

Water Resources Research

RESEARCH ARTICLE

10.1029/2018WR024604

Key Points:

- Developed expected flood potential utilizing regressions of streamgage record discharges across areas (zones) with similar flood magnitudes
- Upper prediction limits were used to and identify and rank extreme floods and assess zone variability
- Indices were developed for comparing floods between zones, to help understand how hazard varies

Supporting Information:

- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Figure S8
- Figure S9
- Figure S10
- Figure S11

Correspondence to:

S. E. Yochum,
Steven.Yochum@usda.gov

Citation:

Yochum, S. E., Scott, J. A., & Levinson, D. H. (2019). Methods for assessing expected flood potential and variability: Southern Rocky Mountains region. *Water Resources Research*, 55, 6392–6416. <https://doi.org/10.1029/2018WR024604>


Received 14 DEC 2018

Accepted 21 JUN 2019

Accepted article online 3 JUL 2019

Published online 2 AUG 2019

Methods for Assessing Expected Flood Potential and Variability: Southern Rocky Mountains Region

Steven E. Yochum¹ , Julian A. Scott¹, and David H. Levinson¹

¹National Stream and Aquatic Ecology Center, U.S. Forest Service, Fort Collins, CO, USA

Abstract Enhanced understanding of flood hazards, and how they vary across regions and continents, is needed to help protect lives and develop more resilient communities. Using the greater Southern Rocky Mountains region as a study area, a novel methodology was developed to predict, rank, and communicate expected flood magnitudes across similar responding areas (zones). Using 463 streamgages, up to 93% of the variance was explained by regression models developed for 11 derived zones. These regressions define the *expected flood potential* of each zone, a term introduced to assist practitioners, policy makers, and the public in understanding what flood magnitudes can be expected given the maximum recorded streamgage floods in nearby watersheds. Discharges above the 90% prediction limit, the *maximum likely flood potential*, are considered extreme; departure above this limit denotes the degree of extremity. The seasonality of the largest 5% floods varied substantially between zones, with the greatest frequency in July, August, and September in some zones (due to the North American monsoon) and May and June in other zones (due to snowmelt and rainfall). Using the lowest flood potential zone as an index area, flood potential and hazard indices were developed for comparing flood hazards across broad regions. The largest floods occur in the southern portion of the eastern slopes of the Southern Rocky Mountains and the adjacent Great Plains, with these events being 15 times larger than floods experienced in central Colorado and New Mexico mountain valleys, on average for a given watershed area.

Plain Language Summary Increased understanding of flood hazards, and how they vary, is needed to help protect lives and develop more resilient communities. Using the greater Southern Rocky Mountains region as a study area, a novel methodology was developed to predict, rank, and communicate flood hazards. Specifically, it is assumed that floods recorded in the last 135 years across similar areas referred to as zones can be used to predict what can be expected in the future. The *expected flood potential* was developed to communicate these predictions, with about half of the experienced floods in each of eleven zones being greater than this line, and the other half being less. The largest floods were identified and ranked using the *maximum likely flood potential*. Floods greater than this are unlikely but still possible and extreme. In contrast, the expected flood potential provides the flood magnitudes that are generally expected, given the historic record of floods. This method offers complimentary terminology to the commonly used but confusing term *100-year flood*. Indices were also developed to compare how large and hazardous floods are in different zones, with floods in some areas being up to 15 times greater than other areas for streams with the same watershed size.

1. Introduction

Enhanced understanding of flood hazards is essential for protecting human lives, infrastructure, and homes and businesses. Insight into the expected magnitude and spatial variation of floods is valuable for discerning flood extents, the geomorphic form and erosion hazards of streams and floodplains, the stability and inherent risk of stream restoration in a given area, the relative risk for flooding and water quality problems after wildfires, and the variability in probable maximum precipitation. At gaged locations, widely varying record lengths and periods complicate interpretations, while ungaged streams require alternative flood prediction methods. Current practice relies heavily on fitted statistical distributions of annual peak flow streamgage data at gaged sites (IACWD, 1982; England et al., 2018) and extrapolating these fits to ungaged streams through regional statistical regressions (Capesius & Stephens, 2009; Kenney et al., 2007; Kohn et al., 2016; Miller, 2003; Waltemeyer, 2008). However, such an approach relies on individual streamgage fits that can be poor and are susceptible to inadequate records, uses language that is problematic for laypeople and managers, and is limited by an assumption that floods in a given area do not have upper bounds.

Additionally, variability in flood magnitudes and flashiness are not easily compared across regions using the results of regional regression studies. As a result, there can be misunderstanding of flood hazards and how these hazards vary across landscapes. Additional methods for analyzing and interpreting flood hazards would be valuable.

Based on historical streamgauge records, as well as paleoflood and other peak flow estimates across the greater Southern Rocky Mountains region, this work develops a method for quantifying the magnitude of expected floods, introduces methods for identifying the spatial variability of large floods and the occurrence and ranking of extreme floods, and utilizes flood indices to quantify flood magnitude variability across broad geographic extents. The method utilizes the maximum record discharge at longer-term streamgages across zones of similar flood potential, with watershed delineations and other topographic features used to delineate zone boundaries. In a variation on the envelope curve method, regressions of the maximum experienced discharges as a function of drainage area (and, in some areas, an additional explanatory variable) were fit for each zone to provide a tool that can be used to quantify flood potentials and systematically identify extreme floods. It is proposed that the predicted magnitude is the *expected flood potential* of a given watershed and that the upper 90% prediction limit is the *maximum likely flood potential*. Additional analyses to quantify flood seasonality and identify trends were performed using the largest 5% floods.

By regressing the floods of record from many watersheds throughout a given geographic area, this space-for-time substitution avoids traditional flood frequency analysis and provides an alternative tool for consistently quantifying experienced flood severity on a zone basis, for inferring what can be expected in the future using comparisons with similarly responding watersheds. This paper describes the new procedure and applies it across the study area to illustrate how neighboring streamgages can be used to understand what size floods can be expected for watersheds in a given area, during what season they can typically be expected, and how the magnitudes of expected floods vary across and between regions.

2. Background

Flood hazards can be poorly communicated to those people who are at risk. There is frequent disconnect between the scientific analysis of flood attributes and public perception, and an overreliance on flood frequency analyses that are considered by some specialists to be problematic (Baker, 1994; Baker, 1998; Klemes, 1986; Klemes, 1989; Serinaldi, 2015). Consequently, potential exists for serious misunderstandings of expected flood hazards by decision makers and the general public, with the flood frequency paradigm and the use of such terms as the *100-year flood* arguably considered “erroneous as science and misleading/destructive as public policy/communication” (Baker, 2008). The use of updated language, such as the equivalent 0.010 annual exceedance probability (England et al., 2018), is more technically accurate than return intervals but may be even less clear for communicating hazards to the public. Less technical language for communicating expected flood hazards would be valuable to help the public and decision makers better understand risk.

In general, there are three methods for estimating flood magnitudes: (1) flood frequency statistical methods that fit statistical distributions to annual peak discharge data, for estimating the magnitude and frequency of floods at streamgages (England et al., 2018; IACWD, 1982), and regional regression methods for flood frequency estimation at ungaged locations; (2) rainfall-runoff analyses, using precipitation frequency estimates and assumed rainfall rate distributions (Moore et al., 2016; Sitterson et al., 2017); and (3) the development of empirically derived relationships between flood discharges and watershed characteristics (for extrapolation to ungaged locations). This last method has been most frequently developed as envelope curves of recorded floods from systematic streamgauge and paleoflood data (Enzel et al., 1993) and is also related to regional regressions. Additionally, probable maximum precipitation studies in combination with rainfall-runoff analyses are valuable for understanding extreme flood risks, with recent work performed in the Southern Rocky Mountains region to advance this science (DNR OSE, 2018).

Flood frequency methods for quantifying floods focus on at-a-station relationships of flood frequency and typically assume unbounded flood magnitudes, with the tail of the frequency distributions estimating magnitudes that are often much greater than the largest measured values. This can result in assigning a nonzero exceedance probability to a flood of any magnitude, no matter how large (Enzel et al., 1993). Instead of floods being open ended in magnitude (though very infrequent as the magnitude increases), it can be expected that

there is an upper limit to flood magnitudes due to physical limits in precipitation and watershed responses (Costa, 1987; Enzel et al., 1993; Wolman & Costa, 1984). Additional recognized research needs for flood frequency analyses include methods for (1) identifying and treating mixed distributions; (2) defining expected flood magnitudes in ungaged or insufficiently gaged watersheds; (3) including watershed physical processes, precipitation data, and rainfall-runoff models in the analyses; (4) addressing alteration imposed by urbanization, wildfires, and other watershed changes; and (5) addressing potential nonstationarity in flood response imposed by climate change and other causes for peak flow trends (England et al., 2018). An additional potential complication with flood frequency analyses is periodicity in flood (and drought) occurrences (McMahon & Kiem, 2018; Tipton, 1937), which is poorly understood over multidecadal or longer periods.

Methods for estimating flood frequency distributions at ungaged locations have been published by the U.S. Geological Survey (USGS) in a number of relevant publications for this study region (Capesius & Stephens, 2009; Kenney et al., 2007; Kohn et al., 2016; Miller, 2003; Waltemeyer, 2008). These analyses were performed by utilizing log-Pearson analyses of streamgage data (England et al., 2018; IACWD, 1982) and performing regressions across hydrophysiographic regions using a number of predictors, including some combination of watershed area; average watershed slope, elevation, outlet elevation, and mean annual and monthly precipitation; maximum precipitation intensity; 100-year precipitation frequency; area covered by herbaceous upland and dominated by clay; mean soils hydrologic index; percent of watershed above an index elevation; and the latitude and longitude of the basin outlet.

Empirically derived relationships of maximum recorded and mean annual flood discharges with watershed characteristics, frequently presented as envelope curves relating flood magnitude to watershed area, are additional tools for understanding flood risk. Envelope curves within and in the vicinity of the Southern Rocky Mountains have been investigated by numerous workers (Tipton, 1937; Crippen & Bue, 1977; Asquith & Slade, 1995; Herschy, 2002; Kenney et al., 2007; Michaud et al., 2001; O'Connor & Costa, 2004; Saharia, Kirstetter, Vergara, Gourley, & Hong, 2017; Waltemeyer, 2008). Early work in the spatial characterization of floods in this region was performed by Tipton (1937), who identified zones of dominant flood type (snowmelt, multiday rain events, and cloudburst rainstorms) to make predictions on flood magnitude with the limited data available at that time. Additionally, streamflow characteristics have been predicted using regressions with a variety of watershed characteristics as predictors in Arizona (Moosburner, 1970), with similar work performed in England and Scotland (Natural Environment Research Council, 1975) and Greece (Mimikou & Gordios, 1989). At a continental scale, Crippen and Bue (1977) delineated regional boundaries and developed envelope curves that provided expected limits for maximum floods for 17 regions across the continental United States. Also, O'Connor and Costa (2004) presented the top 10% annual peak discharge floods for nearly 15,000 streamgaging stations and included the 90th and 99th percentiles of this data subset, and Smith et al. (2018) performed a detailed analysis of the upper tail of flood peak distributions for more than 8,000 stations. At the global scale Herschy (2002) cataloged the largest floods for 1,500 stations in more than 100 countries. Typically, the flood frequency of events predicted using regional envelope curves is unknown; however, both Castellarin et al. (2005) and Vogel et al. (2007) have proposed methods for estimating flood frequency from such data sets.

Flood characterization work has identified the spatial distribution of the largest rainfall-runoff floods in the United States (O'Connor & Costa, 2004). The highest unit discharges (peak discharge per unit contributing area) are experienced across a widespread area, with half of the occurrences experienced in Texas, Puerto Rico, and Hawaii (which are susceptible to extreme precipitation events associated with tropical cyclones and convective thunderstorms). The southern Midwest, portions of the Appalachians, and the western flanks of the Pacific coastal mountain systems also have enhanced flood risk compared to other parts of the nation. Additionally O'Connor and Costa (2004) point out that the intermountain west has experienced some of the highest unit discharges, despite the presence of unexceptional rainfall volumes and rates compared to the humid eastern half of the continent. This could be attributed to bedrock exposure, thin soils, and steep relief in these arid and semiarid areas, as well as to bias introduced by data scarcity. At the regional scale, disparity has been noted between uncommonly large rainfall events and rare floods in the semiarid American West. Compared to more humid areas of the U.S., rainfall depths for the 100-year return interval in the semiarid west, for example, are smaller, while peak flow and flash flood potential is generally larger (Osterkamp & Friedman, 2000), with the cause hypothesized to be due to vegetation and soil conditions in semiarid areas. These findings conflict to an extent with the analyses of Smith and Smith (2015), specifically

in regard to spatial variability of flashy watersheds; this may be due to the relatively sparse streamgage records in semiarid areas. An additional consideration in the American Southwest is the prevalence of highly variable summertime monsoonal and convective storms.

Paleoflood and nonexceedance bound flood data add additional knowledge of flood risk, by extending the record of floods past the (maximum) 135 years of systematic streamgage records within this region. Paleofloods are events that occurred without being recorded by a systematic data collection apparatus; rather, flood stage record is preserved by slackwater sediment deposits and other paleostage indicators that can be age dated and modeled to estimate flood magnitude (Baker, 2008). Nonexceedance bounds are conservative peak flow estimates and are defined as a discharge that has not been exceeded during a specific time interval (Kohn et al., 2016; Levish, 2002). A number of studies have been performed that include portions of the greater Southern Rocky Mountains area (Godaire et al., 2013; Jarrett, 1990; Jarrett & Tomlinson, 2000; Kohn et al., 2016; Levish, 2002); when added to the systematic and historic records, these events can help develop greater understanding of flood potential.

