soil Survey Geographic (gNATSGO) Database for the Conterminous United States (Soil Survey Staff, 2019). Newman et al. (2015) provide daily actual evapotranspiration values from the Sacramento Soil Moisture Accounting Model. Addor et al. (2017) extended the CAMELS data set by Newman et al. (2015) to include continuous and categorical catchment attributes in six thematic groups: topography, climate, soil, vegetation, geology, and streamflow indices. Figure 7 offers a full list of attributes. Detailed descriptions and definitions for each attribute can be found in Tables 1–6 in Addor et al. (2017). Examples of continuous attributes are mean annual precipitation or fraction of the catchment covered by forest. Categorical attributes were not considered in this study, since the statistics method could not account for categorical attributes. This way both methods analyzed the same catchment attributes. We additionally calculated catchment shape as represented by the elongation ratio (Schumm, 1956). A value closer to 1 indicates a round catchment; a value closer to zero an elongated catchment.

2.2. Flood Process Classification

Flood events were identified using a peaks-over-threshold approach. It identifies the highest independent streamflow peaks in the time series. The number of peaks varies depending on the threshold which can be set to find a certain number of peaks per year. The R function "findPeaks" from the package "quantmod" (Ryan & Ulrich, 2019) was used to identify all peak streamflow days. Only independent flood peaks higher than mean daily streamflow were kept for further analysis. For any flood peak identified by "quantmod" to be independent from another, the time difference between both peaks has to be larger than the mean rising time calculated from 5 "clean hydrographs" (Cunnane, 1979). They are the 5 highest peaks with a large time difference to the previous peak. An additional independence criterion was that a trough between two peaks needed to be less than 2/3 of the first peak (Cunnane, 1979). Two subsets of peaks with different magnitudes were identified: One with an average of one event per year (larger peaks) and one with an average of three (smaller peaks) events per year to compare if a difference in magnitude has an effect. We included this option because several studies indicate that land use or storage capacity are more influential for floods of smaller magnitude (Rogger et al., 2012; van Dijk et al., 2009; Wood et al., 1990). If there were more events per year than specified, the smallest peak events were removed.

We classified the identified flood events in each catchment into one of five hydrometeorological generating processes (Stein et al., 2019): excess rainfall, short rainfall, long rainfall, snowmelt, and a combination of rain and snowmelt (rainfall/snowmelt). A decision tree (Figure 1) evaluated hydrometeorological conditions in a 7-day time period before any flood event. It used the date of the flood event and hydrometeorological input data, as well as soil moisture and snowmelt estimates obtained from a simple lumped model routine run at a daily time step (Stein et al., 2019). Critical temperature for snowfall and melt was set to 1 C (Jennings et al., 2018). The thresholds of the tree are based on the hydrometeorological time series of each catchment. This methodology enabled us to classify a large sample of flood events across various climatic regions without prior knowledge about dominant flood generating processes for each catchment. The tree is structured to first evaluate if snowmelt and rainfall occur simultaneously. This is a simplified definition of rain-on-snow floods, as it does not take the snowpack energy balance into account (Marks et al., 1998; Pomeroy et al., 2016). In a next step it checks if snowmelt was higher than the threshold, which would indicate a snowmelt flood event. If neither was the case the tree evaluates if higher than average multiday

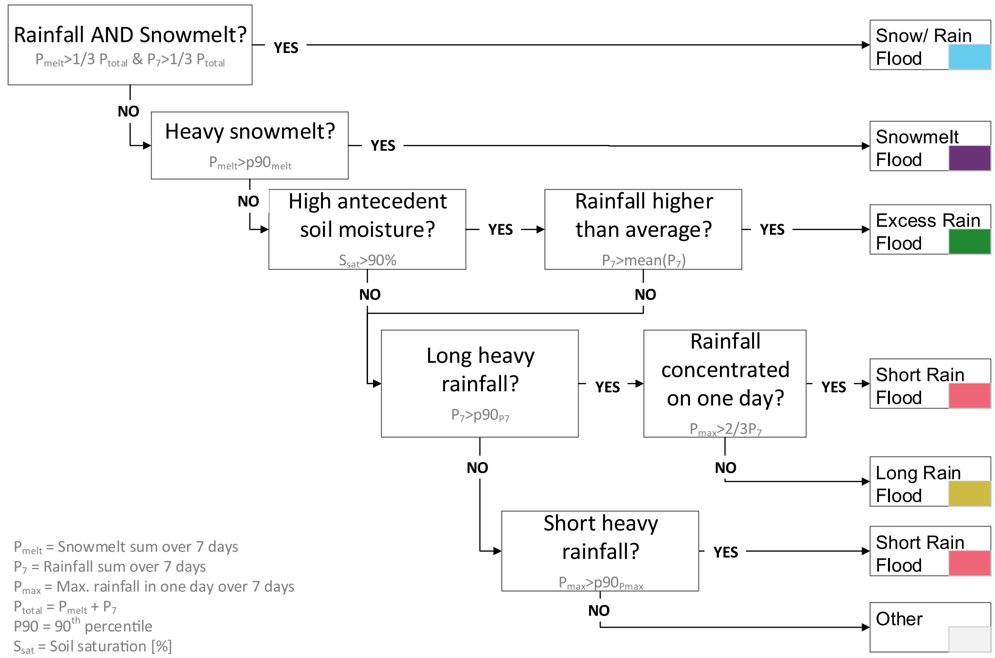

Figure 1. Conceptual decision tree for location independent flood process classification adapted from Stein et al. (2019).

rainfall fell on previously saturated ground to determine if the flood event was an excess rainfall flood. If that was not the case it evaluates whether the thresholds for long rainfall and then short rainfall are exceeded. If no process could be identified, the class "other" will be assigned. Events classified as other were not considered in this analysis. For an in-depth description and evaluation of this methodology please refer to Stein et al. (2019).

2.3. Climate Type Definition

Climatic catchment attributes influence catchment flow behavior (Addor et al., 2018; Berghuijs, Sivapalan, et al., 2014; Jehn et al., 2020; Knoben et al., 2018). We wanted to determine whether the importance of other catchment attributes varies between the different climate types. The CAMELS data is well suited to answer this as the catchments lie within various different climatic regions. In regard to flood process distribution, Berghuijs et al. (2016) note the influence of aridity on the distribution of excess rainfall and short/long rainfall floods. There will be very few or no snowmelt or rainfall/snowmelt floods in catchments with small or zero fraction of precipitation falling as snow. Since the importance of these two attributes (aridity and fraction of snow) is already known, we split the data set into different climate types to evaluate the interaction of these attributes with others. Based on two climatic indices from the CAMELS data set (Addor et al., 2017), the catchments were separated into three different groups: wet, dry, and snow influenced catchments. Wet catchments were defined as catchments with an aridity index $PET/P < 1$. Potential evapotranspiration in those catchments is lower than precipitation (i.e., energy-limited catchments). Dry catchments have an aridity index >1 respectively with mean potential evapotranspiration larger than mean precipitation (i.e., water-limited catchments). All catchments with a fraction of precipitation falling as snow higher than 20% were designated as snow catchments, regardless of their aridity. Twenty percent was chosen as value representing different literature options, while also ensuring approximately equal numbers of catchments in each climatic region. There is no clear definition which snow fraction indicates a snow-dominated catchment. Berghuijs, Woods, and Hrachowitz (2014) use 15% as a threshold value. Davenport et al. (2020) on the other hand, show a change from peak streamflow from spring/summer to winter from a winter precipitation fraction of snow of 30%. The flood process classification showed that these thresholds deliver roughly similar numbers of catchments for all climate types while grouping catchments with snowmelt flood contributions together The distribution of catchments for each climate type is depicted in Figure 2.

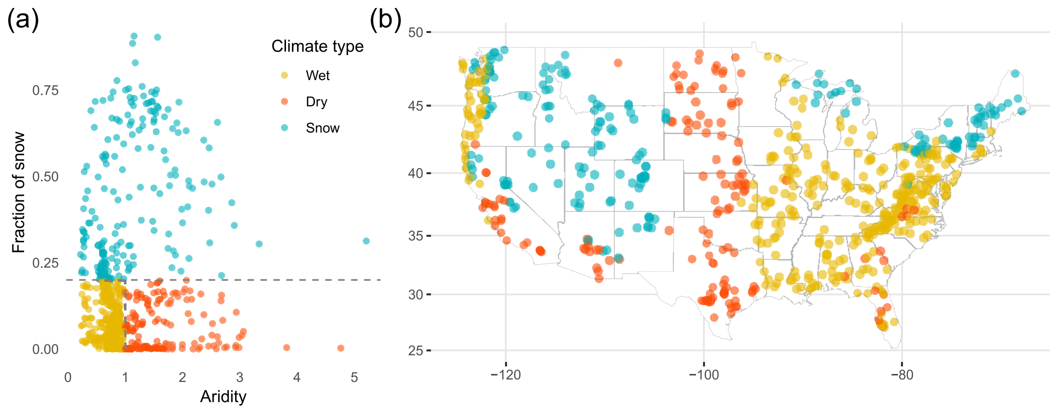

Figure 2. (a) Classification of 671 CAMELS catchments into three climate types wet, dry, and snow based on aridity and fraction of snow. Climate type thresholds are indicated through dashed lines. Aridity and fraction of snow taken from Addor et al. (2017). (b) Spatial distribution of the three climate types for the CAMELS catchments.

