Superfloods in the Western U.S.

Emilie Tarouilly[1], Dongyue Li[1], and Dennis P. Lettenmaier[2,1]

[1] Department of Civil and Environmental Engineering, University of California Los Angeles

[2] Department of Geography, University of California Los Angeles

Corresponding author: Dennis Lettenmaier (dlettenm@ucla.edu)

Key Points:

- We identify superfloods in the Western U.S. as historical floods of very large magnitudes at multiple stations.
- Superfloods fall into three categories: winter (Atmospheric Rivers), spring snowmelt (some with rainfall) and summer (tropical storms).
- Superfloods are caused by both the extreme nature of their primary driver (e.g. rainfall amounts) and unusual combinations of drivers.

Abstract

We examine the characteristics of extremely severe and extensive floods across the Western U.S., which we term superfloods. We develop a system to score each day from 1950 to 2010, across the conterminous U.S. west of the Continental Divide, according to the severity of flooding and the extent of the area affected on that day. In order to augment the available stream gauge data, we incorporate model output runoff from the Variable Infiltration Capacity (VIC) model aggregated to the HUC-8 basin level. The superfloods we identify include well-known events such as the Christmas floods of December 1964. Once the superfloods are identified, we use modeled soil moisture and SWE in combination with gridded observed precipitation and air temperature to identify superflood drivers. We find that the superfloods fall into three distinct categories: a) winter events, most of which are associated with Atmospheric Rivers (ARs), many enhanced by rain-on-snow; b) late spring floods associated with snowmelt, including some events with rainfall contributions; and (c) rainfall-driven warm season floods associated with tropical storms (primarily in the Southwest). We find that rainfall is the primary driver of superfloods, although snowmelt also contributes in rain-snow transition zones. Regardless of superflood characteristics, extreme conditions in combination are a common feature, including combinations of rainfall and snowmelt as well as successive rainstorms. Our results provide a regional view of flood processes and how they have combined to produce superfloods across the region over our 60-year study period.

1. Introduction

In the Western U.S., the seasonal and climate-driven variability of streamflow has led to reliance on a large number of dams and levees for both water supply and flood control. Reliable flood prediction and the accurate estimation of design floods are key to improving the performance and safety of this infrastructure (Bales et al., 2006; Dettinger et al., 2011). This in turn motivates process

understanding of flood generation mechanisms that can supplement statistically based risk estimation methods (Klemes, 1993; Merz et al., 2014; Sivapalan et al., 2003). Flood generation involves a complex interaction of meteorological inputs (including precipitation volumes, timing, and phase) and catchment conditions (e.g. antecedent soil moisture and snowpacks). These interactions challenge statistically based flood frequency approaches (Alila & Mtiraoui, 2002; Apipattanavis et al., 2010). Furthermore, for temperature-dependent processes such as snowmelt, there may be climate-driven long-term trends in the frequencies of the mechanisms. Addressing these challenges requires understanding flood drivers and the controls they exert on the relative importance of mechanisms (rainfall, snowmelt) and resulting flood dynamics. Individually, causes of floods in the Western U.S. are well understood. Rainfall generation (e.g. convective versus frontal) can be a key driver of flood characteristics due to different signatures (intensity, spatial extent) of associated precipitation (Ashley & Ashley, 2008). In the coastal Western U.S., most major floods are associated with stratiform precipitation in winter which is orographically enhanced (James & Houze, 2005). Many of the extreme orographically enhanced stratiform precipitation events that lead to flooding are now understood to be atmospheric rivers (e.g., Ralph et al., 2006; Leung & Qian, 2009). Precipitation inputs alone, however, are not a good predictor of floods (Berghuijs et al., 2016). Antecedent basin conditions can be at least as important as storm intensity and duration (Berthet et al., 2009; Brocca et al., 2008; Ivancic and Shaw, 2015; Cao et al., 2019). In the mountainous Western U.S., both clear sky snowmelt and rain-on-snow (ROS), which depend on altitude and warm temperatures during AR storms, can also be important contributors to flooding (Harr, 1981; Musselman et al., 2018, Li et al., 2019).

Comprehensive sets of indicators that classify complex flood drivers at the regional scale using a combination of precipitation inputs (amount, duration) and basin conditions (soil moisture, snowpack) have been developed e.g. in Canada (Buttle et al., 2016) and Europe (Parajka et al., 2010; Merz & Blöschl, 2003). For example, Merz & Blöschl (2003) categorize events into long-rain floods, short-rain floods, flash floods, ROS floods and snowmelt floods according to multiple indicators, including flood timing, storm duration, rainfall and snowmelt amounts, runoff response and spatial coherence. In the U.S. however, regional studies to date have only used the seasonal timing of floods as a proxy to infer dominant mechanisms (Berghuijs et al., 2016). Further work is required to refine this analysis by using hydrologic information (precipitation, snowpack) in addition to flood timing (Berghuijs et al., 2019). The spatial footprint of flood drivers and the streamflow response across multiple stream gauges has not yet been characterized, despite the fact that the U.S. has an excellent

stream gauge network. The Western U.S. is prone to large scale floods caused by successive storms that affect multiple basins. For example, Fang & Pomeroy (2016) and Pomeroy et al. (2016a, b) show that altitude-dependent snow accumulation and melt dynamics shaped the June 2013 flood in the Canadian Rockies. Detailed studies of this kind are valuable in determining the unusual succession of events that lead to extreme floods. Understanding flood dynamics will require detailed characterization of flood progression and extent, which would shed light on how large floods are produced at the regional scale.

Analysis of flood drivers has most often examined a range of high flows, typically annual maxima or peaks-over-threshold (including the work cited above: Merz & Bloschl, 2003, Berghuijs et al., 2016). Studies of exceptional floods, such as Dettinger & Ingram's (2013) who investigated low frequency (~200 years) "megafloods" in California, have focused on recurrence intervals rather than flood drivers. Other studies have characterized rare, extreme floods in terms of their hydroclimatic drivers but how they translate into basin level mechanisms has not been examined (Najibi et al., 2017; Steinschneider & Lall, 2016). Of the few studies that investigate the drivers of extreme floods, O'Connor & Costa (2003) define "large" floods as the largest annual peaks relative to drainage area and identify Pacific moisture and topography as the key controls on how large a flood can be in the U.S. West. Similarly, Smith et al. (2018) define "strange" floods as the largest peaks relative to the 10-year flood and find that floods 200 times that magnitude can be produced by both uncommon flood drivers for a given watershed and extremes of common drivers. Examining causes of the largest historical floods in the Western U.S. is a necessary step in understanding how large floods of different types can be generated.

We aim here to better quantify the spatial extent of flooding in relation to the distribution of flood drivers and topographic patterns at the continental scale. To do this, we define "superfloods", large-scale events that caused record or near-record flooding across a substantial portion of the Western U.S. We define superfloods according to their rare occurrence and large geographic footprint, which we determine without regard to mechanism, and then identify the underlying mechanisms that generated these extreme and extensive floods. This allows us to address the question of which processes (or combination of processes) dominate their generation, both seasonally and geographically.

2. Methods

2.1 Overview

Our study domain is the conterminous U.S. west of the Continental Divide. To identify the superfloods, we used a combination of USGS stream gage records from the GAGES (Geospatial Attributes of Gages for Evaluating Streamflow) data set (version 2) (Falcone et al., 2010), and gridded runoff from the Variable Infiltration Capacity (VIC) model (Liang et al., 1994) (at $1/16^{th}$ degree spatial resolution, aggregated to HUC-8 river basins) to achieve more complete coverage of the domain. Below, we describe how time series of both observed and modeled streamflow are constructed, compared, and merged into a single dataset to identify superfloods. Our general approach is based on ranks of the observed and modeled streamflow rather than the values themselves. Our motivation in doing so is that ranks provide a sense of the relative order of extreme flows (hence avoiding, for instance, issues of model bias) and to allow for comparison of events regionally at scales where topography and climate exert strong influences. Therefore, at each location (each stream gage or HUC-8 basin for which we developed model-simulated flows) we rank the values by magnitude in descending order (i.e. rank zero being the flood of record) over the study period. Then, we take a peaks-over-threshold (POT) approach and retain only ranks between 0-50 i.e. the 50 largest days in the record for a given location. While we retained the entire POT series for superflood identification, the independence of the resulting superfloods, identified by a scoring system that accounts for both flood magnitudes across the domain and the number of flooded locations, was ensured by requiring a minimum separation interval between superfloods, as in Cao et al. (2019) and Mondal and Mujumdar (2015). In our case, we chose an interval of nine days following the work of Fish et al (2019) on the persistence of atmospheric rivers. As a result, if any of the superfloods occurred within nine days of each other, we assumed that they were part of the same event and only the largest peak in the 9-day interval was retained as a superflood.