From a hazard perspective, flood flashiness is an additional consideration to the expected flood magnitude. Flood flashiness refers to both how large a flood is compared to typical annual floods (Beard, 1975; Enzel et al., 1993), as well as the rate of short-term change in streamflow (Poff et al., 1997; Saharia, Kirstetter, Vergara, Gourley, Hong, & Giroud, 2017). A greater difference between the magnitude of large floods and more normal (and typically expected) annual peak streamflow, as well as rapid rises in hydrographs, results in more hazardous conditions—when large floods occur, they can be unexpectedly large and rapid rising, compared to more typical events. Flash floods have been investigated at the continental scale (Gourley et al., 2013; Michaud et al., 2001; Saharia, Kirstetter, Vergara, Gourley, & Hong, 2017; Saharia, Kirstetter, Vergara, Gourley, Hong, & Giroud, 2017; Smith & Smith, 2015). The flashiest watersheds in the contiguous U.S. were identified by Smith and Smith (2015) through the use of USGS instantaneous streamgage records (principally from the mid-1980s to 2013), with most instances of floods with unit discharges in excess of $1 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^2$ occurring along the West Coast, Southern Midwest, the Gulf Coast states, the Appalachians, and the piedmont physiographic province along the East Coast. Saharia, Kirstetter, Vergara, Gourley, and Hong (2017), Saharia, Kirstetter, Vergara, Gourley, Hong, and Giroud (2017) utilized National Weather Service flood definitions to identify regions at risk for the greatest floods. They incorporated seasonality, related flood magnitudes, and a flashiness index to climatology, geomorphology, and topography and mapped flash flood severity across the continental United States. Six regions where flash floods have been more frequent were identified, including the eastern slopes of the Southern Rocky Mountains and adjacent Great Plains (Saharia, Kirstetter, Vergara, Gourley, Hong, & Giroud, 2017). There have been a number of flashiness indices proposed, with computations based on the variability of peak flows (Baker et al., 2004; Beard, 1975; Smith & Smith, 2015), characteristics of the flood hydrographs (Gourley et al., 2013; Saharia, Kirstetter, Vergara, Gourley, Hong, & Giroud, 2017), as well as predictions based on watershed characteristics (Smith, 2010; Zogg & Deitsch, 2013). Additionally, Patton and Baker (1976) utilized a flood potential index based on the Beard flash flood magnitude index (Beard, 1975) and related this to watershed characteristics.

In summary, current methods for estimating, comparing, and communicating expected flood magnitudes within specific watersheds and across broad geographic areas are insufficient. Flood frequency analyses and regional regression studies have shortcomings that are often overlooked and do not provide tools for easily comparing spatial variability. Empirically derived relations between flood discharges and watershed characteristics offer the most opportunity for additional development, though the existing envelope curve approach is insufficient in that it only provides estimates of the most extreme floods. Clear language and additional tools are needed, with paleoflood data potentially adding substantive contributions for understanding risk. Additional methodologies for estimating expected large flood magnitudes and illuminating variability for comparing and ranking floods across wide areas would be valuable for more effective public communication and resource management.

3. Methods

3.1. Study Region

The study region included areas within and in the vicinity of the Southern Rocky Mountains (Figure 1), including a southern portion of the Central Rockies. Watersheds of up to $8,550 \text{ km}^2$ were included in the

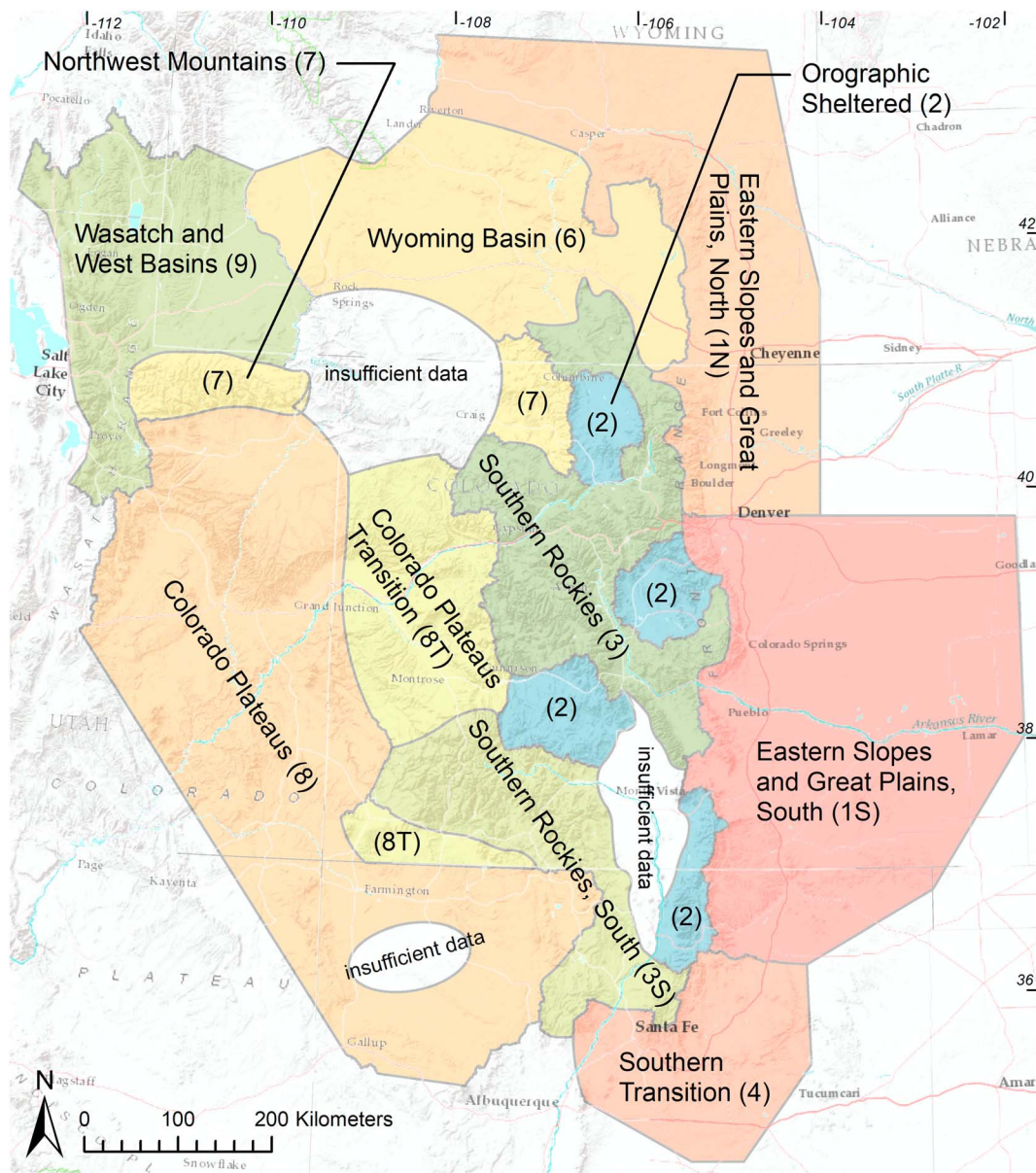


Figure 1. Study extent, with 11 delineated flood potential zones and areas with insufficient streamgage data.

analysis. This extent includes the Great Plains of eastern Colorado to the Great Salt Lake and from Casper, Wyoming, to Albuquerque, New Mexico. Represented physiographic provinces (Fenneman, 1931) are the Southern Rocky Mountains, Great Plains, Colorado Plateaus, Middle Rocky Mountains, Wyoming Basin, and Basin and Range. This region is susceptible to a wide variety of flood magnitudes, which vary by zone. Summary references documenting floods in the region include Tipton (1937), Follansbee and Sawyer (1948), Hansen et al. (1988), and Dollman (2017).

The overall watershed mean average precipitation is 610 mm, though zone averages vary substantially, from 400 and 410 mm (Wyoming Basin, Zone 6, and Colorado Plateaus, Zone 8) to 840 mm (Northwest Mountains, Zone 7). The region is susceptible to both rain- and snowmelt-induced floods, with rain on snow also being a possibility.

The North American monsoon is a driver of summertime rainfall and flooding across much of the study area, with greater influence in the southern portion. Sources for this monsoon moisture include the

eastern tropical Pacific Ocean and the Gulf of California, as well as the Gulf of Mexico, with temporal and spatial variability likely due to gulf surges and the latitudinal position of the midtropospheric subtropical ridge over southwestern North America (Adams & Comrie, 1997). Zones that experience the largest floods, specifically in the foothills and high plains on the eastern portion of the analysis extent (Zones 1N, 1S, and 4), are susceptible to large precipitation events that draw moisture from a variety of sources; for example, the September 2013 flood drew moisture from the Gulf of Mexico, Caribbean Sea, and tropical eastern Pacific Ocean (Gochis et al., 2015), and the June 1965 flood drew moisture from the Gulf of Mexico (Schwarz, 1967). Periodically, atmospheric rivers enhance moisture transport into this region from the eastern Pacific and Gulf of California (Rutz et al., 2014). However, the complex terrain of the Southwest and Southern Rocky Mountains usually diminishes the intensity and frequency of their inland penetration (Rutz et al., 2015). Based on data covering the period 1980–2017, Ralph et al. (2019) determined that moderate to strong atmospheric river events typically occur approximately once every 4–5 years over the Upper Colorado River basin and have a pronounced seasonal peak in winter. Additionally, the southern portion of the analysis extent is susceptible to flooding from tropical storm remnants that develop in the eastern Pacific Ocean and recurve to the north and northeastward into the southwestern U.S.; widespread flooding in Southwest Colorado in October 1911 (USGS, 2017a) may have been due to such an event.

3.2. Streamgage and Other Discharge Data

Annual peak discharge values are the fundamental flow data used; these data were obtained primarily from the U.S. Geological Survey (USGS, 2017a) but were also collected from additional sources, including the Colorado Division of Water Resources (Colorado Division of Water Resources, 2017), the Wyoming State Engineers Office (State Engineers Office, 2017), and the U.S. Bureau of Reclamation (Reclamation, 2017). Data through water year 2016 were used. All streamgages with at least 40 years of peak data were required to be included in the analysis, with exceptions for redundant sites (streamgages on the same stream within close vicinity), heavily urbanized watersheds (as identified using aerial imagery), and sites dominated by attenuation from upstream reservoirs that have little or no preimpoundment data. Specific peak flow values noted to be influenced by dam failures were also excluded. A 40-year period was ad hoc selected, with an implicit assumption that this record length balances the need for a sufficient number of data points while still encompasses multidecadal climate periodicity cycles (McMahon & Kiem, 2018). However, many areas have an insufficient number of streamgages with at least 40 years of record, especially for smaller-sized watersheds; in these areas streamgages that had experienced substantial peak discharges with records as short as 10 years were included in the analysis. The primary analyses used only the flood of record for each streamgage. For evaluating seasonality and large flood trends, annual peak flow data with a Hazen plotting position $\leq 0.05[(i - 0.5)/n \leq 0.05]$ were used.

Paleoflood and nonexceedance bound flood data were compared to the streamgaged-based zone analysis, to place these zonal analyses within a longer temporal perspective. These data were obtained from Jarrett and Tomlinson (2000), Kohn et al. (2016), and Godaire personal communication Godaire (2018).

Sites with measured peak flow data that experienced the highest unit discharges during the 2013 Colorado Front Range flood were used for comparison with the results of the zone analysis. Streamgages where the 2013 flood was the record peak discharge were excluded from this dataset, since this would be redundant. These data were obtained from Yochum and Moore (2013), UDFCD (2014), Schram et al. (2014), Kimbrough and Holmes (2015), Moody (2016), Yochum (2015), Brogan et al. (2017), and Yochum et al. (2017). Additionally, a peak flow value from the 1997 Fort Collins Flood (Grigg et al., 1998) was included in this assessment.

3.3. Zone and Watershed Delineation

Zones of relatively consistent flood hazards were identified, for communicating the consistency (within zone) and variability (between zones) of expected flood magnitudes and other attributes. Zone boundaries reflect watershed delineations and other topographic characteristics for the available streamgage data sets. These zone boundaries are geographically similar to the hydrologic region boundaries for regional regression equations (Capesius & Stephens, 2009; Kenney et al., 2007; Kohn et al., 2016; Miller, 2003;

Waltemeyer, 2008), though do differ. Importantly, the boundaries are approximate and may shift over time, as new floods occur and as climate change potentially modifies precipitation and runoff patterns.

Zones were heuristically delineated using the maximum measured (record) discharge at each streamgage, with the procedure consisting of the following:

1. Based on physiographic provinces and sections, watershed boundaries and topographic features, and a flash flood index (Beard, 1975), the center of an assumed zone was identified.
2. From this assumed geographic center, a plot of the maximum measured streamflow versus watershed area was created, using watersheds of a variety of scales and including all streamgages with at least 40 years of data.
3. Exploratory regressions were evaluated as individual streamgages were added, with a visual assessment of fit within the existing variability and changes in the explained variance (R^2) utilized to judge, as each was added, if this watershed is more likely part of the same zone or, alternatively, is more likely part of an adjacent zone.
4. Assumed centers of adjacent zones were identified and steps 1 to 3 were applied for each new zone
5. In boundary areas, consistency of adjacent streamgages with the discharge versus drainage area fits were used to decide which zones a particular watershed should be included within, or very infrequently, if a streamgage represented a watershed with a mixed flood response between two zones and was inappropriate for inclusion in either regressions. In the two instances where this occurred in this study (eastern portion of Zone 6), these data were combined with topographic information to refine the zone boundaries.
6. Iterative checks on fit were performed (as additional streamgages were assigned to each zone and with assessment of paleoflood and nonexceedance data) to reevaluate if the zone boundaries are most appropriate, with revisions that swapped streamgages between zones, identified new zones, and consolidated zones.