2.4. Estimation of the Influence of Climate and Catchment Attributes on Flood Generating Processes

We employed two methods to evaluate which attributes influence the flood generating process distribution. A statistics-based approach comparing frequency distributions and an interpretable machine learning approach called accumulated local effects (ALE). An overview over the benefits and drawbacks for both methods is presented in Table 2. Both methods are described in detail further below.

2.4.1. Comparison of Frequency Distributions

We first evaluated the influence of each continuous attribute on each flood generating process. We stratified this analysis by climate type (Section 2.3) because it is plausible that the influence of attributes varies with environmental setting. For this, we applied a comparative hydrology approach (Falkenmark & Chapman, 1989; Gaál et al., 2012). To assess the influence of one attribute on one process, we compared two distributions of the same attribute with each other sampled across all catchments within one climate type. Each catchment can contribute the same attribute multiple times depending on the number of events. We compared the empirical cumulative distribution function (ecdf) of the attribute, sampled from all catchments with events driven by that process, with the ecdf associated with all events (independent of process) (Figure 3a/b). If the two distributions differ, we inferred that this attribute influences the occurrence of that process (Gaál et al., 2012; Merz et al., 2006; Pianosi & Wagener, 2015). For example, if a catchment had 15 excess rain events and the mean annual rainfall attribute in that catchment was 400 mm per year, then this catchment would contribute the value 400 mm 15 times to the specific process distribution. If the next catchment had 10 excess rainfall events and a mean annual rainfall of 350 mm per year, it would contribute the value 350 mm 10 times to the same distribution. Similar methods have been applied by Merz et al. (2006) to study runoff coefficients and by Gaál et al. (2012) to evaluate flood duration.

To make the distributions comparable, all attribute values were normalized (min-max-normalization). To summarize the divergence between the two distribution functions, we calculated the mean difference between 100 values along the ecdf curve for each process and the curve for all events. The resulting value may be either positive or negative. Figure 3c illustrates that a negative (positive) value indicates an increased occurrence of the process for smaller (larger) values of the attribute. Figure 3d displays how the mean difference between each process curve and the full range curve translates into a single metric. We used cumulative frequency density functions instead of frequency density functions as they can be calculated without any prior parameter assumptions (Pianosi & Wagener, 2015).

We chose this approach over a correlation-based analysis since a simple correlation analysis would only be able to determine linear relationships between catchment and climatic attributes and flood generating processes. The comparison of ecdf curves is able to indicate both linear and nonlinear relationships by taking into account variations across the whole attribute space. Although rank correlation would be able to

10.1029/2020WR028300

Table 2
Benefits and Drawbacks of the Two Methods Used to Evaluate Attribute Influence

Strengths	FDM	ALE
Compare influence between different processes	x	
Compare influence between different attributes		x
Shows direction of influence	x	
Able to account for nonlinear relationships	x	x
Weaknesses		
Sensitive to correlated attributes	x	
Sensitive to unequal sample size	x	
Sensitive to small sample sizes	x	
Sensitive to unevenly distributed data		x
Needs domain knowledge for interpretation	x	
Uncertainty of model prediction		x

Note. FDM refers to the frequency distribution method, ALE to the accumulated local effects.

give similar results as the curve summary statistics, the comparison over the whole curve additionally allows a visual interpretation of influential attribute ranges (see Figures S5–S9 in the supporting information).

A drawback of this approach is that the distribution functions are sensitive to unequal sample size and to small samples (e.g., the overall number of snowmelt and rainfall/snowmelt flood events in dry catchments is small). If one sample is much larger than the others, it dominates the comparison distribution (e.g., there are much more excess rainfall events in wet catchments than any other process). A small sample size may lead to uncertainties in the real distribution function (an example in Figure 3b is the distribution of snowmelt events in dry catchments). For this reason, more weight should be given to distributions based on a larger sample size. In Figure 3d, this is taken into account by adjusting the point size according to the sample size. Another limitation is that only continuous variables can be analyzed in this way. Lastly, a limitation of the applied summary statistic is that attributes that reverse their influence (i.e., the ecdf curves cross one another) would sum to zero. We visually checked all curves and there is only one case (influence of mean annual precipitation on rainfall/snowmelt floods in "snow" climate) where this occurs.

2.4.2. Random Forests

A random forest is a machine learning model approach that creates and combines multiple regression trees (Breiman, 2001). Addor et al. (2018) list the benefits of random forest models as allowing multiple predictors, being able to incorporate nonlinear relationships, flexibility, a reduced risk of data overfitting compared to individual regression trees, interpretability, and computational efficiency. These benefits make random forest models particularly suitable for predicting hydrological signatures in space where that is possible (Addor et al., 2018; Booker & Woods, 2014).

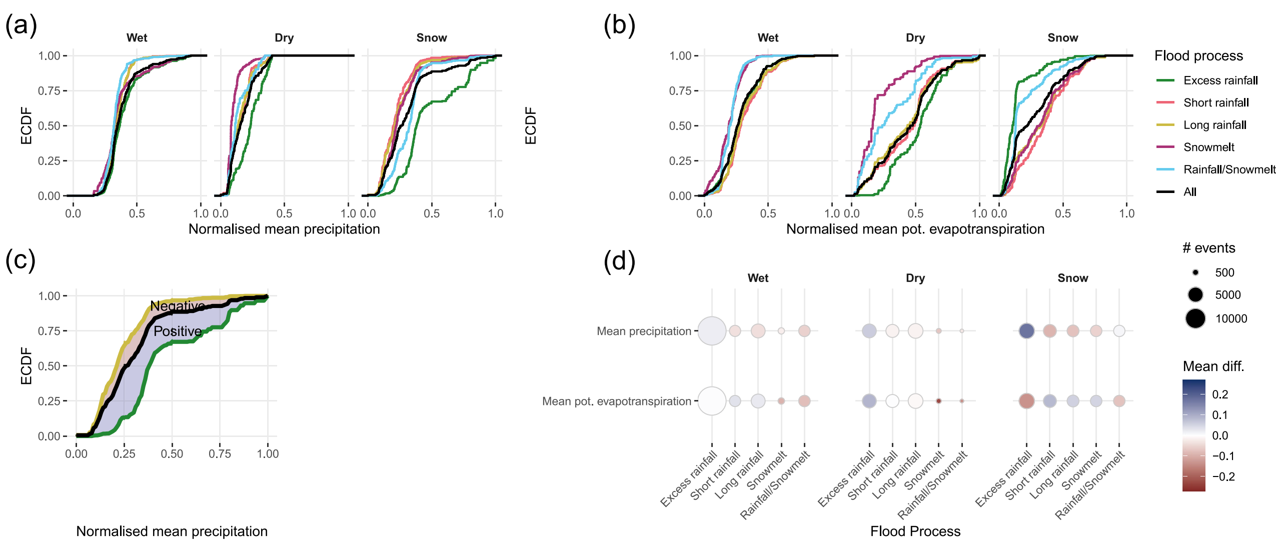

Figure 3. Example figure to explain distribution comparison and difference between distribution value. (a) Empirical distribution functions of normalized mean annual rainfall for each flood generating process compared to all events (black). (b) Empirical distribution functions of normalized mean annual potential evapotranspiration for each flood generating process compared to all events (black). (c) Example distribution differences—the colored space indicates the difference between the long rainfall events and all events (red, difference between distributions is negative) and difference between excess rainfall events and all events (blue, difference between distributions is positive). (d) Summary statistic—Mean difference between distribution value for both example attributes. Color indicates the direction and strength of difference. The number of events that contribute to a distribution is indicated through point size.

We used a random forest model in two ways: (1) to explain flood generating processes using interpretable machine learning; and (2) to predict flood process distributions from catchment attributes. The latter can demonstrate both that catchment attributes influence the distribution of flood-generating processes, and that it is possible to predict the flood process distribution in ungauged catchments. For each climate and process a separate random forest model was used. Each estimated the percentage of events contributed by the respective process. Prediction accuracy was evaluated using 10-fold cross-validation (see e.g., Addor et al., 2018). Random forest models tend to overfit on training data and cross validation gives a better evaluation of prediction accuracy than performance evaluation based on training data (Dormann et al., 2013). Therefore, the data set was split into 10 equal-sized samples. Ten random forest models were trained with nine parts of the data and evaluated on the respective tenth part. This way prediction accuracy could be evaluated for all catchments. Prediction accuracy was calculated as the coefficient of determination (R^2). Robustness of the random forest model was checked by repeating random forest generation with 50 different initial random seeds.