2.2 Data

2.2.1. Stream gage data

The USGS GAGES Version 2 (GAGES II) reference stations are identified by the USGS as being the "least disturbed" locations among the USGS network (Falcone et al., 2010). Any gages with contributing area smaller than one VIC grid cell (about 36 km^2) were excluded. This resulted in unaffected time series for a total of 221 basins across the study domain which, despite higher spatial

density in mountainous areas, are present at a lower density across the entire region (Figure 1a). The average size of the GAGES II basins is 522 km^2. Our study period is from 1950 to 2010, the period for which records exist for the majority of basins (Figure 1b).

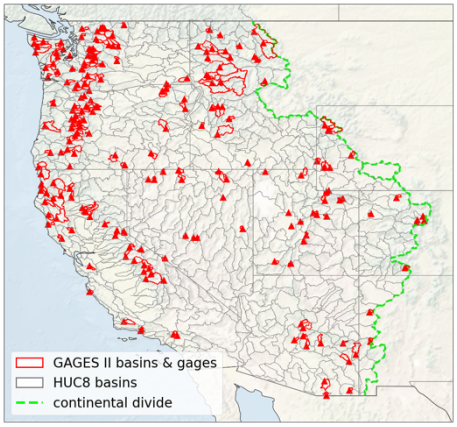


Figure 1a: Location of the 221 GAGES II stream gages (red triangles) and basin (red outlines) relative to HUC-8 basins (gray outlines).

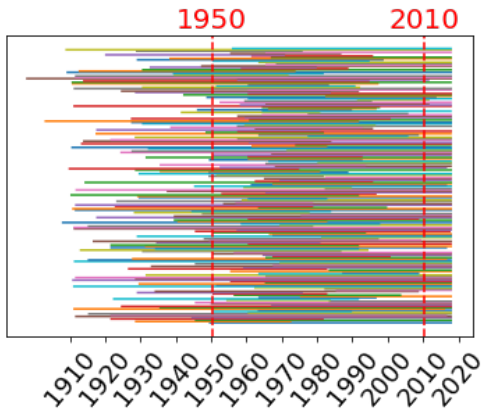


Figure 1b: Period of record for the 221 GAGES II streamflow records.

2.2.2. Modeled streamflow

Because the stream gage data are limited to 221 GAGES II locations, we created a streamflow proxy using model-derived runoff and baseflow, which we then merged with the stream gage observations. To do so, we used gridded daily runoff and baseflow at 1/16° (~6-km) spatial resolution, obtained from the VIC model runs of Li et al. (2019) for the period 1950-2010. We aggregated runoff and baseflow over all pixels within the drainage area of each HUC-8 basin to obtain modeled streamflow at their outlets. We performed comparisons between observed and modeled streamflow which showed modest differences in streamflow magnitudes (Supplement S1). Using modeled streamflow allowed us to extend the GAGES II dataset for the identification of superfloods over the entire region, unconstrained by the lack of stream gage locations.

2.2.3 Construction of the merged data set over the study domain

We merged the observed and modeled streamflows as follows. If a GAGES II basin is contained within a HUC-8 basin, we replaced the modeled streamflow ranks for that HUC-8 basin with the observed (GAGES II) streamflow ranks. Because the GAGES II basins are often smaller than the HUC-8 basins (see Figure 1a), in those cases where one HUC-8 basin contained more than one GAGES II basin, the streamflow rank for the largest of the GAGES II basins was retained. There were no cases where a GAGES II basin covered more than one HUC-8 basin. This produced a collection of streamflow time series for each HUC-8 unit across the Western U.S., with each time series taken either from observed GAGES II streamflow, if a gage exists within that HUC-8 basin, or VIC-derived streamflow. The resulting merged dataset was the basis of our identification of superfloods.

2.3 Superflood identification

The procedure for superflood identification, using the merged dataset, is as follows. Our goal was to select events that were associated with near-record flooding over a large part of the domain, i.e. to identify which days have high-ranking flows at a large number of locations. We did this by summing, on each day, the inverse of the ranks (so that larger flows have higher values) over all the basins that meet the POT criterion (rank < 50) on that day. This is what we call the "superflood score", which reflects both the intensity and extent of flooding for each day that has basins meeting the rank < 50 streamflow criteria. The POT threshold of rank 50 controls whether the superfloods we identify have larger extents by including more intermediate-ranking flows (e.g. if the threshold was 60) versus superfloods that have fewer, very high-ranking values (e.g. if the threshold was 40). Above rank 30, we found that the threshold value has very little impact on the superfloods we identified because it controls the characteristics of all superfloods, not their position relative to each other. Regardless of the threshold, basins that have POT flows with large ranks will add comparatively little to the score. The score will increase, however, with both a few very high-ranking flows and with a large number of locations. It will be particularly large on days that have many high-ranking flows. Once scores were calculated for all days in the record, they were rescaled to 0-1 relative to the largest score. These steps generated a scored list of superfloods, i.e. a scored list of days that have record or near-record flooding over multiple basins.

2.4 Impact of the datasets used on superflood identification

We compared the superfloods identified using the observed and modeled streamflows. We found that the two datasets score the superfloods similarly, but that the spatial extent of the superfloods is much better resolved after the inclusion of modeled streamflows.

We first compared superflood scores identified using observed and modeled streamflows at the GAGES II locations to examine the impact of any differences in the datasets at a given location (essentially controlling for the fact that the two datasets represent different numbers of basins). Figure 2a shows the correlation between the two datasets for all days in the record (before applying the POT threshold and independence criteria, blue dots) and for the subset used in the analysis (after applying the POT threshold and independence criteria, red dots). The superflood scores obtained using observed and modeled streamflows have a correlation of 0.7. In fact, 7 of the largest 12 superfloods are identified by both the observed and modeled datasets when the same locations (GAGES II basins) are used. This suggests that the use of modeled streamflows does not substantially alter the superflood scores compared to using observed streamflows.

Next, we assessed the influence of using the modeled streamflows to extend the number of available locations from the GAGES II basins only to the entire domain. We did so by comparing the superflood scores obtained using the modeled streamflows (in both cases) constructed over the GAGES II basins as well as the HUC-8 basins for the entire domain. Figure 2b shows that a number of superfloods identified using modeled streamflow over all HUC-8 basins score lower when only the GAGES II locations are considered (below the diagonal). The October 1972 flood (Figure 3) was a severe event caused by the remnants of hurricane Joanne (Stockton et al., 1989). Examination of the map shows this event was for the most part missed by the GAGES-II dataset due to low stream gage density in the region. In particular, the inclusion of modeled streamflow for the HUC-8 basins allowed us to identify late summer superfloods, typically in or near Arizona, which scored much lower using only the GAGES II dataset. We argue that these events likely would have scored similarly in the GAGES II-based analysis, had there been more stream gauges in this part of the domain.

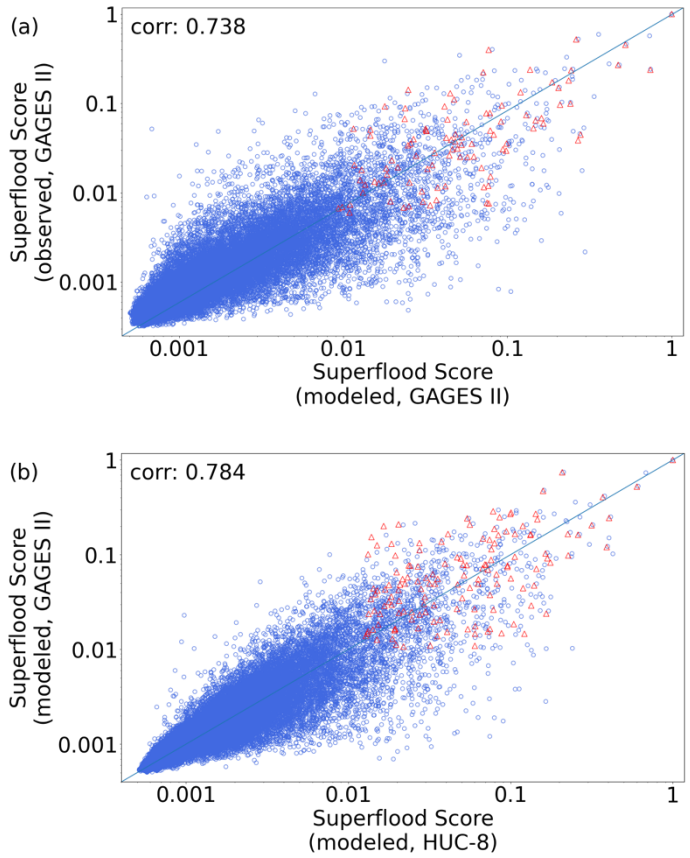


Figure 2: Comparison of the superflood scores obtained from different datasets. (a) comparison of the superflood scores obtained using the observed versus modeled streamflows at the same locations (GAGES II basins), (b) comparison of the superfloods scores obtained using the modeled streamflows at the GAGES II locations only versus the entire domain (using HUC-8 basins). The red dots show the superfloods after the rank < 50 and independence criteria were applied (i.e., those retained for subsequent analysis). Blue dots are superfloods identified using all days in the record.