All streamgages with 40 or more years of data were required to be included in the analyses. Outliers excluded from the regressions, typically due to low values within the periods of record, are noted in the results.

Watershed delineations were created for each streamgage by modifying available watershed boundary data sets (USGS, 2017b) for the specific gage locations. HUC12 boundaries were used. Drainage areas were measured in a Geographic Information System using the most appropriate UTM zone for the specific watershed. These areas were compared to USGS records for each streamgage, for quality control, with the delineated area computations used.

3.4. Watershed Characteristics

In addition to drainage area, other watershed characteristics were tested for significance in predictions. For each watershed, the area-weighted precipitation was generated in R from 30-year (1981–2010), 800 m PRISM grids (Daly et al., 2008; PRISM, 2018) on an annual and monthly basis. Additionally, 30-m national elevation data sets (USGS, 2018) were obtained, with the arithmetic mean elevation, maximum elevation, and arithmetic mean slope and aspect calculated for each watershed in ArcGIS (version 10.5).

3.5. Indices

Four relatively simple flood indices were utilized in this analysis, including both prior published and new methods.

A flood potential index (P_f) was developed for between-zone comparison of expected flood magnitudes for identical watershed sizes. Specifically, from the zonal regressions of the maximum recorded discharge (Q) at long-term streamgages (the expected flood potential), discharge predictions were made for fixed watershed areas (A) of 20, 200, and 2,000 km². For each zone, the predicted discharges were normalized by the respective predicted discharges values for Zone 2 (Orographic Sheltered), which experienced the smallest flood magnitudes within the study extent t . The P_f is the average ratio:

$$P_f = \frac{\left(\frac{Q_{20}}{Q_{20zone2}} + \frac{Q_{200}}{Q_{200zone2}} + \frac{Q_{2000}}{Q_{2000zone2}} \right)}{3}$$

where Q_{20} is the fitted discharge for a 20-km² watershed for a specific zone, $Q_{20zone2}$ is the fitted discharge for

a 20-km² watershed in Zone 2, and with 200 and 2,000 noting likewise computations for 200- and 2,000-km² watersheds. The Zone 2 fitted discharges are $Q_{20\text{zone}2} = 4.15 \text{ m}^3/\text{s}$, $Q_{200\text{zone}2} = 21.0 \text{ m}^3/\text{s}$, and $Q_{2000\text{zone}2} = 106 \text{ m}^3/\text{s}$.

The flood variability index (V_f) is computed as a ratio of the intercepts for the regressions for the maximum likely flood potential (mlf) and the expected flood potential (efp), specifically $V_f = a_{\text{mlf}}/a_{\text{efp}}$, where a is the intercept term in the regression equations $Q = aA^b$.

The Beard flash flood index (F ; Beard, 1975) is computed as

$$F = \left[\frac{\sum (X_m - M)^2}{N-1} \right]^{1/2}$$

where X_m is the natural logarithm of the annual maximum flood, M is the logarithm of the mean annual maximum flood, and N is the number of annual events. This index is the standard deviation of the natural logarithms of the annual peak flows at each streamgage. Importantly, this index is a measure of the annual variability of floods (Beard, 1975; Enzel et al., 1993), rather than the steepness of a hydrograph (Poff et al., 1997; Saharia, Kirstetter, Vergara, Gourley, Hong, & Giroud, 2017).

Finally, it is proposed that a flood hazard index (H_f), based on both flood magnitude and flashiness, is valuable for developing a greater understanding of flood hazards, for comparison between gaged watersheds and flood zones. It is computed as

$$H_f = P_f * F$$

3.6. Statistical Analyses

Simple linear and multivariable regressions of the record peak discharges as a function of drainage area (and additional predictors) were performed using the R software (R Core Team, 2017; version 3.4.3). Natural logarithmic transformations were applied, which generally provided good adherence to regression assumptions of linearity, homoscedasticity, and independent and normally distributed residuals. Outliers were assumed to not be errors in discharge computations and were identified using Cook's distance (D) measurement of influence. High outliers were typically retained in the models to maintain conservative predictions, with the few exceptions noted in the results and discussion section. In zones where low outliers were excluded, these points were identified where $D > 1.9\mu$, where μ is the mean D . Low outlier exclusion was necessary to avoid developing models with less conservative flood predictions.

Prediction limits at 95%, 90%, 85%, and 80% were computed, for use in defining variability and extreme floods. Equations estimating the prediction limits were developed using a second regression of the fitted upper limit values. Significance tests utilized $\alpha = 0.05$. Explained variance was measured using the adjusted R squared.

Basic trend analyses were performed on the annual frequency of the largest 5% floods, for each zone, with $\alpha = 0.05$ determining significance. These trends were computed on both adjusted (for streamgage year frequency) values and unadjusted values. An adjustment is warranted to address the active yearly streamgage count peaking at 363 in 1968 and 225 in 2016. The record length over which trends were computed was 1940 to 2016 (due to pre-1940 streamgage data being much sparser). This simple adjustment was computed as

$$\text{Freq}_{\text{adj}} = \text{Freq} \frac{m_{\text{min}}}{m}$$

where Freq_{adj} is the adjusted annual frequency of the largest 5% floods, Freq is the unadjusted frequency of the largest 5% floods, m is the number of streamgages with annual peak discharge data in any given year, and m_{min} is the minimum number of streamgages operating in any of the years 1940 to 2016.

To test for significant differences between the expected flood potential and 100-year discharges, a Wilcoxon signed rank test was used.

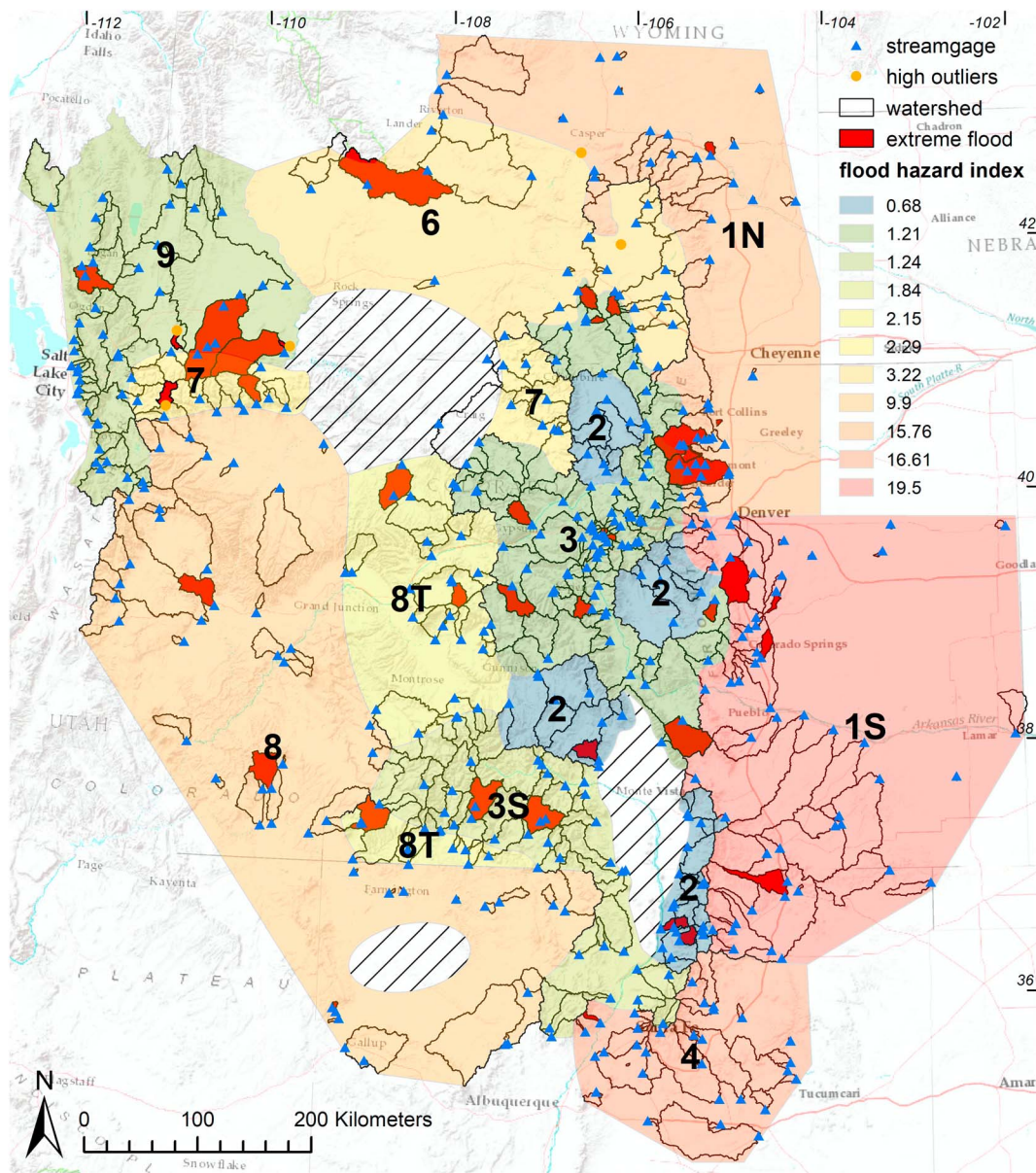


Figure 2. Flood hazard index values for the zones within the greater Southern Rocky Mountains analysis extent, with streamgauge locations, watershed delineations for each streamgauge, high outliers, and watersheds that have experienced extreme floods. The crosshatched areas have insufficient streamgauge data for method application.

3.7. Results and Discussion

This study developed methodologies for predicting flood magnitudes at ungaged locations, identifying and ranking extreme floods in a systematic way, and comparing widely varying flood potential across continental-scale areas. The methods utilize spatial aggregations of streamgauge records to gain insight on flood magnitudes that can be expected within areas referred to as zones, in a space-for-time substitution assessment of flood risk. Using regressions of streamgauge maximum record discharges across zones of similar flood response, the *expected flood potential* concept was developed to allow prediction at ungaged and insufficiently gaged watersheds using methods independent of flood frequency analyses and subsequent regional regression studies. The term *zone* is used to help differentiate this method from conventional flood frequency regional analyses. Upper prediction limits were used to assess variability, and identify and rank extreme floods, with departure above the limit determining the degree of extremity.

Table 1
Analysis Extent and Zone Characteristics for the Greater Southern Rocky Mountains Region

	Overall	1S	1N	2	3	3S
# streamgages	463	52	42	40	89	45
# gage-years with Hazen Prob. ≤ 0.05	1,261	126	100	126	265	149
Record length maximum average minimum	135	107	135	103	128	127
	54	47	46	63	59	66
	10	10	11	10	14	21
Watershed maximum area average minimum	8,551	7,442	2,730	3,699	7,933	3,407
	751	1,037	507	516	702	597
	1.5	4.5	2.2	5.7	4.0	89
maximum experienced flood (m^3/s)	4,600	4,600	900	190	610	710
R^2 , Q_{efp} with area		0.65	0.75	0.89	0.92	0.66
p value		3.79E-12	2.71E-13	<2.2E-16	<2.2E-16	7.51E-12
F statistic		88	120	282	961	86
Flood potential index, P_f		15.0	13.8	1.0	2.3	3.0
Flood variability index, V_f		2.77	1.76	1.52	1.62	1.94
Beard flash flood index, F		1.30	1.10	0.69	0.54	0.61
Flood hazard index, H_f		19.5	15.8	0.69	1.2	1.8
Q_{efp} (for 1,000 km^2 ; m^3/s)		750	450	65	150	220
Q_{mir} (1,000 km^2 ; m^3/s)		2,000	790	98	240	430
R^2 , Q_{efp} with area and paleofloods		0.62	-----	-----	0.92	0.66
p value		1.96E-12	-----	-----	<2.2E-16	3.62E-12
F statistic		86	-----	-----	1015	89
R^2 , Q_{efp} with area+1		0.70	-----	-----	0.93	0.71
Additional predictor		Nov. precip.	-----	-----	ave. precip.	ave. precip.
p value		9.91E-13	-----	-----	<2.2E-16	1.63E-12
F statistic		55	-----	-----	583	55
Significant trends? (none, up/down)		none	none	down	none	none
Months of flood occurrence		8,6.7	(5,6),9.7	5,6.7	6	6,5.9
Number of extreme floods:	40	4	4	3	9	3
Most extreme floods: (with year-month of occurrence and USGS identification numbe)	----	(3) Plum Creek, CO (1965-6), 06709500 (2) Kiowa Creek, CO (1935-5), 06758000 (1) Jimmy Camp Creek, CO (1965-6), 07105900	(3) Big Thompson River, CO (1976-7), 06738000 (2) Spring Creek, CO (1997-7), 06651800 (1) Sand Creek, WY (1955-8), 08268500	(3) Arroyo Hondo, NM (1948-7), 08268500 (2) Rio Pueblo de Taos, NM (1979-5), 08269000 (1) Camero Creek, CO (1945-7), 08230500	(3) Rock Creek, WY (2010-6), 06632400 (2) Pass Creek, WY (1984-5), 06628900 (1) Sweetwater Creek (1976-7), 09061450	(2) Los Pinos River, CO (1957-7), 09353500 (1) San Juan River, CO (1911-10), 09342500 (2) Vallecito Creek, CO (1970-9), 09352900

Note. Definitions: efp, expected flood potential; mlf, maximum likely flood potential; area+1, watershed area plus one additional watershed characteristic used as a predictor. For all significance tests, $\alpha = 0.05$.