2.4.3. Accumulated Local Effects Applied to Random Forest

To interpret a random forest model, Addor et al. (2018) refer to the possibility of determining the influence of an attribute on the outcome through variable importance (the increase in error when a predictor is shuffled). However, this metric is unsuitable for data sets with correlated features (Degenhardt et al., 2019; Dormann et al., 2013; Toloşi & Lengauer, 2011) such as the CAMELS data set (Jehn et al., 2020). Instead we use an interpretable machine learning approach, accumulated local effects (ALE), that is not biased against correlated attributes (Apley & Zhu, 2016). We provide a brief introduction into interpretable machine learning and a more in depth introduction into ALE in the supporting information. ALE plots improve the application of more commonly used partial dependence plots (Anchang et al., 2020; Friedman, 2001; Molnar, 2019). After a model was fit to the data, ALE plots evaluate the change in model prediction over a small interval of an input variable (Apley & Zhu, 2016). Interval size is determined by quantiles in the distribution (Molnar, 2019). For all observed data points in that interval, differences in prediction between the interval boundaries are calculated. This way the change in the variable of interest (local effect) is recorded, disregarding any correlation effect by other variables. The local effects for each boundary are accumulated into a curve and centered around zero. Example ALE curves are displayed in Figure 4a (black lines) for mean precipitation, mean potential evapotranspiration, and water fraction in the soil. Any divergence from zero reveals a conditional effect of the attribute on the prediction outcome. Blue bars in Figure 4a plot the divergence which indicate a conditional effect, from now on referred to as influence. The random forest model has been implemented using the "randomForest" package in R (Liaw & Wiener, 2002) and the ALE were calculated using the package "iml" (Molnar et al., 2018).

ALE are a relatively new method. They have proven their applicability in several fields (e.g., Anchang et al., 2020; Brown et al., 2020; Konapala et al., 2020). One limitation is that ALE evaluate the reaction of a model to changes in an attribute. Results are not directly based on data. ALE calculated on a model with low performance will yield less reliable results (Zhao & Hastie, 2019). Another limitation is that ALE plots do not give reliable results in attribute ranges with scarce data (Molnar, 2019). Interval size over which the ALE is calculated is not regular but instead is based on an equal number of observations per interval. In unevenly distributed data this can lead to large interval sizes. In the CAMELS data set that is the case for the attributes water fraction, organic fraction, and carbonate rocks fraction. Figure 4 demonstrates how the unevenly distributed fraction of water in the soil data (Figure 4b) translates into only two intervals, one at zero and one at 10 (Figure 4a, blue bars).

To summarize the influence an attribute has into one number (instead of a curve), we calculated the mean absolute values of the ALEs (blue bars in Figure 4a). This value is comparable between attributes of the same model, but not between different models. Therefore, the summarized values were normalized (min-max-normalization) for comparability. In the example given mean evapotranspiration would rank as most influential with a value of 0.93, followed by mean precipitation at 0.84. Due to the uneven distribution water fraction would still have a relatively high importance at 0.78.

We applied two approaches to test the robustness of the results. ALE were calculated for each of the 50 random forest iterations using different random seeds. We recalculated the ALE with each attribute missing in turn. Both results are included in the supporting information.

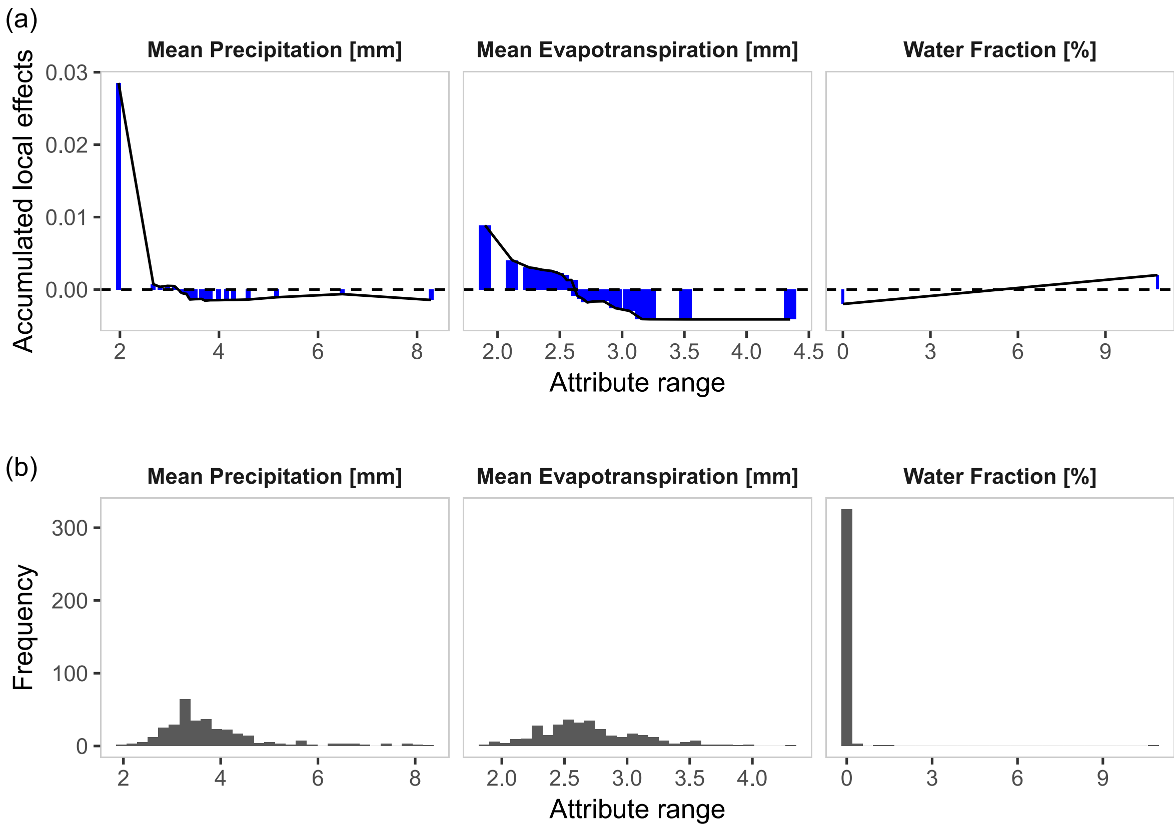

Figure 4. Example figure accumulated local effects plot and its limitation. (a) Accumulated local effects plot for predictions of snowmelt floods in wet climate catchments. The dashed line is the zero line. Blue bars indicate interval locations identified by the ALE algorithm. Their divergence from zero is calculated and the mean is taken as a summary value. (b) Data distribution for each of the example attributes.

3. Results

3.1. Event Classification

Figure 5 illustrates the contribution for each flood generating process in each catchment. Excess rainfall floods are most common in the eastern and north-western United States. Short rainfall floods occur most often in the western United States. Snowmelt floods are most common in the western-central United States where the Rocky Mountains are. In the north-eastern United States rainfall/snowmelt floods are common. Long rainfall floods are most common in the great plains area in the central US. Out of all (61,764) identified flood events the majority of events are excess rainfall floods (Figure 6b). In wet climates excess rainfall floods occur in every catchment (Figure 6a). In drier regions short and long rainfall flood events are more common, with fewer or no events classified as excess rainfall. The combination of rainfall and snowmelt rarely occurs, but several snowmelt floods were identified. Catchments with the climate type snow accordingly classify more events as snowmelt and rainfall/snowmelt. Several catchments with large percentages of snowmelt floods also classify large contributions from short rainfall/long rainfall events (Figure 6a). The majority of catchments (62%) has a contribution of less than 5% of events classified as "other." The catchment with the highest fraction of this kind of events has 50% of events classified as "other."

3.2. Distribution Comparison

The distribution of each catchment attribute for each process was compared with the distribution of each attribute across all processes. The more different the distribution, the more influential an attribute is for that specific process. The results are detailed in Figure 7a. The plotted empirical distributions functions are shown in the supporting information (see Figures S5–S9).

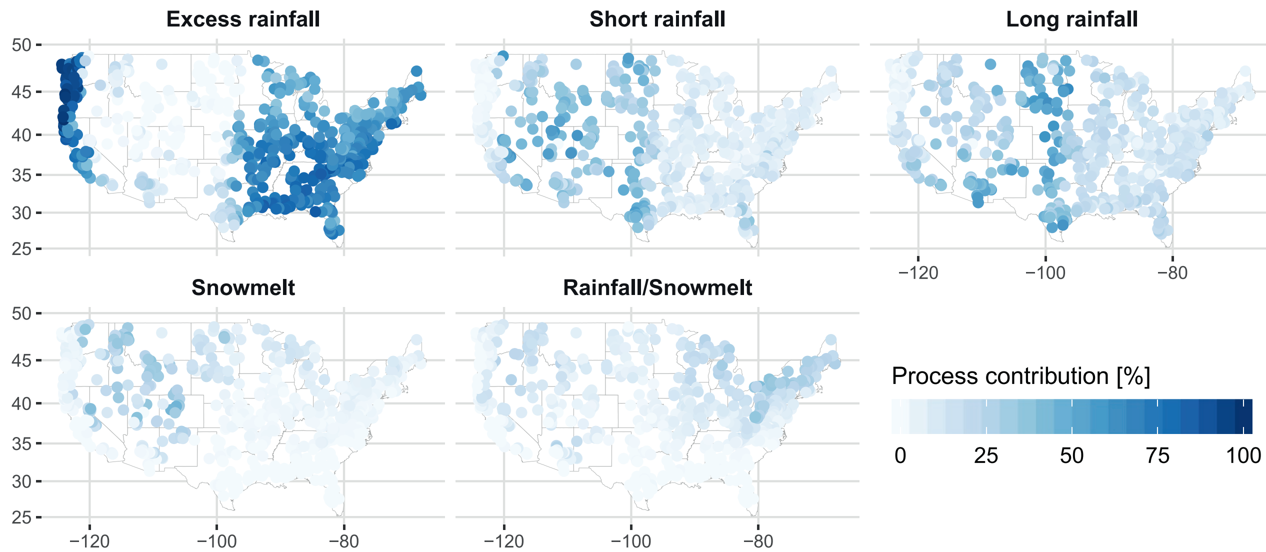

Figure 5. Contribution in percent for each flood generating process across the CAMELS catchment data set. Flood events are defined as peak-over-threshold with an average of 3 events per year.