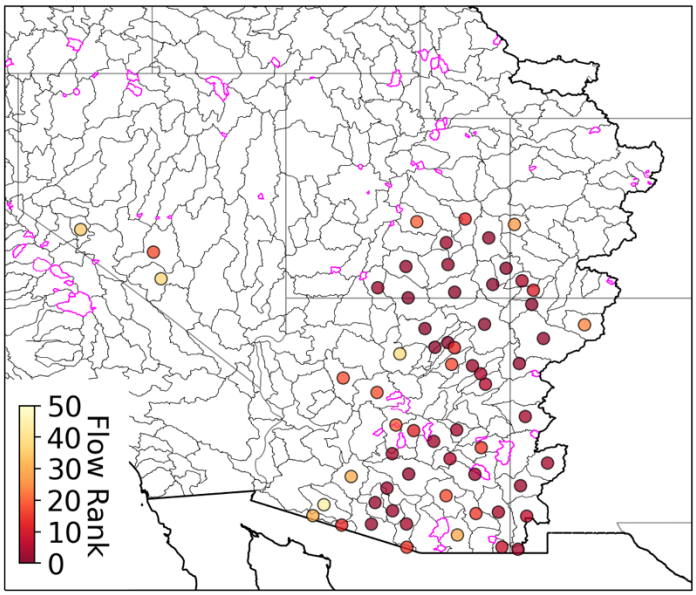


Figure 3: Flow ranks for each flooded basin for the October 1972 Arizona superflood. Gray outlines are HUC-8 basins, pink outlines are the available GAGES II basins. The colored dots show the streamflow ranks for those flooded basins (whether identified by observed or modeled streamflows), many of which were identified by modeled streamflow because they fall outside of the available GAGES II basins.

2.5 Categorization of flood drivers

We categorized the superfloods we identified based on multiple characteristics: (1) their season and region of occurrence, (2) their primary causes (rainfall, snowmelt or ROS), (3) whether they were driven by an AR, and (4) the shape of the flood peak. We carried out this analysis on both a large set of the 100 largest superfloods to identify clusters by season, location and drivers, and on a smaller group of the largest 12 events (first the largest 12 overall, and then the largest 12 in each season) in order to derive more precise understanding of spatial variability and temporal response during some particularly severe superfloods. We describe below how the necessary information was obtained.

We used gridded daily hydrologic variables from the Li et al. (2019) VIC model runs, in particular, snow water equivalent (SWE), snowmelt, rainfall, snowfall, air temperature and total column soil moisture. For each variable, we calculated percentiles relative to the same day of the year for all years in the record, in order to identify how extreme a given contributing mechanism was. We determined whether ROS occurred using the criteria of Li et al. (2019) (developed in Freudiger et al.,

2014), also applied at each pixel. The criteria are that at least 3 mm of rain falls on a snowpack with at least 10 mm SWE, and snowmelt constituted at least 20% of the sum of rainfall and snowmelt for the event. While AR intensity and duration have been shown to control flood severity (Ralph et al., 2019) along the U.S. West Coast, the impact of AR extent on flood footprint has not been examined as much. We obtained AR outlines and gridded integrated vapor transport, or IVT (and averaged the IVT over the extent of the AR that coincides with the study domain i.e. over land) from the AR catalog of Gershunov et al. (2017) and examined AR extent and intensity for those superfloods that coincided with an AR. Then, because the 3-day flood volume is a widely used designed criterion for hydrologic structures (Gangrade et al, 2018), we calculated the 3-day runoff volume and the runoff peak to 3-day volume ratio (PV ratio) for each basin to examine the shape of the flood peaks.

3. Results & Discussion

3.1 Largest superfloods

We first consider the traits of the most exceptional superfloods by examining the 12 largest events we identified overall (Figure 4). These floods are extraordinary in that they all have record or near-record flows (within the 50 largest for a given basin) over wide areas, some spanning much of the Western U.S. They all correspond to well-known events for which impacts, both hydrological and economic, have been well documented. Among them are the December 1964 and January 1997 floods along the north Pacific coast, and the October 1972 Hurricane Joanne floods in Arizona. Further information about the drivers and characteristics of each event, as described in the existing literature, is provided in Supplement S2.

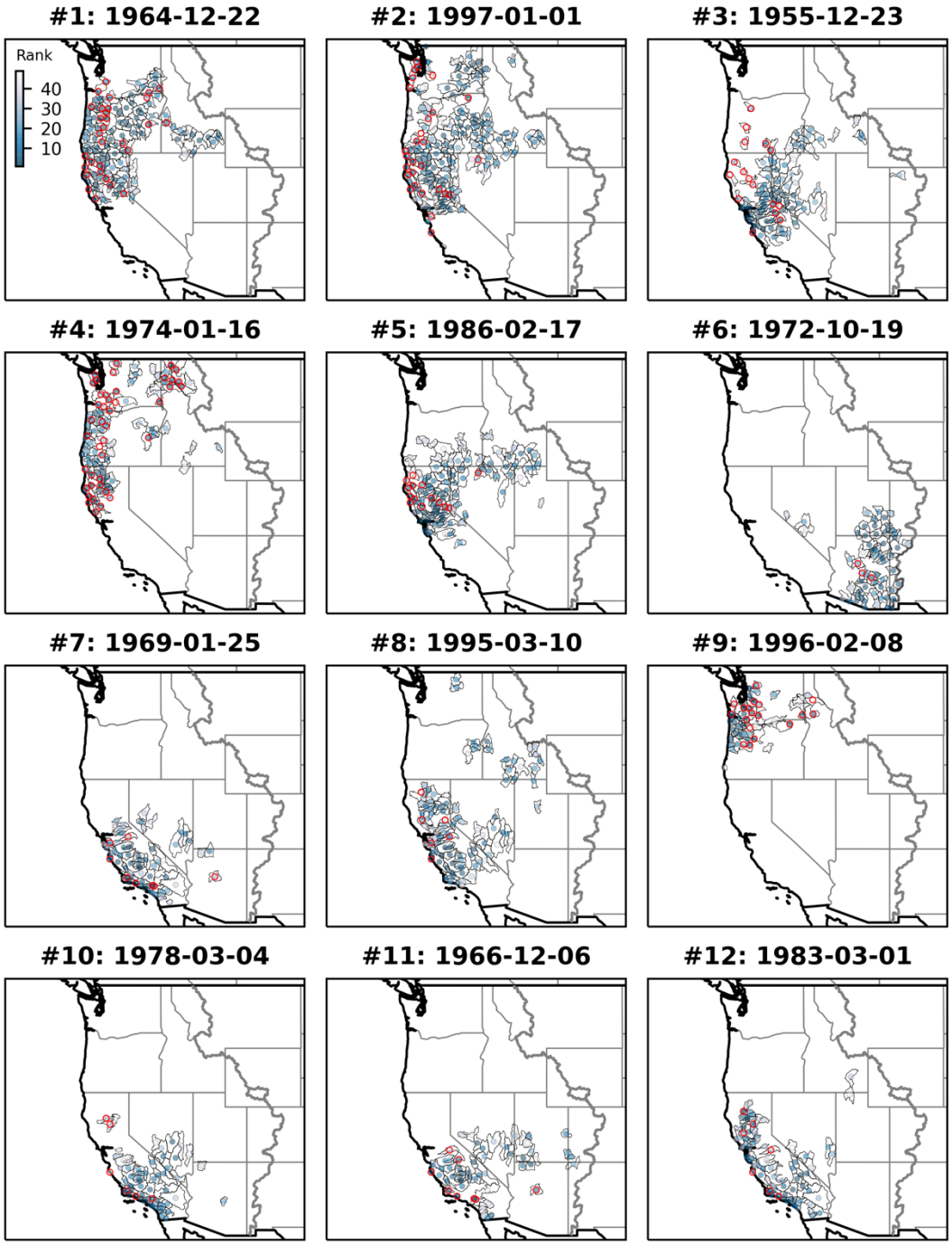


Figure 4: Maps of the largest superflood identified using the merged observed and modeled streamflows, in score order. Colored circles show the locations of the basins that make up the superflood, darker colors indicate larger flows (lower ranks). Basins circled in red are identified using stream gage data.