Table 1
(continued)

	4	6	7	8T	8	9
# streamgages	29	21	24	29	45	45
# gage-years with Hazen Prob. ≤ 0.05	73	39	84	71	98	133
Record length (years)						
maximum	97	86	112	100	81	121
average	50	36	68	49	38	61
minimum	21	12	20	10	10	22
Watershed area (km ²)						
maximum	2,697	6,163	8,551	2,767	4,503	8,080
average	481	1,608	872	527	682	1,053
minimum	1.5	1.6	22	14.2	2.7	8.1
maximum experienced flood (m ³ /s)	2,100	170	710	200	1,190	280
R^2 , Q_{efp} w/Area						
p value	0.82	0.87	0.82	0.81	0.69	0.79
F statistic	1.44E-10	6.07E-10	4.53E-09	1.04E-10	2.71E-12	2.97E-15
Flood potential index, P_f	113	130	96	112	96	153
Flood variability index, V_f	14.0	3.6	4.7	2.7	9.0	1.9
Beard flash flood index, F	1.97	1.43	1.38	1.50	1.88	1.78
Flood hazard index, H_f	1.19	0.90	0.49	0.80	1.10	0.66
Q_{efp} (for 1,000 km ² ; m ³ /s)	16.6	3.2	2.3	2.2	9.9	1.2
Q_{mlf} (1,000 km ² ; m ³ /s)	630	79	200	120	280	81
R^2 , Q_{efp} with area and paleofloods	1,200	110	280	180	520	140
p value	---	---	0.83	0.85	0.70	0.79
F statistic	---	---	9.63E-16	4.30E-13	3.84E-13	5.71E-16
R^2 , Q_{efp} with area+1	0.85	0.90	185	160	104	161
Additional predictor	aspect	ave. elev.	Sep. precip.	---	0.74	---
p value	1.17E-10	2.89E-10	6.07E-09	---	ave. elev.	---
F statistic	73	94	60	---	6.87E-13	---
Significant trends? (none, up/down)	down	none	up	none	61	---
Months of flood occurrence	8,7,9	6,4	6,5	5,6	down	none
(ordered with most frequent first):					8,9,7	5,6
Number of extreme floods:	1	2	2	4	4	4
Most extreme floods: (with USGS identification number)	(1) Rito de Los Frijoles, NM (2013-13), 08313350	(2) Bobcat Draw, WY (1977-5), 06638090	(1) Sweetwater River, WY (1980-4), (2005-5), 09299500	(3) Plateau Creek, CO (1922-5), 09096500	(2) Chuska Wash, NM (1967-9), 09367860	(1) Cottonwood Wash, UT (1968-8), 09378720
			(2) Whiterocks River, UT (2005-5), 09299500	(1) Yellow Creek, CO (1978-9), 09315500	(3) Blacks Fork tributary, WY (1973-7), 09224600	(2) trib. to the Green River, UT (1959-7), 09225300
			(1) N.F. Duchesne River, UT (1963-6), 09274000	(1) McElmo Creek, CO (1927-9), 09371500	(1) Saleratus Wash, UT (1962-9), 09315500	(1) Sulphur Creek, UT (1983-6), 10015700

Simple indices were developed to compare flood potential between zones, to help understand how flood hazard varies across wide and diverse geographic extents. This work is similar to the contributions of Moosburner (1970), Natural Environment Research Council (1975), and Mimikou and Gordios (1989), though focuses on the largest discharges instead of the mean annual and other streamflow characteristics, and takes the approach further by introducing tools to identify and rank extreme floods, and compare flood magnitude potentials.

The greater southern Rocky Mountains region was used to develop the methodology and provide examples on potential use (Figure 1). The analyses were performed for watersheds with less than 8,600 km² of contributing area in the states of Colorado (CO), Wyoming (WY), New Mexico (NM), Utah (UT), Idaho (ID), and Arizona (AZ).

Given the streamgage records of record peak discharges at neighboring streamgages, the developed methods assist practitioners and researchers with answering such questions as

1. What large flood magnitudes can be expected at a given ungaged location?
2. How reasonable are the results of regional flood frequency regression equations?
3. Is a flood frequency analysis at a specific streamgage providing reasonable results, or are results biased due to the presence or absence of a large flood?
4. What areas are inherently prone to larger or smaller floods (have a larger or smaller flood potential index)? Such understanding is valuable for more informed decisions regarding the erosion hazards and geomorphic form of stream corridors and the inherent risks of stream restoration and wildfire-induced flooding on communities.
5. Is a specific flood extreme or rather a typical large flood?
6. Compared to other floods in the area, how extreme is a flood?

3.8. Overview of Streamgage Record and Largest Floods

Annual peak discharge data for 463 streamgages were utilized in the analysis (Figure 2). The longest individual streamgage record is 135 years (Table 1), with an average record length of 54 years and the earliest annual peak flood recorded in 1882 (Cache la Poudre River at Mouth of Canyon, CO). There are a wide variety of record lengths and periods, with the total active yearly streamgage count peaking at 363 in 1968 and dropping to 225 by 2016. Three zones (2, 3S, and 6; central mountain Colorado and New Mexico valleys, southern portion of the Southern Rocky Mountains, and Wyoming Basin) have had extended periods of large flood inactivity through 2016. Additionally, trends in the frequency of the largest 5% floods are significant (downward) in three zones (Table 1). However, interpretations regarding these patterns are complicated by variability in the streamgage record (Table 2).

Large floods occur more frequently over a multiyear period, followed by a gap of lesser activity before another active period initiates (Table 2). Similar observations have been previously noted (McMahon & Kiem, 2018; Tipton, 1937) and may be related to large-scale interannual to interdecadal ocean-atmospheric variability such as the El Niño/Southern Oscillation and the Pacific Decadal Oscillation (Ely et al., 1993; Ward et al., 2014). However, ocean conditions and atmospheric teleconnections have complex interactions with the terrain of the Intermountain West (Wise, 2012). Low-level and midlevel moisture from the Pacific must travel over the coastal mountain ranges but does so with significant air mass modification (Bryson & Hare, 1974) as the dominant flow of Pacific air around the southern end of the Sierra Nevada is typically warmer and drier. This situation modifies during El Niño years, when the subtropical storm track intensifies across the southwest U.S., providing an enhanced Pacific moisture source to the southern Intermountain West (Cayan et al., 1999).

A plot of the largest 5% floods illustrates the wide variability in magnitudes on a per watershed area basis (Figure S1 in the supporting information) in this study extent; this is similar to the presentation in O'Connor and Costa (2004) for the entire United States. Months of annual peak flow occurrence range from January to November, with the highest frequency in June and May (Figure S1). A plot of record peak discharges for the database (Figure S2 also shows wide variability ($R^2 = 0.41$) for the project extent. These approaches for understanding observed flood magnitudes, while helpful for placing individual floods within a regional context, are too coarse for predicting flood risk in particular watersheds.

Table 2

For Each Zone, Years With the Frequency of Floods > Annual Average (Shaded), for the Top 5% of the Measured Floods

Year	Gage count	Zone/threshold exceeded for frequency/month										
		1S	1N	2	3	3S	4	6	7	8T	8	9
1904	s						9					
1905	32											
1906	26											
1907	22											5,6
1908	29											
1909	38				6,8							6
1910	36											
1911	67					10						
1912	75											5,6
1913	68											
1914	73				6							
1915	77											
1916	83											
1917	85				6							
1918	84				6							
1919	78		7,8									
1920	89					5						
1921	88		6,7		6,8	10,6			6			
1922	96											5
1923	103		6,7									
1924	102											
1925	91											
1926	88											
1927	100					6				6,8,9		
1928	97											
1929	104											
1930	110											
1931	109											
1932	117											
1933	122		5,6,7,9									
1934	122									8		
1935	136	5,8										
1936	145			7,8								
1937	154			4,5,6		4,5	6					
1938	167		9				9,10			6		
1939	171											
1940	192									8,9		
1941	194			5,7		5,9,10				5,10	8	
1942	197	4		4,5		4,5	4,9			4,5		
1943	198											
1944	204			5,7								
1945	207											
1946	216											
1947	223			5,6,7								
1948	228											
1949	234					6						
1950	241											
1951	249		5,6,7,8									
1952	258				6	5,6					5,8	5
1953	271										7,8	
1954	272	7,8					7,8					
1955	275	5					7,9					
1956	273											
1957	276			6,7,8	6,7	6,7	8	6			7,8,11	
1958	277											
1959	291											
1960	299											
1961	311						8,9	7			8	

Table 2 (continued)

Year	Gage count	Zone/threshold exceeded for frequency/month										
		1S	1N	2	3	3S	4	6	7	8T	8	9
1962	318							2,4			2,9	
1963	323											
1964	331											
1965	356	6,7	5, 6		6,7		7,9	6	6			6
1966	354						8					
1967	357		6,7				8				7,9	
1968	363										7,8	
1969	357		5									
1970	354		6			7,9		6,8		9	9	
1971	345						7					
1972	344							6,8		10	7,9,10	
1973	340		5,6					5		5,6	3,7	
1974	340								4,5			
1975	338											5,6,7
1976	348											
1977	347									7,8		
1978	347	7,8									7	
1979	356	8		5,6		5,6						
1980	356							4			9	
1981	349	7,8									9	
1982	330						6,7,9					4,5,6,9
1983	310			5,6,7	6,7,8			4,6	5,6	5,6		5,6,8
1984	297			5	5,6,7	5,7			5	5,6,8		5
1985	285				5,6	6,7						
1986	290							6	5,6			2,4,5,6
1987	271											
1988	273										6,8	
1989	274											
1990	276											
1991	287			5		5					9	
1992	284					4,8						
1993	289									5		
1994	283			5,6,7		5,6,8	5					
1995	277		5,6	6	6,7				6	3,5,6		
1996	265						7					
1997	248											
1998	245											
1999	263	4,7,8		5,7								
2000	257											
2001	264											
2002	267											
2003	267				5,6							
2004	261	8										
2005	261									5,9		4,5
2006	263	7,8										
2007	234											
2008	255										7,8	
2009	251											
2010	250				6,8		7,8		6			
2011	245				6,7				6,7			4,5
2012	238											
2013	247	9	9				9					
2014	235				5,6							
2015	240	5,6										
2016	225											

Note. Months of flood occurrence are the numbers listed in each shaded year. Note that for each zone, floods occur more frequently over some multiyear periods, with gaps of lesser activity between these times of increased flood occurrences. Large floods have been relatively infrequent for a number of years in Zones 2, 3S, and 6.

3.9. Zones and Flood Potential

Subsets of the data set were delineated into zones of similar experienced large floods (Figure 2). This approach is comparable to the envelope curve work presented for this region by Tipton (1937), as well as the continental scale approach of Crippen and Bue (1977) and global scale of Herschy (2002), though instead of defining an upper boundary, regressions were fit to the maximum recorded discharge at each streamgage (Figures 3–5) at a spatial scale appropriate for prediction at ungaged locations. Using these discharges (black dots), regressions using watershed area as the predictor explained most of the variability in the dataset for each zone (92% to 65%, Table 1). Equations for each of the regressions are provided in Figures 3–5. Each of these fits defines the *expected flood potential* for each zone, which uses the maximum recorded discharges at each streamgage to understand central tendencies and expectations regarding flood magnitudes. This approach sidesteps communication problems associated with the use of such terms as the 100-year flood (Baker, 2008) or the 0.010 annual exceedance probability flood, in preference for defining what floods can be expected at any location given record peak discharges in similar, nearby watersheds.

The highest explained variance are in zones dominated by snowmelt runoff for the largest floods (Southern Rockies, Zone 3, and Northwest Mountains, Zone 7), and the Southern Transition (zone 4), which is rainfall dominated during the North American monsoon. To understand variability, the upper 95%, 90%, 85%, and 80% prediction limits were computed for each regression. The zone percentage of floods above each of these thresholds was computed (Figure 6), with a heuristic decision that the 90% prediction limit was the most appropriate threshold due to the inflection at this point; this limit is defined as the maximum likely flood potential. Equations approximating the maximum likely flood potential are provided in Figures 3–5.

The addition of climate and topographic predictors to watershed area were tested for significance in refining the expected flood potential. In 7 of the 11 zones, the addition of one additional predictor increased the explained variance to between 70% and 93%. These additional watershed characteristics were valuable for increasing prediction accuracy, though were not effective in all zones.

Most of the zones had a few low outliers excluded from the expected flood potential prediction (Figures 3–5). These points were excluded because they expressed excessive influence on expected flood potentials, with their inclusion resulting in reduced (and nonconservative) flood potential predictions. As the purpose of this exercise is to provide an assessment of flooding threat, a conservative approach (from an engineering perspective) warrants excluding these low outliers to avoid underestimation. There are two hypotheses for the presence of these low outliers: (1) These watersheds are in precipitation shadow areas, with inherently less flood potential than adjacent watersheds, and (2) these watersheds have not yet experienced floods of similar (large) size as adjacent watersheds, but they will likely experience such large floods in the future.

Months of occurrence of the largest 5% floods are also provided in Figures 3–5 and Table 1. The flood seasonality within each zone varies, with the largest floods occurring in August in Zones 1S, 4, and 8 (due to the North American monsoon), and the remaining zones most frequently experiencing floods in May and June, due to snowmelt (Zones 3 and 7), as well as rain events.