From the distribution comparison (Figure 7a, read by row), we learn that in wet catchments (P > PET) catchment and climate attributes influence the mix of flood processes only marginally. The distribution of catchment attributes does not differ widely between catchments with different mix of flood generating processes. Mainly since all wet catchments are dominated by excess rainfall. Excess rainfall as a flood process is only slightly influenced by precipitation seasonality and mean precipitation. The other processes see a minor influence by further climate attributes. The only noticeable exception is the positive influence

Figure 6. (a) Contribution in percent for each flood generating process across the CAMELS catchment data set shown per catchment. Flood events are defined as peak-over-threshold with an average of 3 events per year. Catchments are sorted by their catchment ID (Addor et al., 2017) which approximates spatial proximity and an ordering from East to West. (b) Overview of number of events.

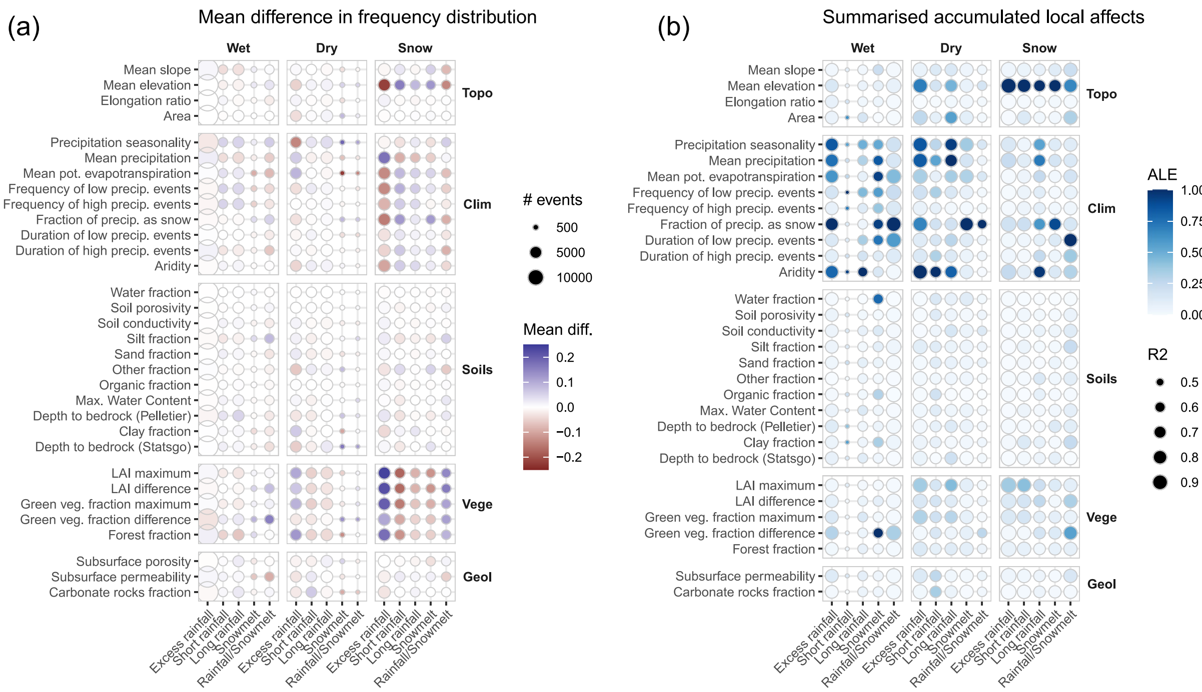

Figure 7. Attribute influence on process distribution. (a) Mean difference between the empirical distribution function (ECDF) of the attributes for a single process and for all events. The larger the absolute value the more different the two ECDF's. Size of points give an overview of how many events contributed to the distribution. Color indicates the direction of influence. Blue values point at a decrease of the process for smaller values of the attribute. Red values at an increase for smaller values of the attribute. Horizontal breaks point at different groups of catchment attributes. (b) Summarized accumulated local effects. For each climate type and flood process the accumulated local effects for all attributes were calculated. The point color shows the mean absolute values for each accumulated local effects curve. Higher values indicate increased importance. Values were normalized for each climate/process to enable comparability between climates/processes. Point sizes represent cross-validation R^2 prediction accuracy for the random forest model.

of differences in green vegetation fraction on snowmelt. We can therefore conclude that, of the attributes we have considered, aridity and fraction of snow, the same attributes we use to divide the data into climate types influence the distribution of flood generating processes. This is confirmed by the difference in distributions between the three climate types demonstrated by Figure 6a.

In drier catchments the differences in attribute distributions are stronger. Excess rainfall floods increase with higher precipitation and potential evapotranspiration and decrease with precipitation seasonality, for example, with a precipitation maximum in the summer. Increased vegetation (fraction of forest, green vegetation fraction, and leaf area index) similarly increase contribution from excess rainfall floods and decrease occurrences of other processes. Snowmelt floods most decrease with increasing potential evapotranspiration and increase with seasonality indicated both by precipitation seasonality and differences in green vegetation fraction.

The strongest differences in distribution can be seen in snowy catchments. Elevation decreases excess rainfall floods and rainfall/snowmelt floods and increases the occurrence of short rainfall, long rainfall and snowmelt floods. Several climatic attributes have a strong effect as well. Vegetation attributes have the strongest effect in Figure 7a in snow-dominated catchments. Similarly to drier climates, we can see with increasing vegetation an increased occurrence of excess rainfall floods and rainfall/snowmelt floods and a decrease in short rain/long rain and snowmelt floods. The same methodology applied to larger floods (peaks-over-threshold with one event per year) yields similar results (see Figure S1 in the supporting information).

3.3. The Influence of Catchment and Climatic Attributes Using Accumulated Local Effects

The summarized ALE are shown in Figure 7b. In contrast to Figure 7a the values here are standardized for each process/climate. Values are not comparable between processes but only between attributes for each process (i.e., read the figure by column). Therefore, for each process it can be assessed how the ranking of

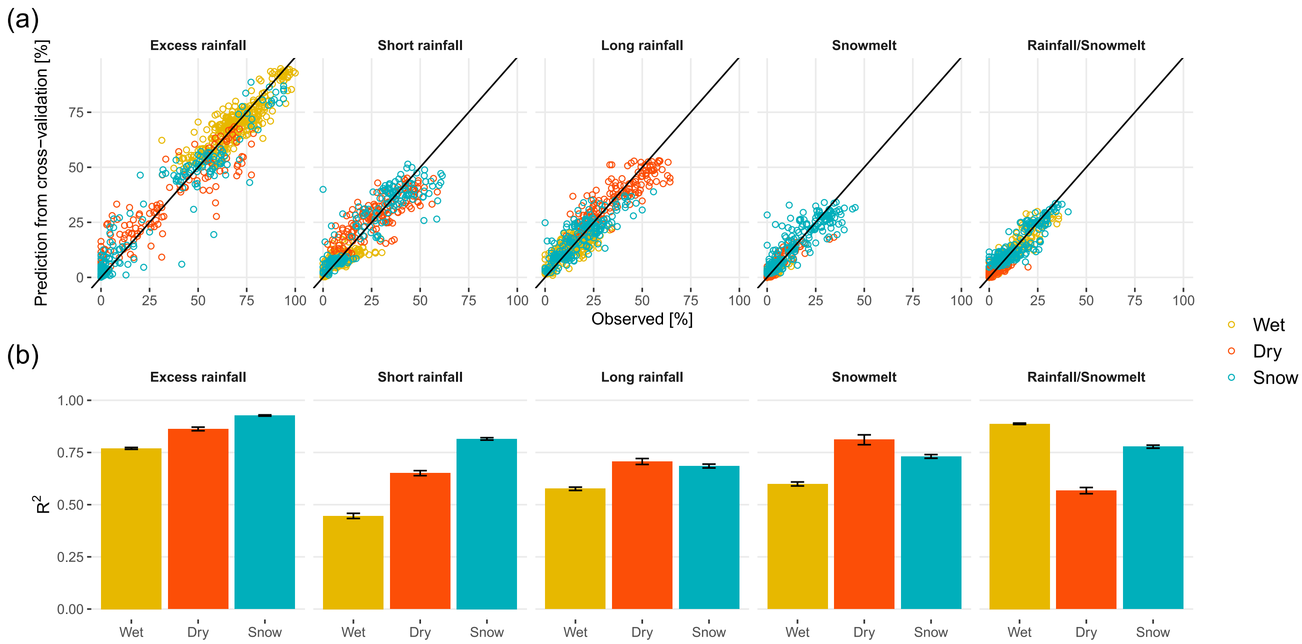

Figure 8. Random forest cross-validation results. For each climate type and flood processes a separate random forest was trained and validated through cross-validation. (a) Validation results in comparison to the observed classification. A black line indicates the perfect fit. (b) Mean and standard deviation R^2 of all 50 random forest models for the cross-validation.

attributes changes between climates. The results are robust over the 50 random forest models (Figure S2). There are some variations in ALE values depending on attribute combination (Figures S3 and S4). More information in that regard is included in the supporting information.