3.2 Location and seasonality of superfloods

Examination of the timing and location for the larger group of 100 superfloods reveals that they occur in markedly separate geographic locations that cluster at different times of year: November to March (winter), April to July (late spring to early summer, referred to as "spring" for brevity) and August to October (late summer to early fall, referred to as "summer"). Based on this grouping, 60 of the largest 100 superfloods occurred in winter, 30 in late spring and 10 in late summer. Figure 5 shows the number of times each basin is part of a superflood in a given season. The 60 winter events occurred primarily in the coastal portion of the domain, sometimes extending further into the interior (Figure 5a). Winter events typically have large footprints (median extent, estimated as the distance between the furthest affected basins: 1017 km), some (e.g., December 1964; January 1997) spanned all the West's mountain ranges (Sierra Nevada, Cascades and all the way to the Rockies). The 30 late spring events on the other hand occurred in the central part of the domain, either in the Rockies or in the interior arid basins and the eastern slopes of the Cascades and Sierra Nevada (Figure 5b), with a median extent of 1122 km. This is in agreement with the observations of Berghuijs et al. (2016) that flooding occurs in winter in coastal basins, early spring in inland California and late spring in the Rockies. Most of the 10 late summer events occurred in the southern portion of the domain, centering around Arizona (Figure 5c). Typically, summer events are more restricted geographically (median extent of 921 km). While winter and spring events affect similarly wide areas, winter floods affect large numbers of basins (median: 40), while spring and summer events affect 22.5 and 26.5 basins, respectively. This causes winter events to predominate among the 12 largest scoring events shown in the previous section (Figure 4), while no spring superflood is part of the largest 12 events overall and only one summer superfloods (October 1972) is #6 largest overall.

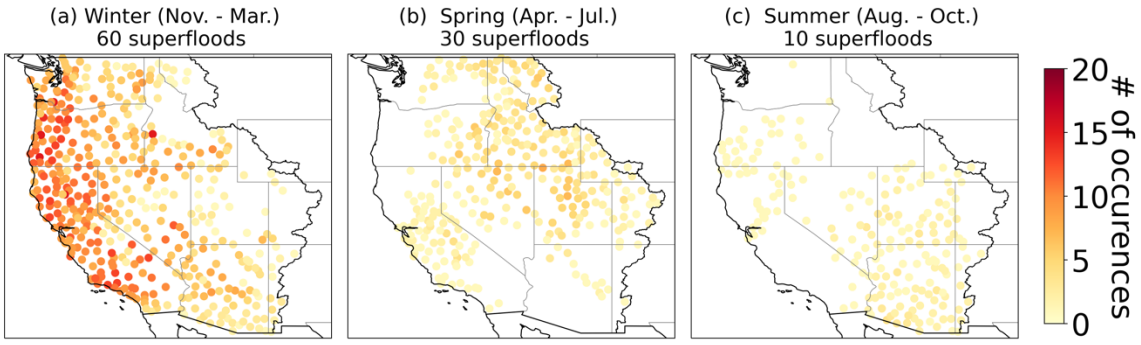
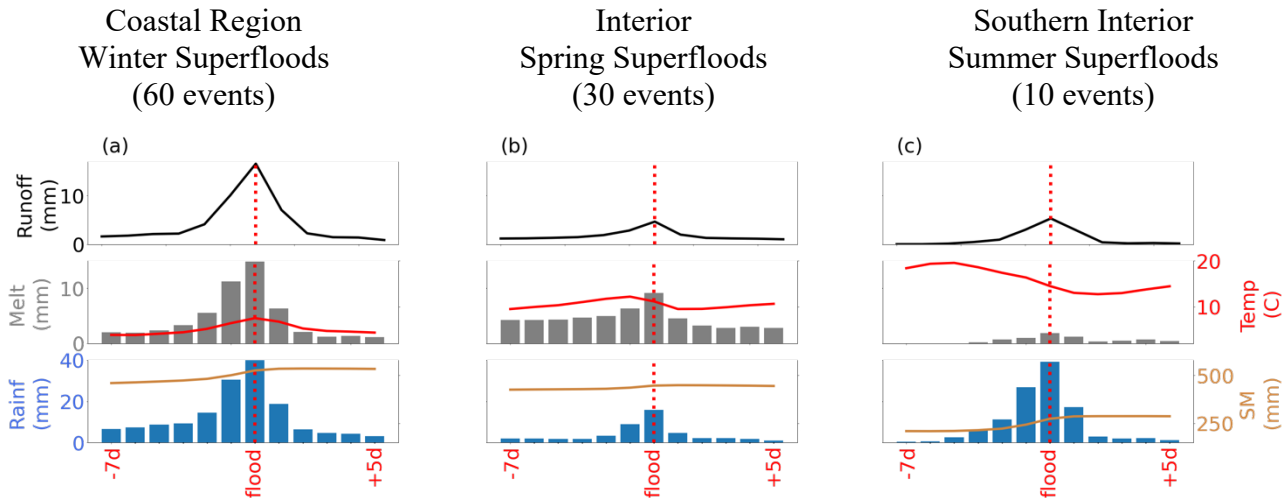


Figure 5: Spatial signatures of the largest 100 superfloods. For each season, each colored dot shows the number of times a given basin is affected by one of the largest 100 superfloods.

3.3 Superflood drivers

Flooding in the Western U.S. is known to result from large precipitation amounts that can be compounded by warm temperatures associated with rainfall and snowmelt where snow is present (Dettinger et al., 2011). Here, we seek to understand which of these drivers prevail across the 60 winter, 30 spring and 10 summer events that make up the largest 100 superfloods we have identified, particularly the relative roles of rainfall and snowmelt contributions by season and location. Figure 6 (top row) shows the progression of hydrologic variables (rainfall, SWE, temperature and snowmelt) and the runoff response aggregated over superfloods in each cluster (winter, spring and summer). The progression during the two weeks before and after the flood differs markedly across the three clusters. Winter superfloods (Figure 6a) have particularly large rainfall peaks. Temperatures are near freezing. Soil moisture is relatively high before the rainfall and flood peaks. These rain-driven winter floods have a much flashier runoff response than late spring floods, which are characterized by small amounts of rain in comparison to snowmelt contributions (Figure 6b). Temperatures are warm, typically around 10°C. Late summer events, on the other hand, mostly have no snowmelt and are almost exclusively caused by rainfall (Figure 6c). We examined the percentiles of the same variables (averaged over the three days leading up to the flood) to evaluate how extreme each contributing variable was, given the season and location (Figure 6, bottom row). Rainfall percentiles are equally extreme in both winter and summer superfloods (95th and 98th percentiles, respectively), while rainfall percentiles are slightly lower in spring (82nd percentile). Snowmelt amounts are equally extreme in winter and spring (93rd and 92nd percentiles) (Figure 6 d-f).



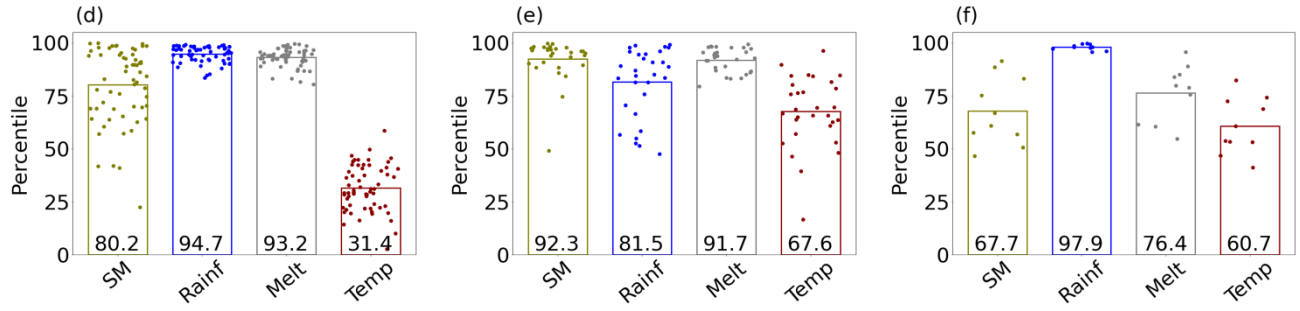


Figure 6: Superflood potential causative variables for the largest 100 events, averaged across superfloods in each season. Top row: timelines of SWE, rainfall, temperature and soil moisture leading up to the flood (averaged over all superfloods in a season). Bottom row: percentiles of the same variables over the three days leading up to the flood. Each dot corresponds to one superflood in the given season, the bar is the average across superfloods.

We specifically assessed the relative contributions of rainfall versus snowmelt in each cluster of superfloods (Figure 7). In all categories, most superfloods are dominated by rainfall (points below the diagonal), while some also have important snowmelt contributions (above the diagonal). Most winter superfloods are dominated by rainfall, with rainfall amounts typically several times larger than snowmelt amounts (Figure 7a). There are a few winter superfloods that have equal amounts of rainfall and snowmelt, but some of those appear to also be the superfloods that had the most ROS (darker colors). The particularly high scoring events (large bubble size) tend to have large contributions from both rainfall and snowmelt.