Zone boundaries were frequently watershed divides. However, for Zone 1N (eastern slopes and Great Plains, north), the western boundary between this high flood potential zone ($P_f = 13.8$) and the adjacent Southern Rockies zone ($P_f = 2.3$) is not a watershed divide but is instead a series of topographic features (Figure S4). This approximate boundary is most appropriate given the available discharge data and is similar to (but varies from) a 2,300-m hypothesis regarding the boundary between snowmelt and rainfall-dominated floods and flood extremes (Capesius & Stephens, 2009; Follansbee & Sawyer, 1948; Jarrett, 1990; Kohn et al., 2016). Mechanistically, it is thought that as a warm, moist air mass is forced upslope by the mountains, this topography initially intensifies precipitation rates due to orographic lifting, but as the air mass rises still higher farther west, available moisture in the air column tends to decrease and precipitation rates decrease (Hansen et al., 1988; Osterkamp & Friedman, 2000). This results in higher-based convection with less moisture availability. However, 2,300 m is approximate, as several watersheds in the Big Thompson watershed indicated during the 2013 flooding; the provided zone boundary was developed based on topographic features and peak discharge data, rather than a 2,300-m contour. Additional exceptions to the 2,300-m hypothesis were noted by Smith et al. (2018).

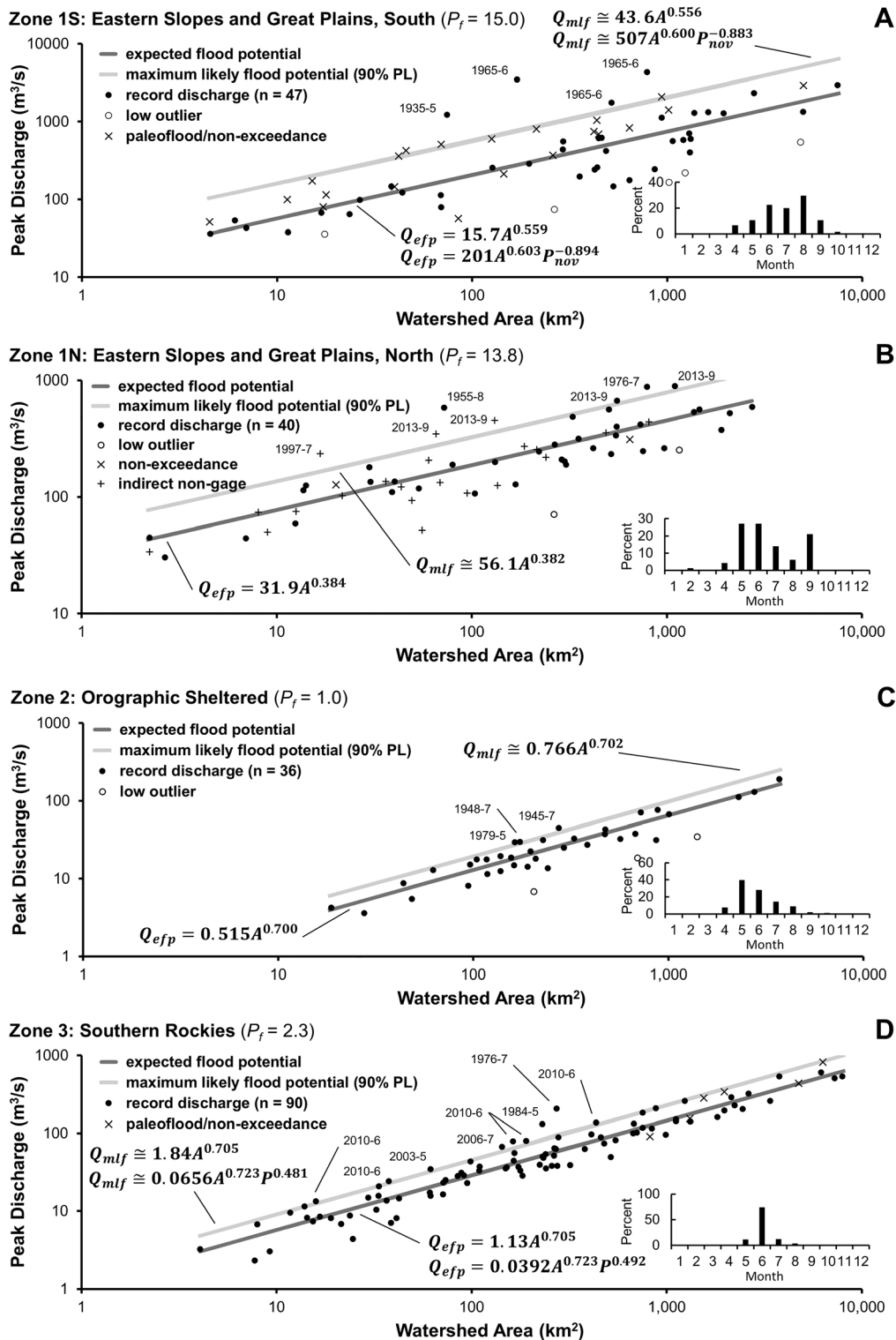


Figure 3. Peak discharge versus watershed area, for Zones 1S, 1N, 2, and 3. P_f is the flood potential index. Black dots indicate the record peak discharges for each streamgage, which were used to develop the expected flood potential regression (dark gray) and the maximum likely flood potential (light gray, 90% prediction limit), with associated prediction equations provided. Where topographic and climate variables were significant in multivariable regression, a second set of equations is provided. Excluded low outliers are noted with circles. Where available, paleoflood and nonexceedance flood estimates (marked with x), and indirect nonstreamgage peak flow estimates (marked with +) are also shown. Months of peak flow occurrence for the highest 5% of floods are also indicated. Q_{efp} : expected flood potential discharge; Q_{mif} : maximum likely flood potential discharge; A : watershed area; P : watershed average annual precipitation (mm); P_{nov} : watershed average monthly precipitation for November (mm).

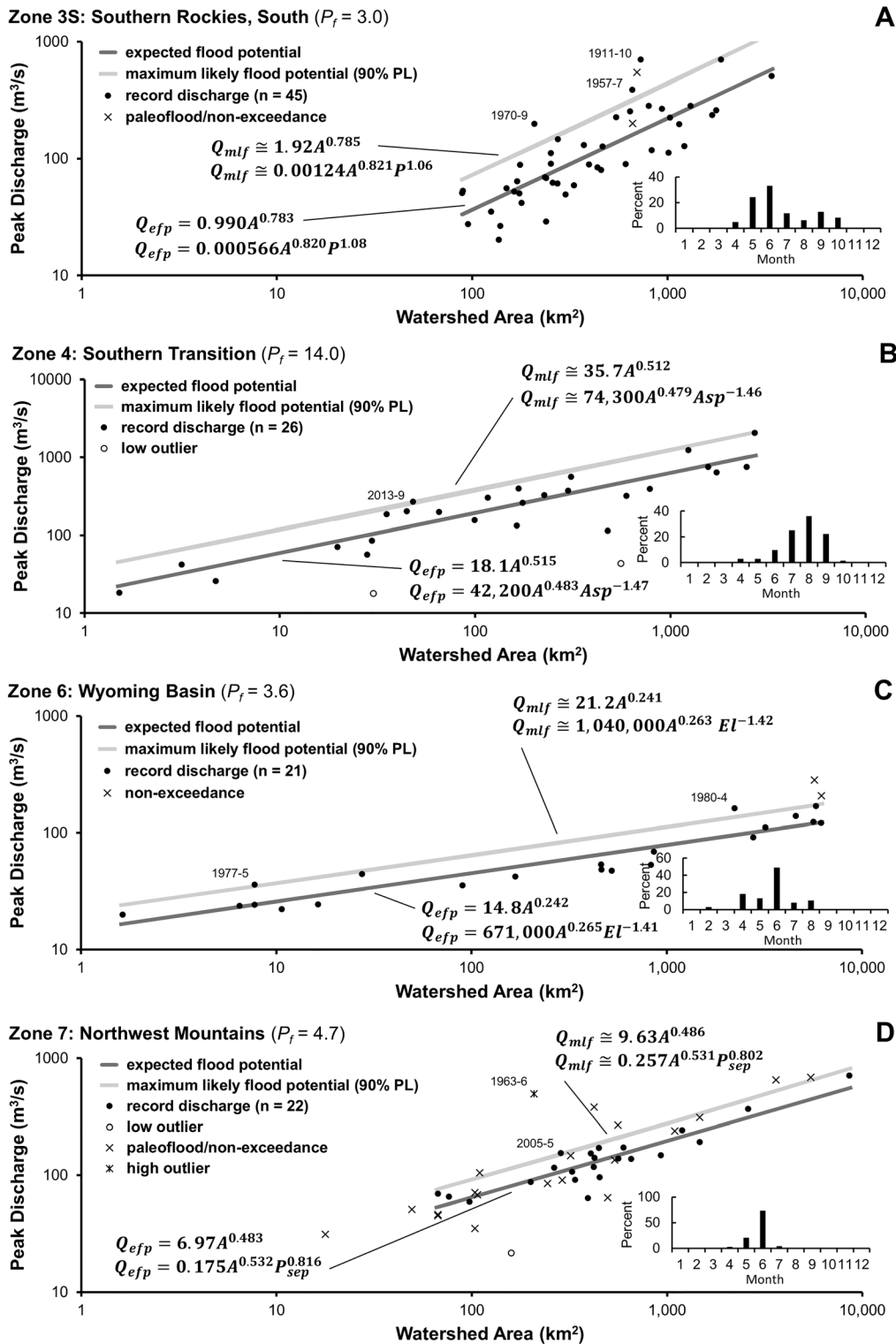


Figure 4. Peak discharge versus watershed area, for Zones 3S, 4, 6, and 7. P_f is the flood potential index. Black dots indicate the record peak discharges for each streamgage, which were used to develop the expected flood potential regression (dark gray) and the maximum likely flood potential (light gray, 90% prediction limit), with associated prediction equations provided. Where topographic and climate variables were significant in multivariable regression, a second set of equations is provided. Excluded low outliers are noted with circles. Where available, paleoflood and nonexceedance flood estimates (marked with x) are also shown. Months of peak flow occurrence for the highest 5% of floods are also indicated. Q_{efp} : expected flood potential discharge; Q_{mlf} : maximum likely flood potential discharge; A : watershed area; P : watershed mean annual precipitation (mm); Asp : watershed average aspect (deg); El : watershed average elevation (m); P_{sep} : watershed mean annual precipitation for September (mm).

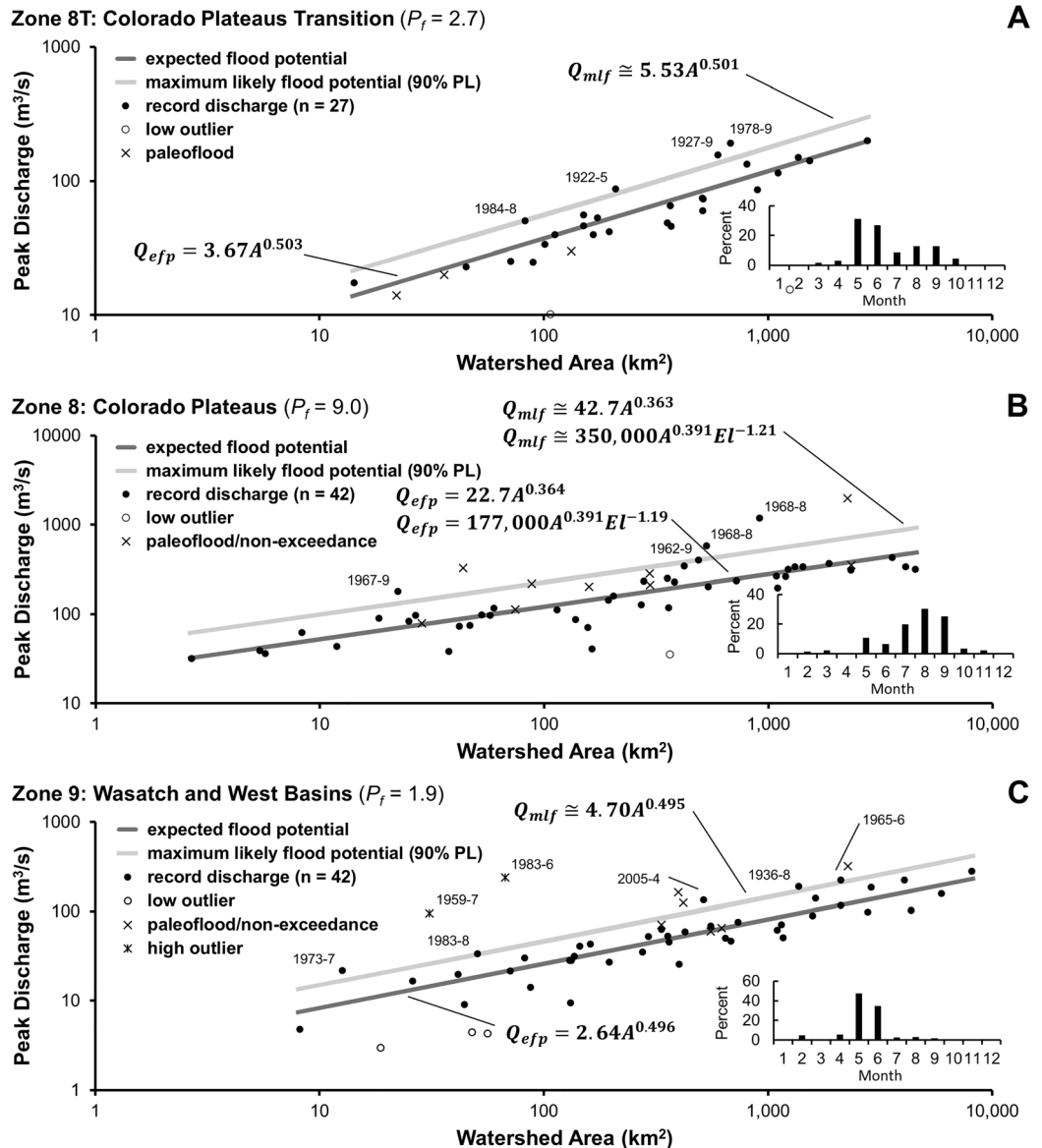


Figure 5. Peak discharge versus watershed area, for Zones 8T, 8, and 9. P_f is the flood potential index. Black dots indicate the record peak discharges for each streamgage, which were used to develop the expected flood potential regression (dark gray) and the maximum likely flood potential (light gray, 90% prediction limit), with associated prediction equations provided. Where topographic and climate variables were significant in multivariable regression, a second set of equations is provided. Excluded low outliers are noted with circles. Where available, paleoflood and nonexceedance flood estimates (marked with x) are also shown. Months of peak flow occurrence for the highest 5% of floods are also indicated. Q_{efp} : expected flood potential discharge; Q_{mLf} : maximum likely flood potential discharge; A : watershed area; El : watershed average elevation (m).