Precipitation seasonality and fraction of snow are ranked influential on predicted excess rainfall floods in wet climates. The process contribution from excess rainfall is influenced by the fraction of snow, since more rainfall/snowmelt floods decrease the contribution by excess rainfall floods (see Figure 6a). In dry climates aridity and mean annual precipitation are important as well as precipitation seasonality. Fraction of snow is less prominent. Climatic attributes in snow-influenced catchments on the other hand do not influence contribution of excess rainfall floods. Instead, elevation is the most relevant attribute for the spatial distribution of excess rainfall floods.

The distribution of short rainfall floods is not well predicted in wet catchments ($R^2 = 0.45$, Figure 8). Any conclusion here is therefore less reliable. However, in dry catchments aridity is the most dominant for predicting this type of event, whereas in snow-dominated catchments elevation is dominant. This is in contrast to long rainfall floods. While aridity is influential in wet climatic conditions, in dry climates precipitation seasonality and mean annual rainfall are more influential than aridity. In snow-dominated climates elevation is similarly important.

The spatial distribution of snowmelt induced flood events under wet climatic conditions is influenced by several catchment attributes. Mostly climate attributes such as fraction of snow and mean annual precipitation and potential evapotranspiration, but the difference in green vegetation fraction influences prediction as well. Water fraction, which refers to the top 1.5 m of soil marked as water in the soil database (STATSGO) (Addor et al., 2017), shows as relevant as well, although this is due to skewness of the data. In a dry climate only fraction of snow is influential and in snow-dominated catchments elevation and fraction of snow dominate.

The contribution of events caused by a combination of rainfall and snowmelt seems to be differently influenced by catchment attributes than sole snowmelt events. In wet and dry climate catchment fraction of snow is the most important attribute. However, in snow-dominated catchments average duration of dry periods seems to be most influential.

3.4. Predictions in Space Using Random Forest

A random forest was used as an unsupervised learning model to predict the distribution of each flood generating process and for each climate. The results of a 10-fold cross validation are presented in Figure 8. It demonstrates that prediction accuracy varies with process and climate. For all processes, higher observed contributions are slightly underestimated and low ones slightly overrated. For all processes there are few outliers. Most occur in snow-dominated catchments. From Figure 8b, we can see that prediction accuracy using all attributes is lowest for short rainfall events in wet climates ($R^2 = 0.45$) and highest for excess rainfall in snowy climate ($R^2 = 0.92$). Except for rainfall/snowmelt floods, prediction accuracy is always lowest in dry climates. The variance in prediction accuracy for models with different random seeds is negligible (black error bars in Figure 8b).

4. Discussion

4.1. Influential Catchment Attributes

A combined interpretation of the two methods takes the direction of influence (positive/negative) from the distribution comparison in Figure 7a. The accumulated local effects (Figure 7b) then confirm if that attribute is influential in comparison to other attributes. In addition to that a comparison between climates is possible using the distribution comparison (Figure 7a) as well. We interpret the combined results in regard to the hypotheses formulated in Section 1.1 and Table 1.

Catchments in snow dominated climate show the strongest influence of catchment attributes on flood process distribution. There is clear difference in catchment attributes between different flood distributions. In contrast, for wet catchments attributes do not vary strongly with different flood process distributions. It is thus important to keep in mind that any importance the ALE attribute to catchments in wet regions stems from minor attribute differences between those catchments. Therefore, aridity and snow fraction which delineate the climate group of wet catchments are by far the most influential attributes for this group.

In dry catchments (P < PET) precipitation seasonality has a slight negative influence on excess rainfall floods. Higher precipitation seasonality values indicate a precipitation peak in summer/warm season and lower a peak in winter/cold season (Addor et al., 2017; Fang & Shen, 2017; Woods, 2009). The colder temperatures prohibit the drying out of soils during the peak rainfall season leading to saturated conditions. In catchments with a precipitation peak in winter we therefore see more excess rainfall floods. Archer (1981) found for the humid catchments in Great Britain that soil moisture deficits in the summer prevent flooding despite rainfall events with high intensity. Instead, flooding is more common in the winter, when soils are saturated. We can conclude that in catchments where the seasonal precipitation peak coincides with low evaporative demand, excess rainfall floods would be even more likely. The effect of precipitation seasonality on excess rainfall floods can be seen for wet and particularly for dry catchments. This is confirmed by the accumulated local effects. Although influence in the distribution comparison is minor, ALE confirms that it is strongest in comparison with the other attributes. In snow catchments precipitation seasonality is less influential. With temperature below freezing in winter, winter seasonal precipitation will instead contribute toward snowpack (Woods, 2009) and not cause floods immediately.

Despite both methods pointing to an influence of elevation as an attribute, it is difficult to distinguish elevation from various other catchment attributes (Dingman, 1981; Merz & Blöschl, 2009), some of which are not included in the analysis. All high elevation catchments in the conterminous United States are located in the Rocky Mountains, Sierra Nevadas, and in the Appalachian Mountains. In general, mountainous catchments are steeper, have a higher fraction of bare soils, are smaller, receive more precipitation and due to a temperature gradient have a higher fraction of snow (Wohl, 2013). Elevation thus affects various aspects of flow behavior (Dingman, 1981). We can therefore not conclude that elevation is influential because it reflects a combination of attributes (fraction of snow, slope, catchment area, etc.) or if it is a proxy for attributes that are either not measured at all (drainage density, infiltration capacity), or measured with a high uncertainty (soil characteristics). The interaction between attributes is outside the scope of this paper (with an exception for aridity and fraction of snow which define the different climate types). Therefore, we can only take elevation as a proxy for mountainous catchments, indicating that in mountainous catchments flood generating processes are more likely to be short rainfall floods and snowmelt than excess rainfall.

If and how forest and vegetation in general affect flood characteristics is widely debated (Bradshaw et al., 2007; Bruijnzeel, 2004; Calder & Aylward, 2006; Rogger et al., 2017) and varies for different processes (Table 1). However, several studies showed that runoff processes can be influenced by land use, for example by decreasing quick rainfall runoff as well as snowpack volume. This affects particularly arid/semiarid and snow-influenced areas (Lull & Reinhart, 1972; Osterkamp & Friedman, 2000; Pariente, 2002; Shafer et al., 2007; Storck et al., 2002; Zhang et al., 2011). The results from the distribution comparison agree with findings in the literature. The comparison shows a stronger influence of vegetation on excess rainfall floods (positive) and short/long rain floods (negative) with increasing vegetation compared to wet catchments. Shafer et al. (2007) notes for desert areas that vegetation increases infiltration capacity of the soil. Additionally, in arid to semiarid areas in Israel shrubs will locally increase soil water retention (Pariente, 2002), this will reduce quick runoff generation leading to less short rain floods. Zhang et al. (2011) describe for the subhumid east Qinghai-Tibet Plateau that forest vegetation in comparison to shrubs increase water retention of the soil. Merz and Blöschl (2003) describe for Austria, that an increased water retention requires larger rainfall amounts or previous saturation to cause flood sized runoff events. This explains why vegetation that increases water retention increases the proportion of excess rainfall floods and decreases short rainfall/long rainfall floods.

However, the distribution comparison approach is sensitive to correlated attributes. In the CAMELS data, the correlation between vegetation and climate attributes is strongest for snow dominated catchments (Figure S10). The effect we are seeing could therefore just be due to correlation and only climate but not vegetation attributes influence flood process distribution. Yet, the ALE approach which is unbiased to correlated features sees a minor influence of vegetation as well. So does vegetation play a role or not? While vegetation attributes do have some influence on flood processes, the influence is small if compared to climate attributes (which is the result ALE shows). This has been noted for other flow behavior as well. Jehn et al. (2020) notice that hydrologic catchment cluster are most strongly shaped by climate but that vegetation and soil information play a role as well. Similar conclusions have been reached by Berghuijs, Sivapalan, et al. (2014) for similarity in a seasonal water balance. Based on their experience it stands to reason that in study areas with very similar climate, vegetation will determine mix of flood generating processes.