The influence of snowmelt relative to rainfall is more evident in spring superfloods, with 10 out of the 30 events having more snowmelt than rainfall (Figure 7b). Unlike winter events, all spring superfloods have at least some snowmelt, consistent with Jenning & Jones (2015) who state that precipitation is normally insufficient to cause severe floods in this region, except in June when combined with snowmelt. Summer superfloods, on the other hand are caused almost exclusively by rainfall (Figure 7c).

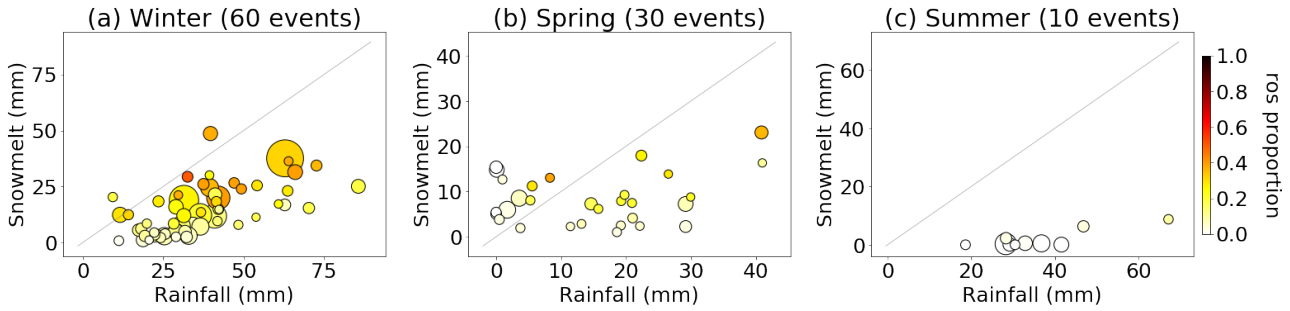


Figure 7: Snowmelt versus rainfall amounts by superflood category. Each point corresponds to one superflood. Rainfall and snowmelt are averaged across all pixels belonging to the basins that make up a superflood during the three days leading up to the flood peak. The diagonal indicates equal amounts of rainfall and snowmelt contributions. The color of the circle denotes the proportion of pixels that had ROS. The size of the circle reflects the superflood score.

3.4 Spatial patterns of flooding and their drivers

The fact that our approach to superflood identification considers basins that flood simultaneously across the Western U.S. domain enables us to examine the spatial pattern of these extensive floods in relation to the footprint of their drivers (rainfall, ROS, snowmelt). Here, we seek to understand how the spatial pattern of identified superfloods drivers controls the number and location of flooded basins. This section focuses on the 12 largest floods for each season (spring and summer floods would be under-represented if we considered only the 12 largest superfloods overall).

3.4.1 Winter Events

Figure 8 shows the hydrologic context maps for the largest 12 winter superfloods (see Supplement S3, for maps of the remainder of the 60 winter superfloods). The extent of any coincident ARs identified by Gershunov et al. (2017) is shown in purple shading and the location of flooded basins by the pink outlines. Of the 60 winter events, we find that all but four (November 1960, December 2005, February 1998 and February 2005) coincide with an AR over the flooded area.

While the winter superfloods that coincide with ARs (Figure 8) always affect the coastal West, their location from North to South and the degree of inland penetration is highly variable across superfloods. Peak rainfall amounts are typically large, sometimes over 200 mm/day locally. There is clear orographic enhancement of the rainfall along the coastal mountain ranges (e.g. large rainfall amounts along the Sierra Nevada in most superfloods). The location of the flooded basins coincides with but is often much wider than the areas of heaviest rainfall. The areas of heaviest rainfall are always found under the AR, so the North to South location of the superfloods typically depends on where the AR makes landfall.

We often observe ROS (in red on the maps) at the edges of the snow-covered areas (in green). In many cases, the superflood area is bounded by the edge of the snow-covered area. While winter superfloods are primarily caused by AR-driven rainfall which, together with topography, exerts a strong control on location and extent of rainfall peaks and the most severe flooding, snowmelt typically limits flood extent in areas not conducive to snowmelt, but often augments flooding at the edges of the superflood domain.

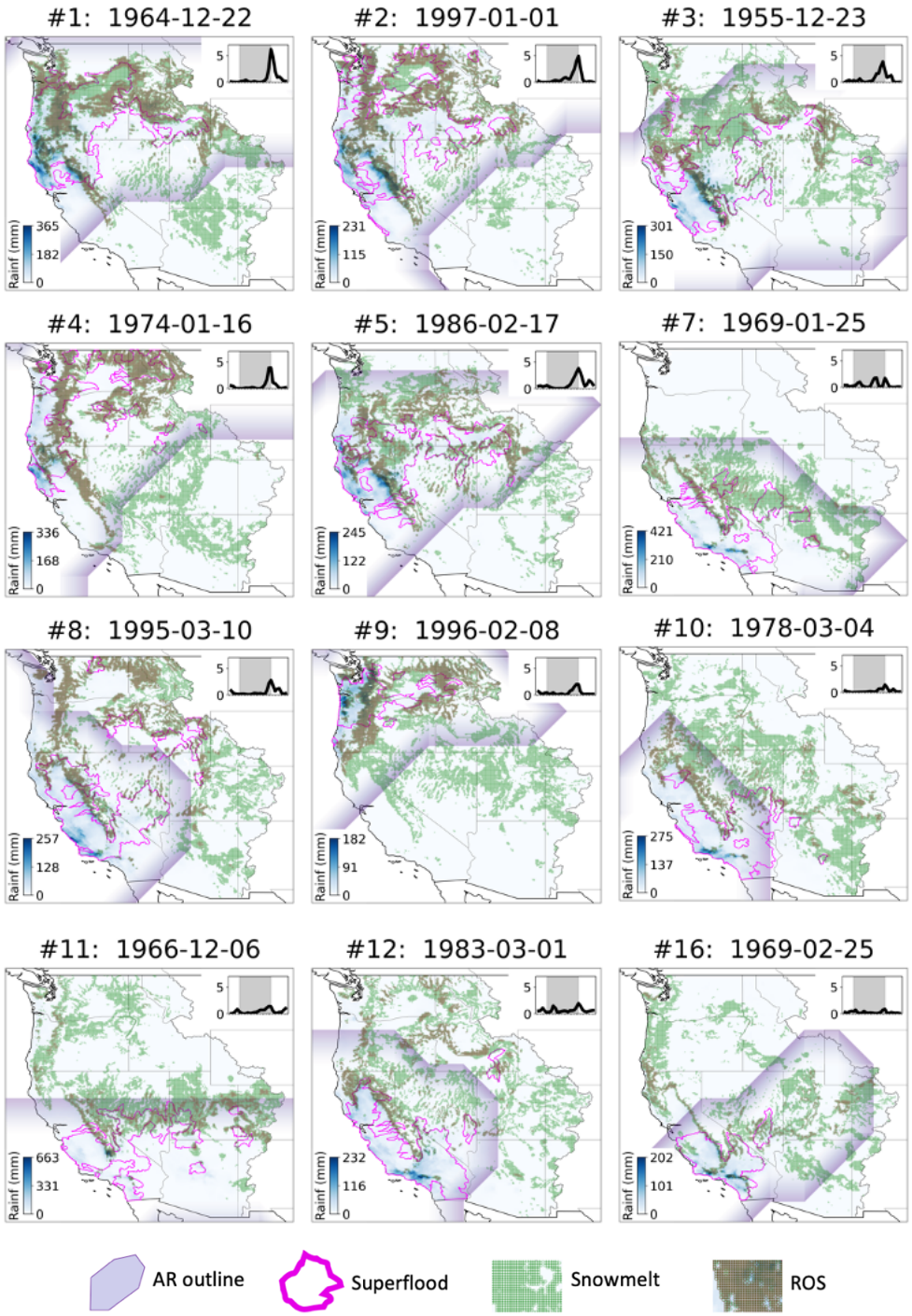


Figure 8: Hydrologic context of the largest winter superfloods. Spatial pattern of snowmelt (green) rainfall (blue) and ROS (red) on the day of each superflood. The flooded basins are shown in pink and the AR extent in purple. The timelines show the evolution of rainfall (mm/d), averaged across all basins that make up each superflood (the two weeks prior to flood day are shaded in gray for scale).

To further examine the role of ARs in controlling the extent and severity of superfloods, we compared flood extent to AR extent, and IVT to rainfall over all 60 winter superfloods (minus the four that do not coincide with an AR). Figure 9a shows that flood extent is closely related to AR extent over the domain. The most intense ARs (measured by IVT) on the other hand, do not systematically result in the heaviest rainfall (Figure 9b). This is probably because rainfall intensities are controlled by multiple factors, including but not limited to IVT (e.g. orographic effects).