3.10. Extreme Floods

Using the maximum likely flood potential, as defined by the 90% prediction limit (with drainage area as the sole predictor), extreme floods were identified using a quantifiable approach for each zone (Table 1 and Figures S3 to S10). Floods with greater departure from the maximum likely flood potential are more extreme. Hence, this method provides a systematic approach for defining and ranking extreme floods, on a relative basis for each zone. Using this approach, 9% of the watersheds in the study area have experienced floods that are considered extreme (Figure 2), though, considering the relative zonal basis of the computations, what is extreme greatly varies. To illustrate this variation, the expected flood potential discharge (Q_{efp}) and the

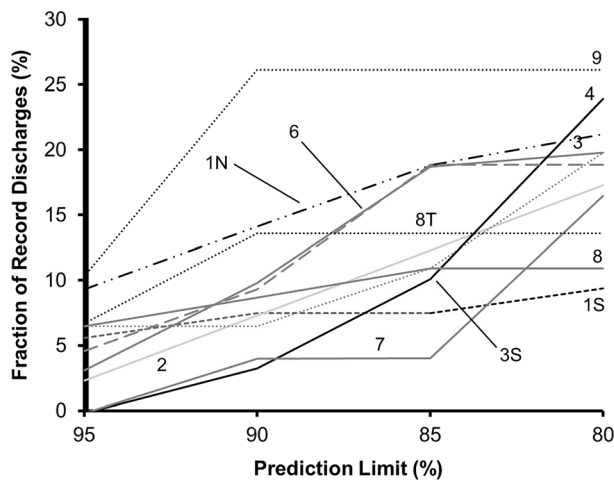


Figure 6. Fraction of record discharges considered extreme with the use of different prediction limits, with zone identifications. A heuristic decision was made that the 90% prediction limit was the most appropriate for use as a threshold, due to inflections at this point.

maximum likely flood discharge (Q_{mlf}) for a standard 1,000-km² watershed area are provided in Table 1. Q_{mlf} varies from 98 (Zone 2) to 2,000 m³/s (Zone 3S). This variation may be due to regional watershed characteristics, such as bedrock exposure, thin soils, vegetative conditions, and steep relief (O'Connor & Costa, 2004; Osterkamp & Friedman, 2000), as well as due to the dominant zonal flood type (rainfall versus snowmelt) along with variability in orographic blockage, water vapor sources, rainfall rates, and convective storm sizes across the study area.

The most extreme floods in the analysis extent have occurred along the southern portion of the eastern slopes of the Southern Rocky Mountains and the adjacent Great Plains (Zone 1S) in May 1935 and June 1965 (Jimmy Camp Creek, CO; Kiowa Creek, CO; and Plum Creek, CO); these floods ranged in magnitude from 12.7 to 6.7 times the expected flood potential for these specific watersheds. Other zones that experienced the most significant extreme floods include Zone 1N (Sand Creek, WY, 1955-8; Spring Creek, CO, 1997-7; and Big Thompson River, CO, 1976-7), Zone 3 (Sweetwater Creek, CO, 1976-7), Zone 3S (San Juan River, CO, 1911-10), Zone 8 (Cottonwood Wash, UT, 1968-8), and Zone

9 (Sulphur Creek, UT, 1983-6; tributary to the Green River, UT, 1959-7), with magnitudes ranging from 4.1 to 1.7 times the expected flood potential.

The finding that the Zone 1S floods were remarkable, as indicated by magnitudes of up to 12.7 times the expected flood potential (substantially greater than the extreme floods experienced in other zones of the analysis extent), is supported by recent work that highlighted this area as having experienced unusually large floods (strange floods; Smith et al., 2018). These floods were linked with large-scale synoptic events, with the 1965 flooding associated with an intense cutoff western low that steered warm, moist, unstable air into Eastern Colorado and a blocking pattern forcing a cold front to be stationary for 3 days, inducing extreme rains (Hirschboeck, 1987). A similar synoptic pattern induced the 2013 Colorado Front Range flood, which occurred over a 4–5-day period (Gochis et al., 2015). The 1935 flooding in Colorado was associated with extreme flooding on the High Plains (Republican River) and in Central Texas, which experienced a world record rainfall rate of 559 mm in 2.75 hr (Smith et al., 2018). These extreme floods, with high multipliers of an already high expected flood potential (see next section), “tend to be distinguished from more common regional floods by the anomalous behavior of the macroscale circulation patterns that drive and steer flood-generating synoptic weather systems” (Hirschboeck, 1987).

There may be spatial clustering of extreme floods within the analysis extent (Figures 2 and S3 to S10). For example, there appears to be an uncommon number of extreme floods experienced along the northern portion of the Colorado Front Range and High Plains (floods of 1976, 1997, and 2013), along the southern faces of the San Juan Mountains, and along the northern foothills and adjacent basins of the Uinta Mountains as well as the northern faces of the northern portion of the Southern Rockies (in the vicinity of Elk Mountain). Some of these extreme floods are likely the result of convection and heavy rainfall during the summer monsoon, but spring snowmelt or rain-on-snow events may be the drivers of extreme floods for those events outside the normal summer monsoon. Previous work has also indicated clustering of the most substantial floods across the western half of the United States, including along the Front Range of the Rocky Mountains (Zones 1S and 1N), and within portions of the High Plains, Black Hills, Columbia Plateau, Olympic Mountains, Southern California, and Texas regions (O'Connor & Costa, 2004; Smith et al., 2018; Smith & Smith, 2015). These patterns may be suggestive of dominant large flood-generating mechanisms within flood potential zones. Generally, extreme floods tend to be associated with intense thunderstorms, squall lines, and shortwave troughs developed within or influenced by synoptic-scale weather systems, and are often associated with atmospheric blocking patterns (Hirschboeck, 1987) that can result in stalled frontal boundaries and semi-stationary cut-off low pressure systems.

It is important to keep in mind that the magnitude of some large floods should not be considered extreme but expected given the conglomeration of record peak discharges that have been measured in each zone

(Figures 3–5). For example, while the 2013 Front Range flood (zone 1N) is generally considered extreme for much of its extent within and along the foothills of the Northern Colorado Front Range, the presented method indicates that other than in portions of the St. Vrain and Little Thompson watersheds (Figure S4), this flooding was not extreme in magnitudes but rather similar or lesser to the expected flood potential over most of the flood extent (Figure 3). Hence, while a landowner or land manager in a canyon on the Front Range may not appreciate the risk associated with living or placing infrastructure within a flood-susceptible area, this methodology can help illuminate that the danger is very real in any given year, and, in fact, happened in nearby watersheds or may occur again in this canyon at an even larger magnitude, with potentially devastating impacts.

3.11. Flood Indices

Flood indices are valuable for comparing flood risk, to understand variability within regions and across continents. Methods that rely on flood frequency analyses and regional regression studies do not facilitate comparison of flood magnitudes across wide geographic extents; the indices presented here ease comparisons, to help users understand flood variability and subsequent impacts regarding such management issues as wildfire-induced flooding, erosion hazards to infrastructure, and the stability of stream restorations. We computed four indices: the flood potential index, flood variability index, a flashiness index, and the product of the flood potential and flashiness indices, a flood hazard index (Figures 2 and 7 and Table 1).

The flood potential index (P_f) was developed for between-zone comparison of expected flood magnitudes using fixed watershed sizes of 20, 200, and 2,000 km². For each zone, this simple method compares expected flood potential to the zone with the smallest flood magnitudes (Orographic Sheltered; Zone 2). This index varied from 15 to 1.0, with the highest flood potential zone (eastern slopes and Great Plains, south; Zone 1S) experiencing floods, on average for a given watershed area, 15 times greater than the index zone (2), and with high flood potential also in the Southern Transition (4), eastern slopes and Great Plains, North (1N), and Colorado Plateaus (8) zones. These areas are susceptible to large rainfall-induced floods and are of particular concern due to occasional large flood extents, depths, and velocities that are experienced, as well as enhanced risk of geomorphic adjustment of near-stream areas (Yochum et al., 2017). In addition to being able to compare average flood sizes to the index zone (2), the flood potential index also allows easy comparison between any zones. For example, Zone 1N ($P_f = 13.8$) can be compared to Zone 3 ($P_f = 2.3$) by dividing $13.8/2.3 = 6.0$ —the northern portion of the eastern slopes and Great Plains on average for a given watershed area experiences floods that are 6 times greater than the adjacent Southern Rockies. Therefore, the use of a consistent reference zone allows simple continental-scale comparison of average flood magnitudes.

The flood variability index describes the within-zone spatial variability of large floods (Table 1), with higher index values indicating greater spatial variability within the zone. Zone 1S has experienced the greatest variability in large floods for any given watershed size ($V_f = 2.77$). Generally, the southern and eastern portions of the analysis extent have experienced the greatest variability in large flood magnitudes (V_f ranging from 2.77 to 1.88), which are zones where the largest floods are rainfall events. The lowest variability in flood magnitudes have occurred in Zone 7, the Northwest Mountains ($V_f = 1.38$), due to snowmelt floods.

The Beard flash flood index (F ; Beard, 1975) was utilized to quantify flashiness, with a greater difference between the magnitude of large floods and more typical annual floods resulting in more hazardous conditions. The index was computed for each streamgaged watershed (Figure S11) and averaged for each zone, with these averages varying from 1.30 to 0.49 (Table 1). However, it is important to note that this flashiness index does not directly quantify the steepness of rapidly rising flood waves but instead quantifies the variability in the annual peak discharges, with higher index values indicating the largest floods are much greater than more typical annual peak discharges. The relationship between these two types of flashiness is unknown.

The product of P_f and F is a flood hazard index (Table 1 and Figure 7), which accounts for both flood magnitudes and flashiness; this index varies from 19.5 (Zone 3S) to 0.60. The zones with greatest flood hazards include the urban areas along the Colorado Front Range, including Fort Collins, Boulder, Denver, and Colorado Springs, as well as the cities of Casper and Cheyenne (WY), Pueblo and Trinidad (CO), Santa Fe and Farmington (NM), and Moab (UT). The lowest flood hazard within the analysis extent is in the

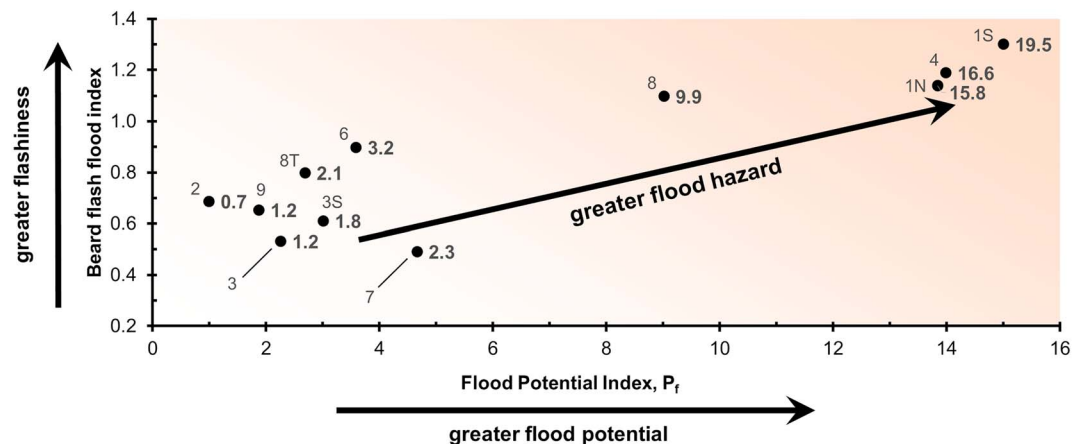


Figure 7. Average zone flood potential, flashiness, and flood hazard index values for the greater Southern Rocky Mountain region. The values to the right of each labeled point are the flood hazard index values, with the smaller font values indicating the zone identification.

Orographic Sheltered Zone 2, with these mountain valleys sheltered from large flood producing meteorological processes and the primary moisture sources of the East Pacific Ocean and the Gulf of Mexico.

3.12. Paleofloods

Paleoflood and nonexceedance bound flood data were obtained for most of the zones in the analysis extent. Where available, both paleoflood and nonexceedance floods are plotted for each zone (Figures 3–5), with these points generally within the cloud of streamgage points and the points that plot high (with respect to the streamgage data) being (conservative) nonexceedance values.