In contrast to vegetation, we did not find area, slope, and shape to be influential. Although these attributes have been shown to influence flood magnitude (Gaál et al., 2012; Murthy, 2002; Padi et al., 2011; A. Smith et al., 2015), we detected no influence on flood process. For the lack of effect of area on snowmelt floods, this has been documented in previous studies (Merz & Blöschl, 2003). A possible explanation that the effect of area on short rain floods was not detected as expected (Table 1) could be that the effect is strongest for flash floods occurring within a few minutes/hours. These types of floods were outside the temporal resolution of the data.

4.2. Predictions in Space Using Random Forest

In addition to evaluating attribute influence, we were able to show that a random forest model is able to predict the spatial distribution of each flood generating process. The accuracy of the prediction varies between climates, and especially in a wet climate several processes are not as well predicted. A possible explanation might be that with excess rainfall being the most common process in these regions, any other processes can be related less to catchment or climate attributes and more to extreme weather events occurring outside the regular flood season, for example severe thunderstorms or tropical cyclones (J. A. Smith et al., 2018). Stein et al. (2019) highlighted that in the southeastern United States, several catchments have a different dominant flood generating process than they do for the most extreme flood event in the time series. These single event contributions from different processes are difficult for a random forest model to predict based on stationary input attributes. Therefore, while the overarching prediction accuracy might be high, the possible uncertainty of extreme flood generating processes should be kept in mind.

4.3. Limitations

The simple snowmelt routine and simplified definition of rain-on-snow floods used here can lead to some misclassifications of snowmelt and rain-on-snow floods on an event basis (Stein et al., 2019). However, Stein et al. (2019) also mention that relative contribution of each flood process within a catchment seemed to be classified correctly. Since this is the information used in this analysis, we assume that individual event misclassification does not have a large impact on the overall result.

We recognize that environmental data is prone to uncertainties. Soil data, in particular, relies on uncertain interpolation of point measurements over space and depth (Addor et al., 2017; Merz & Blöschl, 2009; D. A. Miller & White, 1998). This uncertainty might be a possible explanation for having found little influence of soil attributes on flood processes, despite the influence of soil on storage capacity (Section 1.1, Table 1). Similar uncertainties can be found in large scale geology data sets, especially since the CAMELS data set uses information from global geology data sets (Addor et al., 2017). The evaluated attributes were all taken from the CAMELS data set (Addor et al., 2017) as a consolidated source. Further studies might want to take additional and nonstationary catchment attributes into account. Possible suggestions for additional stationary attributes are drainage density, wetland area, slope aspect, and urbanized areas. Possible suggestions for nonstationary attributes are forest cover, leaf area index, green vegetation fraction, annual precipitation and annual fraction of snow. Furthermore, the data set includes mostly small headwater catchments. It is possible that the conclusions might change if larger catchments are taken into account (FAO, 2002). It is likely that influence of land use attributes, such as vegetation, would be even less noticeable in larger catchments due to confounding relationships with other attributes (Rogger et al., 2017).

5. Conclusions

We employed a statistics-based approach (comparing empirical distribution functions) and a machine learning approach (random forest model combined with accumulated local effects) to evaluate which catchment attributes influence flood generating processes. ALE have only recently be introduced to the field of hydrology (Konapala et al., 2020). The careful combination of the two approaches combined with hydrological theory allowed us to draw conclusions how catchment attributes influence flood generating processes. We recommend that the hydrological community makes more use of the novel methods of interpretable machine learning. Particularly, since correlated attributes are a common occurrence in hydrologic catchment data.

In regard to flood generating processes we found that climatic attributes, such as fraction of snow, aridity, precipitation seasonality and mean precipitation have the strongest influence within the catchment and within space. In comparison, vegetation plays a minor role. This confirmed previous findings that flow behavior across climates is most strongly influenced by climate attributes (Addor et al., 2018; Jehn et al., 2020). In snow influenced catchments, elevation as a proxy for one or more attributes is influential in predicting flood processes across space. Neither of the methods we used found soil or geologic attributes to be influential. This might be due to limitations in data quality or attribute selection for both groups.

With the available catchment attribute information, the mix of flood generating processes can be predicted with relatively high accuracy. A prediction of processes for ungauged catchments is therefore possible, although climate-dependent uncertainties should be taken into account.

Further studies are necessary to evaluate the implication of these findings in regard to changes in climate and land use. Changes in flood magnitude and frequency have been observed, yet direction and magnitude of the trends are not homogeneous (Blöschl et al., 2019; Gudmundsson et al., 2019; Mallakpour & Villarini, 2015; Sharma et al., 2018; Wasko & Nathan, 2019). The results of this study can give an indication why: not all flood processes are influenced by the same climate attributes and the influence of catchment attributes on flood generating processes varies between different climates.

Data Availability Statement

The Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) data set (Addor et al., 2017; Newman et al., 2015) is freely available at https://ral.ucar.edu/solutions/products/camels. The National Soil Geographic Database (NATSGO) used to calculate available soil water storage (Soil Survey Staff, 2019) is freely accessible at https://nrcs.app.box.com/v/soils. Processed NATSGO data are provided in a table in the supporting information. The code needed to replicate this work can be found here: http://doi.org/10.5281/zenodo.4277642.

Acknowledgments
This work was funded as part of the Water Informatics Science and Engineering Centre for Doctoral Training (WISE CDT) under a grant from the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/L016214/1. We thank Maria Xenochristou for her helpful advice on random forest and interpretable machine learning. We are grateful to the editors, Larisa Tarasova, Julia Hall, and two anonymous reviewers for their constructive comments.

References

Acreman, M., & Holden, J. (2013). How wetlands affect floods. *Wetlands*, *33*(5), 773–786. https://doi.org/10.1007/s13157-013-0473-2

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, *54*(11), 8792–8812. https://doi.org/10.1029/2018WR022606

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Anchang, J. Y., Prihodko, L., Ji, W., Kumar, S. S., Ross, C. W., Yu, Q., et al. (2020). Toward operational mapping of woody canopy cover in tropical savannas using Google Earth Engine. *Frontiers in Environmental Science*, *8*, 4. https://doi.org/10.3389/fenvs.2020.00004

Apley, D. W., & Zhu, J. (2016). *Visualizing the effects of predictor variables in black box supervised learning models*. Retrieved from http://arxiv.org/abs/1612.08468

Archer, D. R. (1981). Seasonality of flooding and the assessment of seasonal flood risk. *Proceedings of the Institution of Civil Engineers*, *71*(4), 1023–1035.

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

Berghuijs, W. R., Sivapalan, M., Woods, R. A., & Savenije, H. H. G. (2014). Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales. *Water Resources Research*, *50*(7), 5638–5661. https://doi.org/10.1002/2014WR015692

Berghuijs, W. R., Woods, R. A., & Hrachowitz, M. (2014). A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change*, *4*(7), 583–586.

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

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B, et al. (2017). Changing climate shifts of European floods. *Science*, *357*(6351), 588–590. https://doi.org/10.1126/science.aan2506

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., et al. (2019). Changing climate both increases and decreases European river floods. *Nature*, *573*(7772), 108–111. https://doi.org/10.1038/s41586-019-1495-6

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., & Savenije, H. H. G. (2013). Runoff prediction in ungauged basins: Synthesis across processes, places and scales. *Eos*, *95*, 22. https://doi.org/10.1017/CBO9781139235761

Booker, D. J., & Woods, R. A. (2014). Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. *Journal of Hydrology*, *508*, 227–239. https://doi.org/10.1016/j.jhydrol.2013.11.007

Bradshaw, C. J., Sodhi, N. S., Peh, K. S., & Brook, B. W. (2007). Global evidence that deforestation amplifies flood risk and severity in the developing world. *Global Change Biology*, *13*(11), 2379–2395. https://doi.org/10.1111/j.1365-2486.2007.01446.x

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown, S. C., Wells, K., RoyDufresne, E., Campbell, S., Cooke, B., Cox, T., & Fordham, D. A (2020). Models of spatiotemporal variation in rabbit abundance reveal management hotspots for an invasive species. *Ecological Applications*, *30*, e02083.