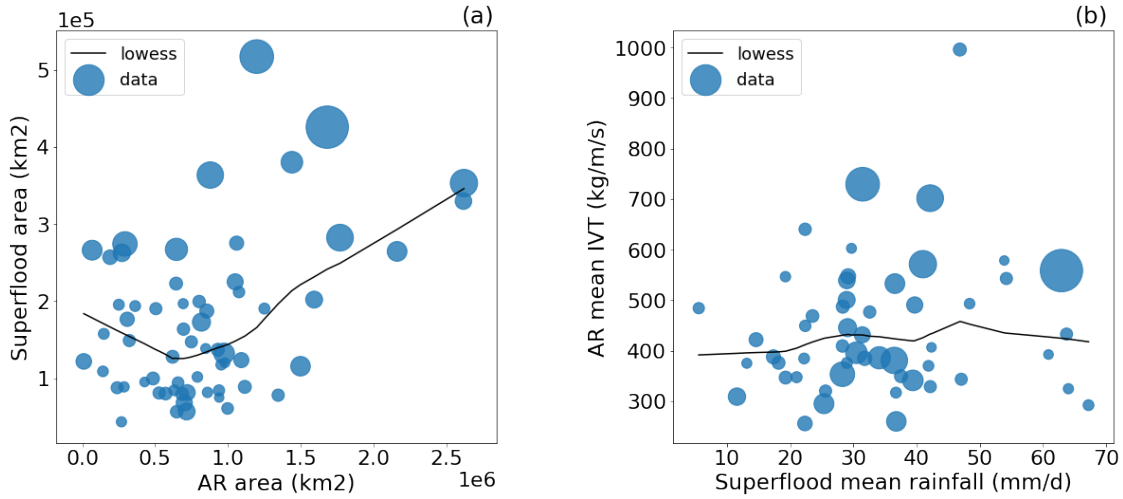


Figure 9: Flood footprint in relation to AR extent (a) and rainfall in relation to AR intensity (b). The AR area is the portion of the AR total area that is over the continental U.S. Mean rainfall is calculated only over the basins included in the superflood. The size of the circles reflects the superflood score.

3.4.2 Spring Events

We examined the spatial footprint of spring superfloods with particular attention to any differences between the rainfall-driven (20) and snowmelt-driven (10) events, as identified in Figure 7b. While both rainfall and snowmelt contribute in varying proportions to each spring superflood, this distinction provides some insights into the dominant cause. Figure 10 shows that in all spring events, rainfall peaks are much lower than in winter events. While SWE at this time of year is limited over most of the domain, melt occurred in many of the affected basins. The spring events typically occurred further in the interior (and in particular, to the east of the Cascades and Sierra Nevada, which effectively are coastal topographic barriers) than winter events. The 20 rainfall-driven superfloods

tend to center further north in the domain and affect the Rockies and share characteristics with winter superfloods which, while snowmelt is important at the edges, are primarily controlled by rainfall. In contrast, the 10 melt-driven superfloods tend to occur most often in the Colorado ranges, but their location is much less predictable, because they can occur at any location where substantial SWE has accumulated and surface conditions (primarily net radiation) are conducive to rapid snowmelt.

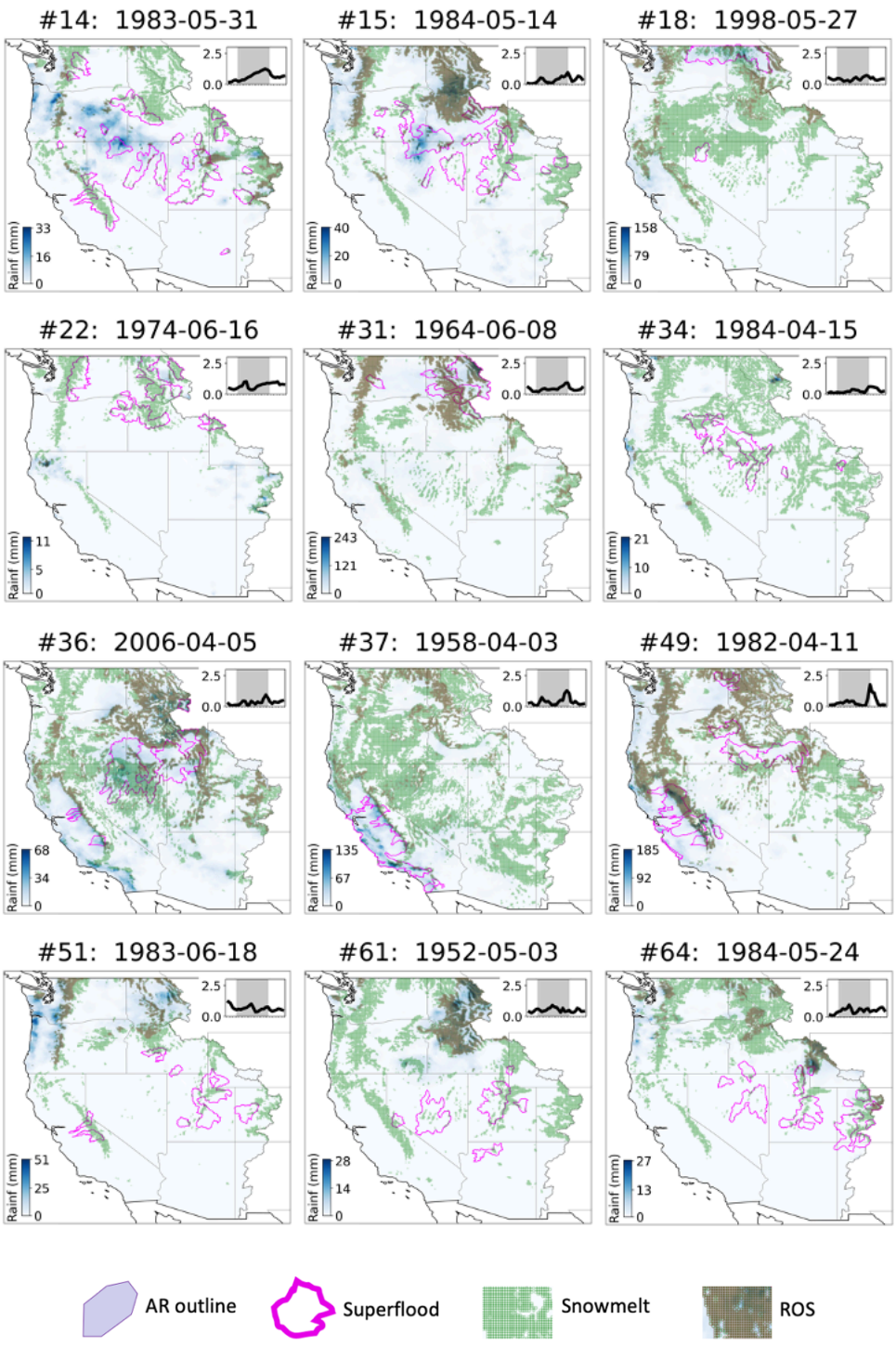


Figure 10: Hydrologic context of the largest spring superfloods (as in Figure 8). Spatial pattern of snowmelt (green) rainfall (blue) and ROS (red) on the day of each superflood. The flooded basins are shown in pink and the AR extent in purple. The timelines show the evolution of rainfall (mm/d), averaged across all basins that make up each superflood (the two weeks prior to flood day are shaded in gray for scale).

3.4.3 Summer Events

The summer superfloods differ from winter and spring events in that their drivers and resulting characteristics are varied despite there being only 10 summer events among the 100 largest superfloods (Figure 11). Previous studies have identified two common causes of flooding in Arizona and surrounding regions: the North American Monsoon (which typically occurs in June – September) and tropical storms (typically August - September) (Durrenberger et al., 1978). The 10 summer floods we identify are caused by diverse drivers (details and references in Supplement S2). Two coincided with tropical storms (Oct. 1983 and Aug. 1977), two with extratropical storms (Oct. 1962 and Oct. 2006), two with hurricanes (Oct. 1972 and Sept. 1982), two with frontal systems (Oct. 2004 and Oct. 1950), one with very heavy rains associated with unusual meteorological circumstances (Aug. 1951) and one with a monsoon thunderstorm (Aug. 1983). The majority of these floods occur in the arid interior, and center around Arizona (Figure 11). It is worth noting that many winter superfloods discussed earlier (e.g. see Dec. 1966 in Figure 4) extend east into the interior region, so floods also occur in the interior in winter, but are due to other causes that were discussed as part of winter floods (Section 3.4.1). Some snowmelt occurred in a few cases in those summer superfloods that occurred in the in the coastal ranges of the PNW.