For example, Zone 1S had 21 paleoflood and nonexceedance bound floods estimated, at a wide variety of watershed scales (Figure 3a). These floods are typically between the expected flood potential and the maximum likely flood potential, with none of these floods surpassing the extreme floods of 1965 and 1935. In Zone 3S, two paleoflood and nonexceedance floods were estimated (Figure 4A), with a paleoflood estimate plotting a bit above the maximum likely flood potential and a nonexceedance flood plotting as extreme, though less so than the San Juan River in October 1911. Additionally, Zone 7 has 16 paleoflood and 3 nonexceedance floods (Figure 4D), with one of these estimates plotting high in the extreme, though less so than the Whiterocks River, UT flood (2005–5).

Inclusion of the paleoflood data in regressions (while excluding the nonexceedance flood estimates) generally provide fits that are slightly higher quality to the streamgage only regressions (Table 1). Inclusion typically has minimal effects, likely due to the small amount of available data. Regardless, there are additional benefits to predictions; for example, the incorporation of paleoflood data can allow for the extension of flood potential predictions to smaller watersheds than the minimum streamgage watershed size (such as in Zone 7).

These results generally support the hypothesis that the unbounded assumption of at-a-station relationships of flood frequency is inappropriate and that there is likely an upper limit to flood magnitudes due to physical limits in precipitation and watershed responses, as asserted by Wolman and Costa (1984), Costa (1987), and Enzel et al. (1993). Probable maximum precipitation studies also provide upper limits on precipitation and, hence, flooding; this work can help inform such studies.

3.13. Flood Frequency Comparison

Published log-Pearson-estimated 100-year flood magnitudes were obtained (where available) from regional flood frequency studies (Barenbrock, 2002; Capesius & Stephens, 2009; Kohn et al., 2016; Waltemeyer, 2008) and compared to expected flood potential values for 147 streamgages (Figure 8). All zones besides 1S and 6 were represented with at least one streamgage. Q_{efp} and Q_{100} are linearly related ($R^2 = 0.8$, $p < 2e-16$), and the means of the two distributions are not statistically different (p

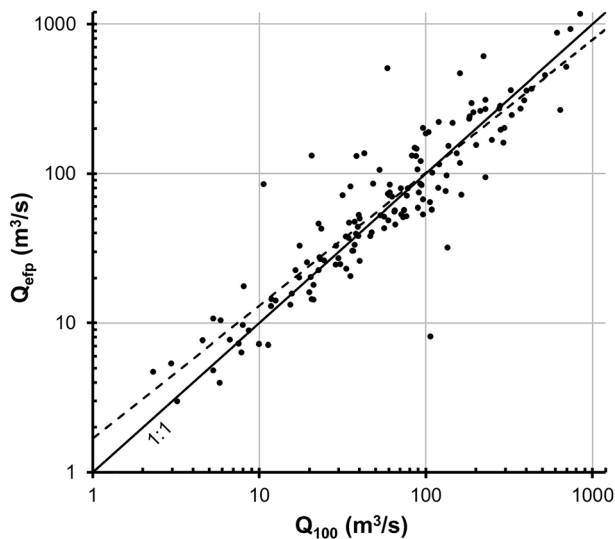


Figure 8. Effective flood potential discharge (Q_{effp}) versus 100-year discharge (Q_{100}) as computed using a log-Pearson analysis of streamgage records, for locations where log-Pearson discharges were available from regional flood frequency studies. The 1:1 line (solid) is provided for comparison to the trendline (dashed).

$= 0.5 \gg \alpha$). Additional analysis were performed for those gages with record length greater and less than 30, 40, and 50 years—in no case was there a statistically significant difference between the means of the two distributions. This result did not change when performed on subsets of the data defined by zone, indicating that there was no statistical evidence for a difference between Q_{effp} and Q_{100} within individual zones.

4. Summary and Conclusions

Greater understanding of flood hazards, and how they vary spatially and temporally across regions and continents, is needed to protect lives, property, and infrastructure and for developing more resilient communities. Additionally, strategic yet simple language for communicating expected flood hazards would be valuable to help citizens and policy makers understand risk. Using the greater Southern Rocky Mountains as a study region, a new method was developed to rank, compare, and predict flood magnitudes and hazard. Key questions that this new methodology helps to answer include the following:

1. Given the record of floods within a derived area, what flood magnitudes can be expected for a specific watershed?
2. What flood magnitudes are most properly referred to as extreme, and what floods are the most extreme?
3. How do expected flood magnitudes, flashiness, and overall hazard in a given area compare to other areas?

Utilizing record peak discharges at longer term streamgages, regressions were performed using watershed area and topographic and climatologic descriptors as predictors for 11 derived zones of similar flood potential. Up to 93% of the variance in the assembled datasets was explained by the regression models. Each of these regressions define the *expected flood potential* of a zone, to help understand what flood magnitudes can be expected for any given watershed within the derived watershed area ranges. This space-for-time substitution allows users to understand the magnitude of flooding that can be expected given the streamgage records within nearby watersheds. This method sidesteps a few issues associated with flood frequency and regional regression analyses, including bimodal and mixed flood peak distributions, complications induced by flood periodicity (a preponderance of floods for several years, followed by a dearth of floods in other years), and problems induced by wide ranging variability in streamgage record periods and lengths. The method also avoids problems associated with rainfall-runoff analyses, including limitations on the understanding of rainfall frequency and distributions (especially in mountainous settings), and the proper simulation of runoff processes. Utilizing flood potential analyses provides a third method for understanding what size floods can be expected for a given watershed.

This paper introduces an additional set of tools for understanding flooding, including defining and ranking extreme floods across zones and comparing and ranking flood potential, flashiness, and general flood hazard across regions and, potentially, continents. Tendencies in flood seasonality, periodicity, and trends for each zone in the greater Southern Rocky Mountains region were also assessed, using the largest 5% floods. Additionally, paleoflood data were combined with the streamgage data, to provide a longer temporal perspective on flood magnitudes and variability. The developed indices simplify comparisons and enhance communication. For example, the zone with the highest flood potential (1S, eastern slopes and Great Plains, South) experiences flood magnitudes that are, on average for a given watershed area, 15 times greater than the adjacent index zone (2, rain-shadowed, high-elevation central Colorado and New Mexico mountain-surrounded valleys)—there is a very wide range in flood magnitudes experienced across this region.

The developed methodologies may also be applicable in others regions, with preliminary analyses indicating the methods are applicable in portions of the New England, Southern Midwest, Gulf Coast, and West Coast

regions of the continental United States, as well as in Puerto Rico. Hence, this method shows promise for enhancing insight into the expected magnitude and spatial variation of floods (and consequently, expected flood extents), as well as for helping to understand variability in forcing mechanisms in regard to geomorphic form and erosion hazards of streams and floodplains, the inherent risk of stream restoration in a given area, the relative risk for flooding and water quality problems after wildfires, and the variability in probable maximum precipitation and floods.

4.1. Future Research Needs

A number of associated research needs have been illuminated by this work, specifically:

1. The presented methodology should be applied in other regions, to assess appropriateness and effectiveness.
2. Investigate methodologies for identification of flood potential zone boundaries. Such methods may be diverse, including numerical methods or methods that rely on geomorphic indicators between high and low flood potential zones (including the use of existing soil survey data, such as alluvial fans).
3. Perform a more comprehensive investigation of flood peak trends and periodicity, including a quantitative assessment of more and less active groupings of flood years.
4. Investigate the causes of the most extreme floods identified in this study area. The mechanisms behind these largest floods can illuminate future threats.
5. More effort is needed to collect paleoflood data and add measurements of large floods, to increase understanding of the spatial variability and causative mechanisms of flooding.
6. Compare flood flashiness as computed by different methodologies, to understand if simpler methods that can utilize the entire streamgage record (such as Beard F) can be used in place of more complex methods that rely on continuous streamgage records for measuring rates of hydrograph change.
7. Evaluate methods for potentially assigning frequencies to expected flood potential values.
8. For prediction in areas where watersheds are in multiple zones, develop tools for distributing predictions across zone boundaries.
9. Link expected flood potential with top-down inundation hydraulic modeling (Wing et al., 2018), to assess flood risk over large geographic areas.
10. Link expected flood potential to geomorphic change estimation techniques (Yochum et al., 2017), to predict the risk for valley adjustment during large floods.

Acknowledgments

Review of this manuscript by Tao Liu and Victor Baker, and two anonymous reviewers, provided a great deal of value for increasing the conciseness and clarity of this manuscript. Their efforts and dedication to peer review are greatly appreciated. Additionally, Paul Patterson and Sara Goeking are appreciated for their statistical thoughts and Victoria Stempniewicz is valued for checking the accuracy of the prediction equations. The many hydrologists, hydraulic engineers, and technicians who have collected and analyzed peak stage and discharge data for more than a century are greatly valued for their contributions to flood science. Additionally, the long-term commitment of the U.S. Geological Survey, the Colorado Division of Water Resources, the Wyoming State Engineers office, the U.S. Bureau of Reclamation, and their funding partners, is greatly appreciated; the vision and commitment for collecting streamgage data established the foundation of flood hydrology. Paleoflood and nonexceedance bound flow estimates from an internal U.S. Bureau of Reclamation database were also valuable; Jeanne Godaire is appreciated for providing these data for the study region. Supporting data for the analyses are provided as supporting information.

References

- Adams, D. K., & Comrie, A. C. (1997). The North American monsoon. *Bulletin of the American Meteorological Society*, 78(10), 2197–2213. [https://doi.org/10.1175/1520-0477\(1997\)078<2197:TNAM>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2)
- Asquith, W. H., & Slade, R. M. (1995). *Documented and potential extreme discharges and relation between potential extreme peak discharges and probable maximum flood peak discharges in Texas*. U.S. Geological Survey, Water-Resources Investigations Report 95-4249.
- Baker, D. B., Richards, R. P., Leftus, T. T., & Kramer, J. W. (2004). A new flashiness index: Characteristics and applications to Midwestern rivers and streams. *Journal of the American Water Resources Association*, 40(2), 503–522. <https://doi.org/10.1111/j.1752-1688.2004.tb01046.x>
- Baker, V. R. (1994). Geomorphological understanding of floods. *Geomorphology*, 10(1-4), 139–156. [https://doi.org/10.1016/0169-555X\(94\)90013-2](https://doi.org/10.1016/0169-555X(94)90013-2)
- Baker, V. R. (1998). Hydrological understanding and societal action. *Journal of the American Water Resources Association*, 34(4), 819–825. <https://doi.org/10.1111/j.1752-1688.1998.tb01518.x>
- Baker, V. R. (2008). Paleoflood hydrology: Origin, progress, prospects. *Geomorphology*, 101(1-2), 1–13. <https://doi.org/10.1016/j.geomorph.2008.05.016>
- Barenbrock, C. (2002). Estimating the magnitude of peak flows at selected recurrence intervals for streams in Idaho. U.S. Geological Survey, Water-Resources Investigations Report 02-4170.
- Beard, L. R. (1975). *Generalized evaluation of flood potential*, (Vol. CRWW-124, pp. 1–27). University of Texas, Austin: Center for Research in Water Resources.
- Bryson, R. A., & Hare, F. K. (1974). The climates of North America. In E. A. Bryson & F. K. Hare (Eds.) *Climates of North America: World Survey of Climatology* (Vol. 11, pp. 1–47). Amsterdam, Elsevier.
- Brogan, D. J., Nelson, P. A., & MacDonald, L. H. (2017). Reconstructing extreme post-wildfire floods: A comparison of convective and mesoscale events. *Earth Surface Processes and Landforms*, 42(15), 2505–2522. <https://doi.org/10.1002/esp.4194>
- Capesius, J. P., & Stephens, V. C. (2009). Regional regression equations for estimation of natural streamflow statistics in Colorado. U.S. Geological Survey Scientific Investigations Report 2009–5136.
- Castellarin, A., Vogel, R. M., & Matalas, N. C. (2005). Probabilistic behavior of a regional curve. *Water Resources Research*, 41, W06018. <https://doi.org/10.1029/2004WR003042>
- Cayan, D. R., Redmond, K. T., & Riddle, L. G. (1999). ENSO and hydrologic extremes in the western United States. *Journal of Climate*, 12, 2881–2893.
- Colorado Division of Water Resources (2017). Colorado's surface water conditions. <http://www.dwr.state.co.us/Surfacewater/default.aspx>.