Bruijnzeel, L. A. (2004). Hydrological functions of tropical forests: Not seeing the soil for the trees? *Agriculture, Ecosystems & Environment*, *104*, 185–228. https://doi.org/10.1016/j.agee.2004.01.015

Bullock, A., & Acreman, M (2003). The role of wetlands in the hydrological cycle. *Hydrology and Earth System Sciences*, *7*(3), 358–389. https://doi.org/10.5194/hess-7-358-2003

Calder, I. R., & Aylward, B. (2006). Forest and floods. *Water International*, *31*(1), 87–99. https://doi.org/10.1080/02508060608691918

Cantón, Y., Domingo, F., Solé-Benet, A., & Puigdefábregas, J. (2002). Influence of soil-surface types on the overall runoff of the Tabernas badlands (south-east Spain): Field data and model approaches. *Hydrological Processes*, *16*(13), 2621–2643. https://doi.org/10.1002/hyp.1052

Chang, H., Johnson, G., Hinkley, T., & Jung, I.-W. (2014). Spatial analysis of annual runoff ratios and their variability across the contiguous U.S. *Journal of Hydrology*, *511*, 387–402. https://doi.org/10.1016/J.JHYDROL.2014.01.066

Clark, M. P., Schaefli, B., Schymanski, S. J., Samaniego, L., Luce, C., Jackson, B. M., et al. (2016). Improving the theoretical underpinnings of process-based hydrologic models. *Water Resources Research*, *52*(3), 2350–2365. https://doi.org/10.1002/2015WR017910

Cunnane, C. (1979). A note on the Poisson assumption in partial duration series models. *Water Resources Research*, *4*(2), 489–494. https://doi.org/10.1029/WR015i002p00489

Davenport, F. V., Herrera-Estrada, J. E., Burke, M., & Diffenbaugh, N. S. (2020). Flood size increases nonlinearly across the western United States in response to lower snow-precipitation ratios. *Water Resources Research*, *56*, e2019WR025571. https://doi.org/10.1029/2019WR025571

David, V., & Davidova, T. (2014). Methodology for flood frequency estimations in small catchments. *Natural Hazards and Earth System Sciences*, *14*(10), 2655–2669. https://doi.org/10.5194/nhess-14-2655-2014

Degenhardt, F., Seifert, S., & Szymczak, S. (2019). Evaluation of variable selection methods for random forests and omics data sets. *Briefings in Bioinformatics*, *20*(2), 492–503. https://doi.org/10.1093/bib/bbx124

Diezig, R., & Weingartner, R. (2007). Hochwasserprozesstypen in der Schweiz. *Wasser und Abfall*, *4*(1), 18–26.

Dingman, S. L. (1981). Elevation: A major influence on the hydrology of New Hampshire and Vermont, USA / L'altitude exerce une influence importante sur l'hydrologie du New Hampshire et du Vermont, Etats-Unis. *Hydrological Sciences Bulletin*, *26*(4), 399–413. https://doi.org/10.1080/02626668109490904

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., et al. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Eaton, B., Church, M., & Ham, D. (2002). Scaling and regionalization of flood flows in British Columbia, Canada. *Hydrological Processes*, *16*(16), 3245–3263. https://doi.org/10.1002/hyp.1100

Falkenmark, M., & Chapman, T. (1989). *Comparative hydrology: An ecological approach to land and water resources*. Paris, France: The Unesco Press. ISBN: 9231025716.

Fang, K., & Shen, C. (2017). Full-flow-regime storage-streamflow correlation patterns provide insights into hydrologic functioning over the continental US. *Water Resources Research*, *53*(9), 8064–8083. https://doi.org/10.1002/2016WR020283

FAO. (2002). *Land-water linkages in rural watersheds electronic workshop. Conclusions and recommendations*. Food and Agriculture Organization of the United Nations.
Fischer, S., Schumann, A., & Schulte, M. (2016). Characterisation of seasonal flood types according to timescales in mixed probability distributions. *Journal of Hydrology*, *539*, 38–56.
French, R. H., & Miller, J. J. (2011). *Flood hazard identification and mitigation in semi- and arid environments*. World Scientific Publishing Co. https://doi.org/10.1142/8175
Freudiger, D., Kohn, I., Stahl, K., & Weiler, M. (2014). Large-scale analysis of changing frequencies of rain-on-snow events with flood-generation potential. *Hydrology and Earth System Sciences*, *18*(7), 2695–2709. https://doi.org/10.5194/hess-18-2695-2014
Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.2307/2699986
Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., & Blöschl, G. (2012). Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resources Research*, *48*, W04511. https://doi.org/10.1029/2011WR011509
Gudmundsson, L., Leonard, M., Do, H. X., Westra, S., & Seneviratne, S. I. (2019). Observed trends in global indicators of mean and extreme streamflow. *Geophysical Research Letters*, *46*(2), 756–766. https://doi.org/10.1029/2018GL079725
Institute of Hydrology (IoH). (1999). *Flood estimation handbook*. Wallingford, UK: IoH.
Jehn, F. U., Bestian, K., Breuer, L., Kraft, P., & Houska, T. (2020). Using hydrological and climatic catchment clusters to explore drivers of catchment behavior. *Hydrology and Earth System Sciences*, *24*(3), 1081–1100. https://doi.org/10.5194/hess-24-1081-2020
Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P. (2018). Spatial variation of the rainsnow temperature threshold across the Northern Hemisphere. *Nature Communications*, *9*(1), 1148. https://doi.org/10.1038/s41467-018-03629-7
Keller, L., Rössler, O., Martius, O., & Weingartner, R. (2018). Delineation of flood generating processes and their hydrological response. *Hydrological Processes*, *32*(2), 228–240. https://doi.org/10.1002/hyp.11407
Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, *42*, W03S04. https://doi.org/10.1029/2005WR004362
Kjeldsen, T. R. (2015). How reliable are design flood estimates in the UK? *Journal of Flood Risk Management*, *8*(3), 237–246.
Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resources Research*, *54*(7), 5088–5109. https://doi.org/10.1029/2018WR022913
Konapala, G., Kao, S. C., Painter, S. L., & Lu, D. (2020). Machine learning assisted hybrid models can improve streamflow simulation in diverse catchments across the conterminous US. *Environmental Research Letters*, *15*(10), 104022. https://doi.org/10.1088/1748-9326/aba927
Li, D., Lettenmaier, D. P., Margulis, S. A., & Andreadis, K. (2019). The role of rain-on-snow in flooding over the conterminous United States. *Water Resources Research*, *55*(11), 8492–8513. https://doi.org/10.1029/2019WR024950
Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.
Lull, H., & Reinhart, K. (1972). *Forests and floods in the eastern United States, Vol. 226*. US Northeastern Forest Experiment Station.
Mallakpour, I., & Villarini, G. (2015). The changing nature of flooding across the central United States. *Nature Climate Change*, *5*(3), 250–254. https://doi.org/10.1038/nclimate2516
Marks, D., Kimball, J., Tingey, D., & Link, T. (1998). The sensitivity of snowmelt processes to climate conditions and forest cover during rain-on-snow: A case study of the 1996 Pacific Northwest flood. *Hydrological Processes*, *12*(10–11), 1569–1587. https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11⟨1569::AID-HYP682⟩3.0.CO;2-L
Marks, D., Link, T., Winstral, A., & Garen, D. (2001). Simulating snowmelt processes during rain-on-snow over a semi-arid mountain basin. *Annals of Glaciology*, *32*, 195–202. https://doi.org/10.3189/172756401781819751
McCabe, G. J., Clark, M. P., & Hay, L. E. (2007). Rain-on-snow events in the western United States. *Bulletin of the American Meteorological Society*, *88*(3), 319–328. https://doi.org/10.1175/BAMS-88-3-319
McCartney, M. P., Neal, C., & Neal, M. (1998). Use of deuterium to understand runoff generation in a headwater catchment containing a dambo. *Hydrology and Earth System Sciences*, *2*(1), 65–76. https://doi.org/10.5194/hess-2-65-1998
Merz, R., & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, *39*, 1340. https://doi.org/10.1029/2002WR001952
Merz, R., & Blöschl, G. (2005). Flood frequency regionalisation spatial proximity vs. catchment attributes. *Journal of Hydrology*, *302*(1), 283–306.
Merz, R., & Blöschl, G. (2009). A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research*, *45*, W01405. https://doi.org/10.1029/2008WR007163
Merz, R., Blöschl, G., & Parajka, J. (2006). Spatio-temporal variability of event runoff coefficients. *Journal of Hydrology*, *331*(3–4), 591–604. https://doi.org/10.1016/J.JHYDROL.2006.06.008
Miller, D. (1964). *Interception processes during snow storms* (Technical Report). Res. Paper PSW-RP-18. Berkeley, CA: Pacific Southwest Forest & Range Experiment.
Miller, D. A., & White, R. A. (1998). A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions*, *2*(2), 1–26. https://doi.org/10.1175/1087-3562(1998)002⟨0001:acusms⟩2.3.co;2
Molnar, C. (2019). *Interpretable machine learning – A guide for making black box models explainable*. Retrieved from https://christophm.github.io/interpretable-ml-book/
Molnar, C., Bischl, B., & Casalicchio, G (2018). iml: An R package for interpretable machine learning. *JOSS*, *3*(26), 786. https://doi.org/10.21105/joss.00786
Murthy, C. (2002). *Water resources engineering: Principles and practice*. Delhi: New Age International (P) Limited.
Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., et al. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. *Nature Climate Change*, *8*(9), 808–812. https://doi.org/10.1038/s41558-018-0236-4
Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A, et al. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, *19*, 209–223. https://doi.org/10.5194/hess-19-209-2015
Osterkamp, W. R., & Friedman, J. M. (2000). The disparity between extreme rainfall events and rare floods – With emphasis on the semi-arid American West. *Hydrological Processes*, *14*(16–17), 2817–2829. https://doi.org/10.1002/1099-1085(200011/12)14:16/17⟨2817::AID-HYP121⟩3.0.CO;2-B
Padi, P. T., Baldassarre, G. D., & Castellarin, A. (2011). Floodplain management in Africa: Large scale analysis of flood data. *Physics and Chemistry of the Earth*, *36*(7-8), 292–298. https://doi.org/10.1016/j.pce.2011.02.002