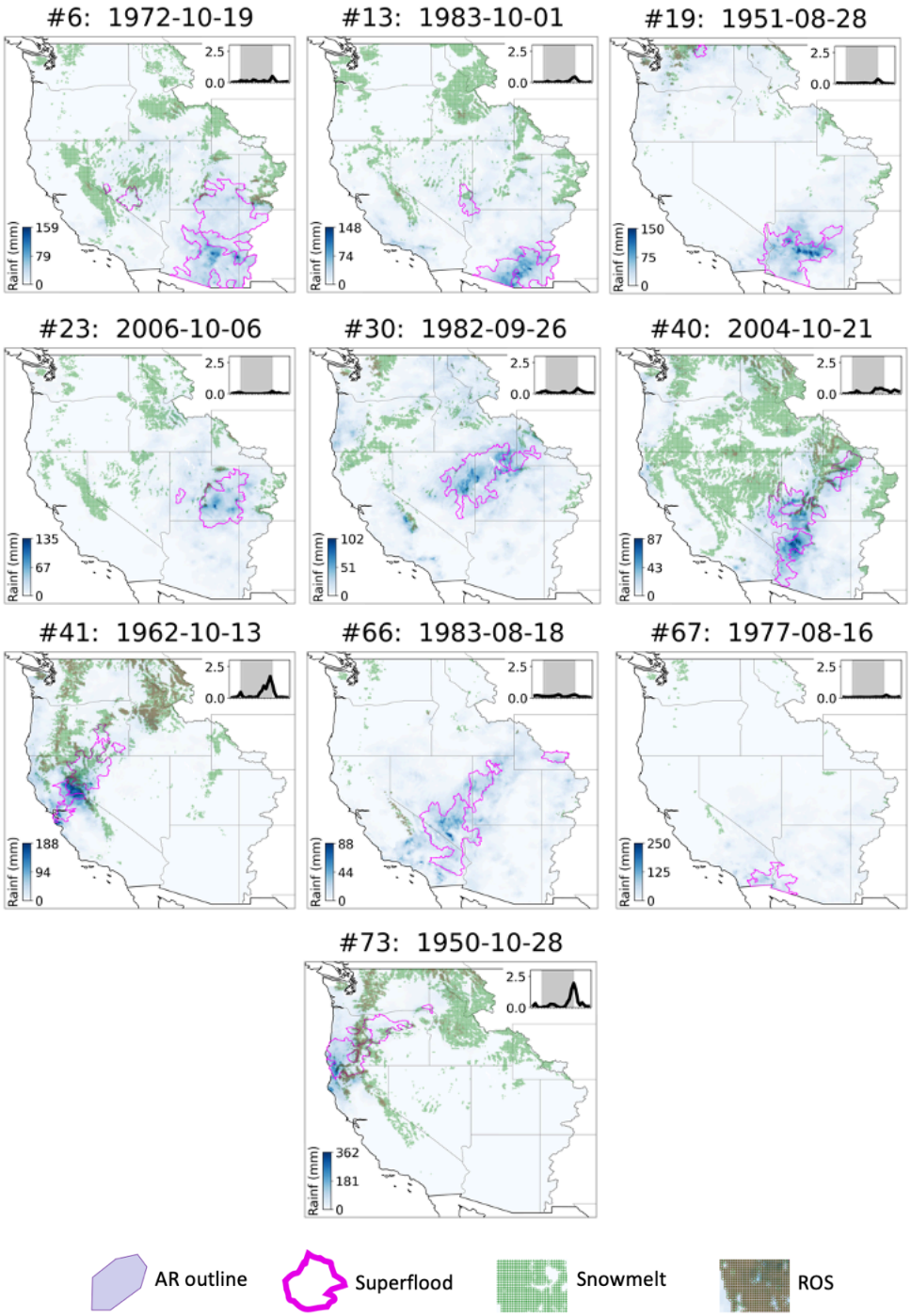


Figure 11: Hydrologic context of the largest summer superfloods (as in Figures 8 and 10). Spatial pattern of snowmelt (green) rainfall (blue) and ROS (red) on the day of each superflood. The flooded basins are shown in pink and the AR extent in purple. The timelines show the evolution of rainfall (mm/d), averaged across all basins that make up each superflood (the two weeks prior to flood day are shaded in gray for scale).

3.5 Runoff response during superfloods

To examine the runoff response in each of the superfloods, we used the peak-to-volume ratio (PV ratio), i.e. the ratio of the runoff peak to the 3-day runoff volume, to characterize the shape of the runoff responses. Figure 12a shows the PV ratios across basins for two high-scoring superfloods. While we had observed that the runoff response differs between the three seasonal clusters (see Figure 6a-c), it also appears that within each superflood, there is a range of PV ratios, from relatively large 3-day volumes and a modest peak runoff, to short duration and high magnitude peaks. For each superflood, we divided basins into those that had low, medium and high peak to volume ratios within a given superflood and examined the shape of the runoff response (timelines in Figure 12a). We found that while the PV ratios varied from basin to basin according to the height of the peak, the width of the runoff peak was similar for basins that belong to the same superflood, suggesting a strong influence of storm characteristics, as opposed to catchment characteristics, on the runoff response. Within a superflood, for the two events examined below, the largest PV ratios are found in basins near the location of intense snowmelt or rainfall, respectively (Figure 12a, by comparison with rainfall and snowmelt distribution in Figures 8 and 11, respectively). Nevertheless, Figure 12b shows that this variability occurs within a relatively restricted range, with PV ratios mostly between 0.4-0.7 across the largest 100 superfloods. Those superfloods that have high PV ratios are not necessarily the highest scoring, as the largest 12 superfloods overall (in amber) have PV ratios from the low to the higher end of the range observed in the wider set of 100 superfloods (in blue).

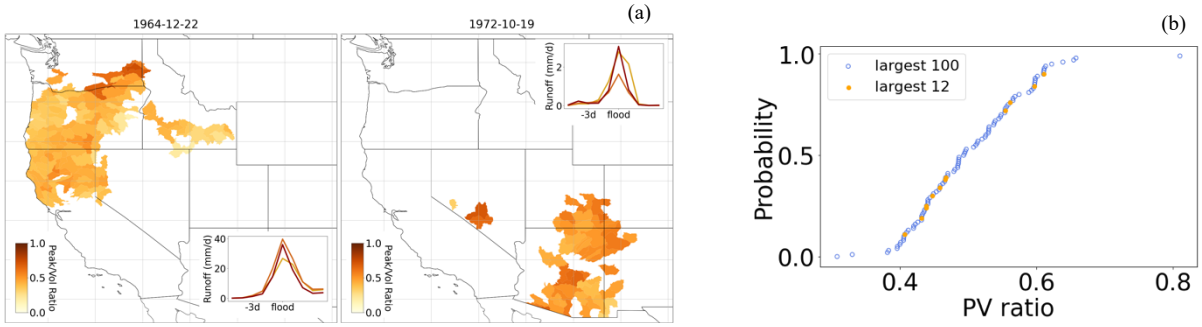


Figure 12: Flood response to superfloods. (a) Peak to volume ratios in each affected basin for two high-scoring superfloods. Timelines show runoff hydrographs for low, medium and high peak to volume ratio categories (note differences in scale). (b) CDF of PV ratios across all 100 largest superfloods. The PV ratio for each superflood is the average of PV ratios across basins that belong to that superflood.

4. Summary & Conclusions

We developed a method that allowed us to identify the most severe and spatially extensive historical floods in the Western U.S. during the period 1950-2010. We defined a superflood score for each day by summing flood magnitude inverse ranks across the entire domain using POT series of modeled or observed streamflow, such that the superflood score reflects both flood magnitude and footprint of flooding over the domain. Among the events we identified, the floods of December 1964 and January 1997 (which both affected the entire Northern half of the West coast) and October 1972 (in Arizona) are particularly prominent. The largest superfloods we identify are only weakly sensitive to the criteria we used.

We identified three general categories of superfloods: (1) 60% are winter events that occur dominantly in the coastal part of the domain. Almost all of these events coincide with ARs, and are characterized by extreme rainfall, and often (over at least part of the region) by ROS; (2) 30% are late spring floods, which are essentially all associated with anomalous winter snow accumulation, and often have components of both radiation melt and some ROS. Their peaks are generally smaller than winter events; (3) 10% are late summer events in and around Arizona, which are often associated with tropical storms and are almost exclusively caused by rainfall.

Considering multiple basins that flood simultaneously sheds light on the spatially variable response to common meteorological drivers. This highlights the multiple processes that shape the hydrologic response to a given storm, including orographic enhancement of rainfall and elevation-controlled snow dynamics. In particular, we find that the most extreme historical floods have been caused by both the extreme nature of each causative mechanism alone (e.g. rainfall amounts) and in many cases, unusual combinations of these factors, such as the interactions between rainfall and snowmelt in the snow/rain transition zone. ARs are also important in controlling the extent of flooding via zones of intense precipitation and snowmelt caused by warm temperatures.