- Costa, J. E. (1987). A comparison of the largest rainfall-runoff floods in the United States with those of the People's Republic of China and the world. *Journal of Hydrology*, 96(1-4), 101–115. [https://doi.org/10.1016/0022-1694\(87\)90146-6](https://doi.org/10.1016/0022-1694(87)90146-6)
- Crippen, J. R., & Bue, C. D. (1977). Maximum floodflows in the conterminous United States. Geological Survey Water Supply Paper 1887.
- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*, 28: 2031–2064, 15, <https://doi.org/10.1002/joc.1688>
- DNR OSE (2018). Colorado–New Mexico regional extreme precipitation study. Colorado Division of Water Resources, Dam Safety Branch, New Mexico Office of the State Engineer, Same Safety Bureau. <http://water.state.co.us/damsafety/dams.asp>.
- Dollman, D. S. (2017). *Colorado's deadliest floods*. Charleston, SC: The History Press.
- Ely, L. L., Enzel, Y., Baker, V. R., & Cayan, D. R. (1993). A 5000-year record of extreme floods and climate change in the southwestern United States. *Science*, 262(5132), 410–412. <https://doi.org/10.1126/science.262.5132.410>
- England, J. F., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas, W. O., Veilleux, A. G., et al. (2018). Guidelines for determining flood flow frequency: Bulletin 17C. U.S. Department of Interior, U.S. Geological Survey, Techniques and Methods 4-B5.
- Enzel, Y., Ely, L. L., House, P. K., Baker, V. R., & Webb, R. H. (1993). Paleoflood evidence for a natural upper bound to flood magnitudes in the Colorado River Basin. *Water Resources Research*, 29(7), 2287–2297. <https://doi.org/10.1029/93WR00411>
- Fenneman, N. M. (1931). *Physiography of the western United States*. New York, London: McGraw-Hill Book Co.
- Follansbee, R., & Sawyer, L. R. (1948). Floods in Colorado. U.S. Department of Interior, Geological Survey, Water Supply Paper 997.
- Gochis, D., Schumacher, R., Friedrich, K., Doesken, N., Kelsch, M., Sun, J., et al. (2015). The Great Colorado Flood of September 2013. *Bulletin of the American Meteorological Society*, 96(9), 1461–1487. <https://doi.org/10.1175/BAMS-D-13-00241.1>
- Godaire, J. (2018). personal communication, U.S. Bureau of Reclamation. Email dated 2018-8-31.
- Godaire, J., Caldwell, J., Novembre, N., Harden, T., & Sankovich, V. (2013). Extreme floods in a changing climate. U.S. Bureau of Reclamation, Dam Safety technology Development Program, Report DSO-2013-02.
- Gourley, J. J., Hong, Y., Flamig, Z. L., Arthur, A., Clark, R., Calianno, M., et al. (2013). A unified flash flood database across the United States. *Bulletin of the American Meteorological Society*, 799–805. <https://doi.org/10.1175/BAMS-D-12-00198.1>
- Grigg, N. S., Doesken, N. J., Frick, D. M., Grimm, M., Hilmes, M., McKee, T. B., & Oltjenbruns, K. A. (1998). Fort Collins Flood 1997: Lessons from an extreme event. Colorado State University Water Center, Water Center Paper No. 98-1.
- Hansen, E. M., Fenn, D. D., Schreiner, L. C., Stodt, R. W., & Miller, J. F. (1988). Probable maximum precipitation estimates—United States between the continental divide and the 103rd meridian. National Oceanic and Atmospheric Administration, Hydrometeorological Report No. 55A.
- Herschey, R. W. (2002). The world's maximum observed floods. *Flow Measurement and Instrumentation*, 13(5-6), 231–235. [https://doi.org/10.1016/S0955-5986\(02\)00054-7](https://doi.org/10.1016/S0955-5986(02)00054-7)
- Hirschboeck, K. K. (1987). In L. Mayer, & D. B. Nash (Eds.), *Catastrophic flooding and atmospheric circulation anomalies*, (pp. 23–56). Allen & Unwin: Catastrophic Flooding.
- IACWD (1982). Flood flow frequency: Bulletin #17B of the hydrology subcommittee. U.S. Department of Interior, Geological Survey, Office of Water Data Coordination, Interagency Advisory Committee on Water Data.
- Jarrett, R. D. (1990). Paleohydrologic techniques used to define spatial occurrence of floods. *Geomorphology*, 3(2), 181–195. [https://doi.org/10.1016/0169-555X\(90\)90044-Q](https://doi.org/10.1016/0169-555X(90)90044-Q)
- Jarrett, R. D., & Tomlinson, E. M. (2000). Regional interdisciplinary paleoflood approach to assess extreme flood potential. *Water Resources Research*, 36(10), 2957–2984. <https://doi.org/10.1029/2000WR900098>
- Kenney, T. A., Wilkowske, C. D., & Wright, S. J. (2007). Methods for estimating magnitude and frequency of peak flows for natural streams in Utah. U. S Geological Survey Scientific Investigations Report 2007–5158.
- Kimrough, R. A., & Holmes, R. R. (2015). Flooding in the South Platte River and Fountain Creek Basins in Eastern Colorado, September 9–18, 2013. U.S. Department of Interior, U.S. Geological Survey, Scientific Investigations Report 2015-5119.
- Klemes, V. (1986). Dilettantism in hydrology: Transition or destiny? *Water Resources Research*, 22(9S), 177S–188S. <https://doi.org/10.1029/WR022i09Sp0177S>
- Klemes, V. (1989). The improbable probabilities of extreme floods and droughts. In M. Molder (Ed.), *Hydrology of disasters*, 0. *Starosolszky and 0*, (pp. 43–51). London, England: James and James.
- Kohn, M. S., Stevens, M. R., Harden, T. M., Godaire, J. E., Klinger, R. E., & Mommandi, A. (2016). Paleoflood investigations to improve peak-streamflow regional-regression equations for natural streamflow in Eastern Colorado, 2015. U.S. Geological Survey Scientific Investigations Report 2016-5099.
- Levish, D. R. (2002). Paleohydrologic bounds: Non-exceedance information for flood hazard assessment. In P. K. House, R. H. Webb, V. R. Baker, & D. R. Levish (Eds.), *Ancient floods, modern hazards: Principles and practices of paleoflood hydrology*, *Water Science and Application*, (Vol. 5, pp. 175–190). Washington, D.C.: American Geophysical Union.
- McMahon, G. M., & Kiem, A. S. (2018). Large floods in South East Queensland, Australia: Is it valid to assume they occur randomly? *Australian Journal of Water Resources*, 22(1), 4–14. <https://doi.org/10.1080/13241583.2018.1446677>
- Michaud, J. D., Hirschboeck, K. K., & Winchell, M. (2001). Regional variations in small-basin floods in the United States. *Water Resources Research*, 37(5), 1405–1416. <https://doi.org/10.1029/2000WR900283>
- Miller, K. A. (2003). Peak-flow characteristics of Wyoming streams. U. S Geological Survey Water-Resources Investigations Report 03-4107.
- Mimikou, M., & Gordios, J. (1989). Predicting the mean annual flood and flood quantities for ungauged catchments in Greece. *Hydrological Sciences Journal*, 34(2), 169–184. <https://doi.org/10.1080/02626668809491322>
- Moody, J. A. (2016). Estimates of peak discharge for 21 sites in the Front Range in Colorado in response to extreme rainfall in September 2013. U.S. Geological Survey Scientific Investigations Report 2016-5003, doi:10.3133/sir20165003
- Moore, D., Todea, N., Cerrelli, G., Yochum, S., Norman, J.B., & Hoeft, C. (2016). Hydrologic analyses of post-wildfire conditions. U.S. Department of Agriculture, Natural Resources Conservation Service, Hydrology Technical Note No. 4.
- Moosburner, O. (1970). A proposed streamflow-data program for Arizona. U.S. Geological Survey, Open File Report 70-231.
- Natural Environment Research Council (1975). *Flood Studies Report*, 5 Volumes. London: Natural Environment Research Council.
- 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, W01107. <https://doi.org/10.1029/2003WR002247>
- 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](https://doi.org/10.1002/1099-1085(200011/12)14:16/17<2817::AID-HYP121>3.0.CO;2-B)

- Patton, P. C., & Baker, V. R. (1976). Morphometry and floods in small drainage basins subject to diverse hydrogeomorphic controls. *Water Resources Research*, 12(5), 941–952.
- Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegard, K. L., Richter, B. D., et al. (1997). The natural flow regime. *Bioscience*, 47(11), 769–784. <https://doi.org/10.2307/1313099>
- PRISM (2018). Climate Group, Northwest alliance for Computational Science and Engineering, 30-year normals. <http://www.prism.oregonstate.edu/normals/> (accessed August 2018).
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>
- Ralph, F. M., Dettinger, M., Cordeira, J. M., Rutz, J. J., Schick, L., Anderson, M., et al. (2019). A scale to characterize the strength and impacts of atmospheric rivers. *Bulletin of the American Meteorological Society*, 100(2), 269–289.
- Reclamation Water Information System (2017). <https://water.usbr.gov/query.php>, (accessed August 2017).
- Rutz, J. J., Steenburgh, W. J., & Ralph, F. M. (2014). Climatological characteristics of atmospheric rivers and their inland penetration over the western United States. *Monthly Weather Review*, 142(2), 905–921. <https://doi.org/10.1175/MWR-D-13-00168.1>
- Rutz, J. J., Steenburgh, W. J., & Ralph, F. M. (2015). The inland penetration of atmospheric rivers over western North America: A Lagrangian analysis. *Monthly Weather Review*, 143(5), 1924–1944. <https://doi.org/10.1175/MWR-D-14-00288.1>
- Saharia, M., Kirstetter, P. E., Vergara, H., Gourley, J. J., & Hong, Y. (2017). Characterization of floods in the United States. *Journal of Hydrology*, 548, 524–535. <https://doi.org/10.1016/j.jhydrol.2017.03.010>
- Saharia, M., Kirstetter, P. E., Vergara, H., Gourley, J. J., Hong, Y., & Giroud, M. (2017). Mapping flash flood severity in the United States. *American Meteorological Society*, 397–411. <https://doi.org/10.1175/JHM-D-16-0082.1>
- Schram, H., Wulliman, J. T., & Rapp, D. N. (2014). Hydrologic evaluation of the St. Vrain Watershed: Post September 2013 flood event. Prepared for the Colorado Department of Transportation, Region 4 Flood Recovery Office. Jacobs Engineering, Denver, Colorado.
- Schwarz, F.K. (1967). The role of persistence, instability, and moisture in the intense rainstorms in Eastern Colorado, June 14–17, 1965. U.S. Department of Commerce, Weather Bureau, Office of Hydrology, Technical Memorandum WBTM HYDRO-3.
- Serinaldi, F. (2015). Dismissing return periods! *Stochastic environmental resources risk assessment*, 29(4), 1179–1189. <https://doi.org/10.1007/s00477-014-0916-1>
- Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Muche, M., & Avant, B. (2017). An overview of rainfall-runoff types. U.S. Environmental Protection Agency, Office of Research and Development. EPA/600/R-14/152.
- Smith, B. K., & Smith, J. A. (2015). The flashiest watersheds in the Contiguous United States. *American Meteorological Society*, 2365–2381. <https://doi.org/10.1175/JHM-D-14-0217.1>
- Smith, G. E. (2010). Development of a flash flood potential index using physiographic data sets within a geographic information system. Masters thesis. Department of Geography, The University of Utah.
- Smith, J. A., Cox, A. A., Baack, M. L., Yang, L., & Bates, P. (2018). Strange floods: The upper tail of flood peaks in the United States. *Water Resources Research*, 54, 6510–6542. <https://doi.org/10.1029/2018WR022539>
- State Engineers Office (2017). AQUARIUS WebPortal 2017. <http://seoflow.wyo.gov/> (retrieved August 2017).
- Tipton, R. J. (1937). Characteristics of floods in the Southern Rocky Mountain Region. *Transactions of the American Geophysical Union*, 18(2), 592–600. <https://doi.org/10.1029/TR018i002p00592-2>
- U.S. Geological Survey (2017a). Peak streamflow for the nation. Retrieved from <https://nwis.waterdata.usgs.gov/usa/nwis/peak> (accessed August 2017).
- U.S. Geological Survey (2017b). Watershed boundary dataset. Retrieved from <https://nhd.usgs.gov/wbd.html> (accessed December 2017).
- U.S. Geological Survey (2018). National elevation datasets. Retrieved from <https://nhd.usgs.gov/wbd.html> (accessed August 2018).
- Urban Drainage and Flood Control District. (UDFCD) (2014). 2013 flood peak estimates. Data accessed 7/14/2014. <http://udfcd.maps.arcgis.com/home/gallery.html>
- Vogel, R. M., Matalas, N. C., England, J. F., & Castellarin, A. (2007). An assessment of exceedance probabilities of envelope curves. *Water Resources Research*, 43, W07403. <https://doi.org/10.1029/2006WR005586>
- Waltemeyer, S. D. (2008). Analysis of the magnitude and frequency of peak discharge and maximum observed peak discharge in New Mexico and surrounding areas. U. S Geological Survey Scientific Investigations Report 2008-5119.
- Ward, P. J., Jongman, B., Kumm, M., Dettinger, M. D., Weiland, F. C. S., & Winsemius, H. C. (2014). Strong influence of El Nino Southern Oscillation on flood risk around the world. *Proceedings National Academies of Science*, 111, 15,659–15,664.
- Wing, O. E. J., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Fargione, J., & Morefield, P. (2018). Estimates of present and future flood risk in the conterminous United States. *Environmental Research Letters*, 13(3). <https://doi.org/10.1088/1748-9326/aaac65>
- Wise, E. K. (2012). Hydroclimatology of the U.S. Intermountain West. *Progress in Physical Geography*, 36(4), 458–479. <https://doi.org/10.1177/0309133312446538>
- Wolman, M. G., & Costa, J. E. (1984). Envelope curves for extreme flood events: Discussion. *Journal of Hydraulic Engineering*, 110(1), 77–78. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1984\)110:1\(77\)](https://doi.org/10.1061/(ASCE)0733-9429(1984)110:1(77))
- Yochum, S. E. (2015). Colorado Front Range Flood of 2013: Peak flows and flood frequencies. Proceedings from the 3rd Joint Federal Interagency Conference on Sedimentation and Hydrologic Modeling, April 19–23, 2015, Reno, Nevada, USA.
- Yochum, S. E., & Moore, D. S. (2013). Colorado Front Range Flood of 2013: Peak flow estimates at selected mountain stream locations. U.S. Department of Agriculture, Natural Resources Conservation Service, Colorado State Office.
- Yochum, S. E., Sholtes, J. S., Scott, J. A., & Bledsoe, B. P. (2017). Stream power framework for predicting geomorphic change: The 2013 Colorado Front Range flood. *Geomorphology*, 292, 178–192. <https://doi.org/10.1016/j.geomorph.2017.03.004>
- Zogg, J., & Deitsch, K. (2013). *The flash flood potential index at WFO Des Moines*. NOAA National Weather Service: Iowa.