Pariente, S (2002). Spatial patterns of soil moisture as affected by shrubs, in different climatic conditions. *Environmental Monitoring and Assessment*, *73*(3), 237–251. https://doi.org/10.1023/A:1013119405441

Pianosi, F., & Wagener, T. (2015). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, *67*, 1–11. https://doi.org/10.1016/J.ENVSOFT.2015.01.004

Pitlick, J. (1994). Relation between peak flows, precipitation, and physiography for five mountainous regions in the western USA. *Journal of Hydrology*, *158*, 219–240. https://doi.org/10.1016/0022-1694(94)90055-8

Pomeroy, J. W., Fang, X., & Marks, D. G. (2016). The cold rain-on-snow event of June 2013 in the Canadian Rockies – Characteristics and diagnosis. *Hydrological Processes*, *30*(17), 2899–2914. https://doi.org/10.1002/hyp.10905

Ries, F., Schmidt, S., Sauter, M., & Lange, J. (2017). Controls on runoff generation along a steep climatic gradient in the Eastern Mediterranean. *Journal of Hydrology: Regional Studies*, *9*, 18–33. https://doi.org/10.1016/j.ejrh.2016.11.001

Rogger, M., Agnoletti, M., Alaoui, A., Bathurst, J. C., Bodner, G., Borga, M., et al. (2017). Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research. *Water Resources Research*, *53*, 5209–5219. https://doi.org/10.1002/2017WR020723

Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., et al. (2012). Step changes in the flood frequency curve: Process controls. *Water Resources Research*, *48*, W05544. https://doi.org/10.1029/2011WR011187

Rosbjerg, D., Blöschl, G., Burn, D. H., Castellarin, A., Croke, B., Di Baldassarre, G., et al. (2013). Prediction of floods in ungauged basins. In G. Blöschl (Ed.), *Runoff prediction in ungauged basins: Synthesis across processes, places and scales* (pp. 189–226). Cambridge University Press.

Ryan, J. A., & Ulrich, J. M. (2019). *quantmod: Quantitative financial modelling framework*. Retrieved from https://cran.r-project.org/package=quantmod

Schumm, S. (1956). *Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey*. Geological Society of America Bulletin. http://gsabulletin.gsapubs.org/content/67/5/597.short

Shafer, D. S., Young, M. H., Zitzer, S. F., Caldwell, T. G., & McDonald, E. V. (2007). Impacts of interrelated biotic and abiotic processes during the past 125 000 years of landscape evolution in the Northern Mojave Desert, Nevada, USA. *Journal of Arid Environments*, *69*(4), 633–657. https://doi.org/10.1016/j.jaridenv.2006.11.011

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If precipitation extremes are increasing, why aren't floods? *Water Resources Research*, *54*, 8545–8551. https://doi.org/10.1029/2018WR023749

Sikorska, A. E., Viviroli, D., & Seibert, J. (2015). Flood-type classification in mountainous catchments using crisp and fuzzy decision trees. *Water Resources Research*, *51*(10), 7959–7976. https://doi.org/10.1002/2015WR017326

Singh, P., Spitzbart, G., Hübl, H., & Weinmeister, H. (1998). The role of snowpack in producing floods under heavy rainfall. In K. Kovar (Ed.), *Hydrology, water resources and ecology in headwaters* (pp. 89–95). IAHS Publ.

Smith, A., Sampson, C., & Bates, P. D. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, *51*(1), 539–553. https://doi.org/10.1002/2014WR015814

Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L., & Bates, P. (2018). Strange floods: The upper tail of flood peaks in the United States. *Water Resources Research*, *54*(9), 6510–6542. https://doi.org/10.1029/2018WR022539

Sohrt, J., Ries, F., Sauter, M., & Lange, J. (2014). Significance of preferential flow at the rock soil interface in a semi-arid karst environment. *Catena*, *123*, 1–10. https://doi.org/10.1016/j.catena.2014.07.003

Soil Survey Staff. (2019). *Gridded National Soil Survey Geographic (gNATSGO) Database for the Conterminous United States*. United States Department of Agriculture, Natural Resources Conservation Service.

Stein, L., Pianosi, F., & Woods, R. A. (2019). Event based classification for global study of river flood generating processes. *Hydrological Processes*, *34*, 1514–1529. https://doi.org/10.1002/hyp.13678

Storck, P., Lettenmaier, D. P., & Bolton, S. M. (2002). Measurement of snow interception and canopy effects on snow accumulation and melt in a mountainous maritime climate, Oregon, United States. *Water Resources Research*, *38*(11), 5-1–5-16. https://doi.org/10.1029/2002wr001281

Sui, J., & Koehler, G. (2001). Rain-on-snow induced flood events in Southern Germany. *Journal of Hydrology*, *252*(1–4), 205–220. https://doi.org/10.1016/S0022-1694(01)00460-7

Tarasova, L., Basso, S., Wendi, D., Viglione, A., Kumar, R., & Merz, R. (2020). A process based framework to characterize and classify runoff events: The event typology of Germany. *Water Resources Research*, *56*(5), e2019WR026951. https://doi.org/10.1029/2019wr026951

Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., et al. (2019). Causative classification of river flood events. *Wiley Interdisciplinary Reviews: Water*, *6*, e1353. https://doi.org/10.1002/wat2.1353

Tetzlaff, D., Seibert, J., McGuire, K. J., Laudon, H., Burns, D. A., Dunn, S. M., & Soulsby, C. (2009). How does landscape structure influence catchment transit time across different geomorphic provinces? *Hydrological Processes*, *23*(6), 945–953. https://doi.org/10.1002/hyp.7240

Toloşi, L., & Lengauer, T. (2011). Classification with correlated features: Unreliability of feature ranking and solutions. *Bioinformatics*, *27*(14), 1986–1994. https://doi.org/10.1093/bioinformatics/btr300

Tooth, S. (2000). Process, form and change in dryland rivers: A review of recent research. *Earth-Science Reviews*, *51*(1–4), 67–107. https://doi.org/10.1016/S0012-8252(00)00014-3

van Dijk, A. I., van Noordwijk, M., Calder, I. R., Bruijnzeel, S. L., Schellekens, J. A., & Chappell, N. A. (2009). Forest-flood relation still tenuous – Comment on 'Global evidence that deforestation amplifies flood risk and severity in the developing world' by C. J. A. Bradshaw, N. S. Sodi, K. S.-H. Peh and B. W. Brook. *Global Change Biology*, *15*, 110–115. https://doi.org/10.1111/j.1365-2486.2008.01708.x

Viglione, A., & Blöschl, G. (2009). On the role of storm duration in the mapping of rainfall to flood return periods. *Hydrology and Earth System Sciences*, *13*(2), 205–216. https://doi.org/10.5194/hess-13-205-2009

Würzer, S., Jonas, T., Wever, N., & Lehning, M. (2016). Influence of initial snowpack properties on runoff formation during rain-on-snow events. *Journal of Hydrometeorology*, *17*(6), 1801–1815. https://doi.org/10.1175/JHM-D-15-0181.1

Ward, R. C. (1978). *Floods – A geographical perspective*. The Macmillan Press.

Wasko, C., & Nathan, R. (2019). Influence of changes in rainfall and soil moisture on trends in flooding. *Journal of Hydrology*, *8*, 432–441. https://doi.org/10.1016/J.JHYDROL.2019.05.054

Weingartner, R., Barbena, M., & Spreafico, M. (2003). Floods in mountain areas – An overview based on examples from Switzerland. *Journal of Hydrology*, *282*(1–4), 10–24. https://doi.org/10.1016/S0022-1694(03)00249-X

Wohl, E (2013). *Mountain rivers revisited, Vol. 19*. Washington, DC: American Geophysical Union. https://doi.org/10.1029/WM019

Wood, E. F., Sivapalan, M., & Beven, K. J. (1990). Similarity and scale in catchment storm response. *Reviews of Geophysics*, *2*(1), 1. https://doi.org/10.1029/RG028i001p00001

Woods, R. A. (2009). Analytical model of seasonal climate impacts on snow hydrology: Continuous snowpacks. *Advances in Water Resources*, *32*(10), 1465–1481. https://doi.org/10.1016/j.advwatres.2009.06.011

Zhang, W., An, S., Xu, Z., Cui, J., & Xu, Q. (2011). The impact of vegetation and soil on runoff regulation in headwater streams on the east Qinghai-Tibet Plateau, China. *Catena*, *87*(2), 182–189. https://doi.org/10.1016/j.catena.2011.05.020
Zhao, Q., & Hastie, T. (2019). Causal interpretations of black-box models. *Journal of Business & Economic Statistics*, *39*, 272–281. https://doi.org/10.1080/07350015.2019.1624293

References From the Supporting Information

Doshi-Velez, F., & Kim, B. (2018). Considerations for evaluation and generalization in interpretable machine learning. In H. Escalante et al. (Eds.), *Explainable and Interpretable Models in Computer Vision and Machine Learning. The Springer Series on Challenges in Machine Learning*. Cham: Springer. https://doi.org/10.1007/978-3-319-98131-4_1
Pearl, J. (1993). Comment: Graphical models, causality and intervention. *Statistical Science*, *8*(3), 266–269. https://doi.org/10.1214/ss/1177010894