These interactions are of particular importance in flood estimation and forecasting because they are typically responsible for “large” or “strange” floods, as identified by O'Connor & Costa (2003) and Smith et al. (2018), who have identified similar features (Pacific moisture, orographic effects, snowmelt and ROS) to the ones we describe, as the reason why the Western U.S. has particularly large floods relative to drainage area compared to other regions. These large historical

floods are the most relevant information available to estimate the upper tails of the flood distribution. This suggests further characterization of how the largest floods are generated, and in particular, the identification of realistic "worst case" sequences of combined temperature and precipitation for a given basin that would generate the largest possible flood. Such mechanistic understanding of flood sensitivity to their generation processes and possible interactions is expected to improve confidence in both our estimation of current large floods and how they will change in the future.

Acknowledgements

The research supported herein was funded in part by the Center for Western Weather and Water Extremes (CW3E) at the Scripps Institution of Oceanography UC San Diego via subcontract to UCLA. The authors declare no real or perceived conflicts of interests. All data used in this study are publicly available online at the following URLs: GAGES II data (https://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml#stdorder), and VIC model output from Li et al. (2019) (https://doi.org/10.6084/m9.figshare.8244398.v2). The VIC hydrologic model is open source and can be downloaded online (from https://github.com/UW-Hydro/VIC/releases/tag/VIC.4.2.d).

References

Alila, Y., & Mtiraoui, A. (2002). Implications of heterogeneous flood-frequency distributions on traditional stream-discharge prediction techniques. *Hydrological Processes*. https://doi.org/10.1002/hyp.346

Apipattanavis, S., Rajagopalan, B., & Lall, U. (2010). Local Polynomial–Based Flood Frequency Estimator for Mixed Population. *Journal of Hydrologic Engineering*. https://doi.org/10.1061/(asce)he.1943-5584.0000242

Ashley, S. T., & Ashley, W. S. (2008). The storm morphology of deadly flooding events in the United States. *International Journal of Climatology*. https://doi.org/10.1002/joc.1554

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States. *Water Resources Research*. https://doi.org/10.1029/2005WR004387

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*. https://doi.org/10.1029/2019WR024841

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

Berthet, L., Andréassian, V., Perrin, C., & Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-13-819-2009

Brocca, L., Melone, F., & Moramarco, T. (2008). On the estimation of antecedent wetness conditions in rainfall-runoff modelling. *Hydrological Processes*. https://doi.org/10.1002/hyp.6629

Buttle, J. M., Allen, D. M., Caissie, D., Davison, B., Hayashi, M., Peters, D. L., Pomeroy, J. W., Simonovic, S., St-Hilaire, A., & Whitfield, P. H. (2016). Flood processes in Canada: Regional and special aspects. *Canadian Water Resources Journal*. https://doi.org/10.1080/07011784.2015.1131629

Cao, Q., Mehran, A., Ralph, F. M., & Lettenmaier, D. (2019). The Role of Hydrological Initial Conditions on Atmospheric River Floods in the Russian River Basin. *Journal of Hydrometeorology 20(8)*.

Dettinger, M. D., Ralph, F. M., Das, T., Neiman, P. J., & Cayan, D. R. (2011). Atmospheric rivers, floods and the water resources of California. *Water*, *3*(2), 445–478.

Dettinger, M. D., & Ingram, B. L. (2013). The coming megafloods. *Scientific American*. https://doi.org/10.1038/scientificamerican0113-64

Durrenberger, R. W.; Ingram, R. S. (1978). *Major Storms and Floods in Arizona 1862-1977*.

Falcone, J. A., Carlisle, D. M., Wolock, D. M., & Meador, M. R. (2010). GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States. *Ecology*. https://doi.org/10.1890/09-0889.1

Fang, X., & Pomeroy, J. W. (2016). Impact of antecedent conditions on simulations of a flood in a mountain headwater basin. *Hydrological Processes*. https://doi.org/10.1002/hyp.10910

Fish, M. A., Wilson, A. M., & Ralph, F. M. (2019). Atmospheric River Families: Definition and Associated Synoptic Conditions. *Journal of Hydrometeorology*.

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*. https://doi.org/10.5194/hess-18-2695-2014

Gangrade, S., Kao, S. C., Naz, B. S., Rastogi, D., Ashfaq, M., Singh, N., & Preston, B. L. (2018). Sensitivity of Probable Maximum Flood in a Changing Environment. *Water Resources Research*. https://doi.org/10.1029/2017WR021987

Gershunov, A., Shulgina, T., Ralph, F. M., Lavers, D. A., & Rutz, J. J. (2017). Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophysical Research Letters*. https://doi.org/10.1002/2017GL074175

Harr, R. D. (1981). Some characteristics and consequences of snowmelt during rainfall in western Oregon. *Journal of Hydrology*. https://doi.org/10.1016/0022-1694(81)90006-8

Ivancic, T. J., & Shaw, S. B. (2015). Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge. *Climatic Change*. https://doi.org/10.1007/s10584-015-1476-1

James, C. N., & Houze, R. A. (2005). Modification of precipitation by coastal orography in storms crossing northern California. *Monthly Weather Review*. https://doi.org/10.1175/MWR3019.1

Jennings, K., & Jones, J. A. (2015). Precipitation-snowmelt timing and snowmelt augmentation of large peak flow events, western Cascades, Oregon. *Water Resources Research*. https://doi.org/10.1002/2014WR016877

Klemes, V. (1993). Probability of extreme hydrometeorological events - a different approach. *Extreme Hydrological Evente: Precipitation, Floods and Droughts*.

Leung, R. L., & Qian, Y. (2009). Atmospheric rivers induced heavy precipitation and flooding in the western U.S. simulated by the WRF regional climate model. *Geophysical Research Letters*. https://doi.org/10.1029/2008GL036445

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*. https://doi.org/10.1029/2019WR024950

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*. https://doi.org/10.1029/94JD00483

Merz, B., Aerts, J., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., Blöschl, G., Bouwer, L. M., Brauer, A., Cioffi, F., Delgado, J. M., Gocht, M., Guzzetti, F., Harrigan, S., Hirschboeck, K., Kilsby, C., Kron, W., Kwon, H. H., Lall, U., … Nied, M. (2014). Floods and climate: Emerging perspectives for flood risk assessment and management. *Natural Hazards and Earth System Sciences*. https://doi.org/10.5194/nhess-14-1921-2014

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

Mondal, A., & Mujumdar, P. P. (2015). Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2014.11.071

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., & Rasmussen, R. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. In *Nature Climate Change*. https://doi.org/10.1038/s41558-018-0236-4

Najibi, N., Devineni, N., & Lu, M. (2017). Hydroclimate drivers and atmospheric teleconnections of long duration floods: An application to large reservoirs in the Missouri River Basin. *Advances in Water Resources*. https://doi.org/10.1016/j.advwatres.2016.12.004

O'Connor, J. E., & Costa, J. E. (2004). Spatial distribution of the largest rainfall-runoff floods from basins between 2.6 and 26,000 km2 in the United States and Puerto Rico. *Water Resources Research*, *40*(1).

Parajka, J., Kohnová, S., Bálint, G., Barbuc, M., Borga, M., Claps, P., Cheval, S., Dumitrescu, A., Gaume, E., Hlavčová, K., Merz, R., Pfaundler, M., Stancalie, G., Szolgay, J., & Blöschl, G. (2010). Seasonal characteristics of flood regimes across the Alpine-Carpathian range. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2010.05.015

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

Pomeroy, J. W., Stewart, R. E., & Whitfield, P. H. (2016b). The 2013 flood event in the South Saskatchewan and Elk River basins: Causes, assessment and damages. *Canadian Water Resources Journal*. https://doi.org/10.1080/07011784.2015.1089190

Ralph, F. M., Rutz, J. J., Cordeira, J. M., Dettinger, M., Anderson, M., Reynolds, D., Schick, L. J., & Smallcomb, C. (2019). A scale to characterize the strength and impacts of atmospheric rivers. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-18-0023.1

Ralph, F. M., Neiman, P. J., Wick, G. A., Gutman, S. I., Dettinger, M. D., Cayan, D. R., & White, A. B. (2006). Flooding on California's Russian River: Role of atmospheric rivers. *Geophysical Research Letters*. https://doi.org/10.1029/2006GL026689

Sivapalan, M., Blöschl, G., Zhang, L., & Vertessy, R. (2003). Downward approach to hydrological prediction. *Hydrological Processes*. https://doi.org/10.1002/hyp.1425

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*. https://doi.org/10.1029/2018WR022539

Steinschneider, S., & Lall, U. (2016). El Niño and the U.S. precipitation and floods: What was expected for the January-March 2016 winter hydroclimate that is now unfolding? *Water Resources Research*. https://doi.org/10.1002/2015WR018470

Stockton, C., Meko, D., & Boggess, W. (1989). *Drought History and Reconstructions from Tree Rings*. Laboratory of Tree-Ring Research, University of Arizona (Tucson, AZ). https://repository.arizona.edu/bitstream/handle/10150/302684/ltrr-0046.pdf?sequence=1&isAllowed=y