

Mapping Flash Flood Severity in the United States

MANABENDRA SAHARIA,^{a,b,c} PIERRE-EMMANUEL KIRSTETTER,^{a,c,d} HUMBERTO VERGARA,^{b,c,d}
JONATHAN J. GOURLEY,^d YANG HONG,^a AND MARINE GIROUD^e

^a School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, Oklahoma

^b Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma

^c Advanced Radar Research Center, University of Oklahoma, Norman, Oklahoma

^d NOAA/National Severe Storms Laboratory, Norman, Oklahoma

^e Department of Risk and Crises Management, École des Mines d'Alès, Alès, France

(Manuscript received 1 April 2016, in final form 3 November 2016)

ABSTRACT

Flash floods, a subset of floods, are a particularly damaging natural hazard worldwide because of their multidisciplinary nature, difficulty in forecasting, and fast onset that limits emergency responses. In this study, a new variable called “flashiness” is introduced as a measure of flood severity. This work utilizes a representative and long archive of flooding events spanning 78 years to map flash flood severity, as quantified by the flashiness variable. Flood severity is then modeled as a function of a large number of geomorphological and climatological variables, which is then used to extend and regionalize the flashiness variable from gauged basins to a high-resolution grid covering the conterminous United States. Six flash flood “hotspots” are identified and additional analysis is presented on the seasonality of flash flooding. The findings from this study are then compared to other related datasets in the United States, including National Weather Service storm reports and a historical flood fatalities database.

1. Introduction

Flash floods are swift flood responses to intense rainfall or release of water over a small area. Inundation over dry land occurs within minutes to a few hours of the rainfall event, potentially causing devastating impact on lives and infrastructure (Hong et al. 2012). For the water year 2014 (from 1 October 2013 to 30 September 2014) alone, direct flood damages totaled \$2.86 billion in the United States, according to the Flood Loss Report compiled by the National Weather Service (NWS; NWS 2014). Pielke and Downton (2000) found that flood damage costs for the United States have steadily increased throughout the twentieth century. Ashley and Ashley (2008) compiled a nationwide database of flood fatalities across the contiguous United States from 1959 to 2005 with detailed event and demographic information. They found that the majority of the fatalities were caused by flash floods. In 2014, 55 flood-related fatalities were recorded, where 29 were attributed to vehicle-related accidents and 39 were attributed to flash flood events (NWS 2014). An increasing

trend of heavy precipitation at both continental (Groisman et al. 2004) and global scale (Groisman et al. 2005), combined with rapid urbanization, is expected to increase the frequency and impact of flash floods.

But despite their disruptions to societies and economies, flash floods have not received the systematic and comprehensive study commensurate with their impacts. To understand how flood severity varies spatially, temporally, and geographically, we must first have a centralized database that collates quantitative information regarding floods. Such databases are not easily available, as the historical records are generally distributed across disparate sources and reports. Despite this limitation, flood characterization studies performed based on case studies and limited databases have provided important insights into spatial and temporal flood dynamics. Costa (1987a) investigated the hydraulic factors behind 12 of the largest floods of small basins ever measured by the U.S. Geological Survey (USGS) in the conterminous United States (CONUS) and related them to basin morphometry of the channels. He found that the peak discharges were controlled by an optimum combination of high rainfall rates and basin physiographic and morphological characteristics, such as elevation ratio,

Corresponding author e-mail: Pierre-Emmanuel Kirstetter, pierre.kirstetter@noaa.gov

drainage ratio, basin magnitude, basin slope, ruggedness number, etc. [Bhaskar et al. \(2000\)](#) devised a flash flood index in order to distinguish between floods and flash floods. The index was based on the characteristics describing the shape of the flood hydrograph, such as the rising curve gradient, the flood magnitude ratio, and the flash flood response time. [Merz and Blöschl \(2003\)](#) identified the causative mechanisms of floods using 11 518 maximum annual flood peaks in 490 Austrian catchments. They used a number of indicators such as timing of the floods, storm duration, rainfall depths, snowmelt, catchment state, runoff response dynamics, and spatial coherence to stratify flood peaks into flood process types. [Gaume et al. \(2009\)](#) reported the compilation of an inventory containing 550 previously undocumented flash flood events in seven hydrometeorological regimes in Europe. In one of the most comprehensive studies over Europe, [Marchi et al. \(2010\)](#) defined flash floods as those with an average time to peak of 6 h for catchments less than 100 km² in size. They identified causative processes behind 25 selected extreme flash floods in Europe and related them to climate and basin morphology. They characterized these events in terms of basin morphology, flood-generating rainfall, peak discharges, runoff coefficient, and response time to improve flash flood risk management. [Perucca and Angileri \(2011\)](#) approximated the basin behavior in response to rainstorms by evaluating the flash flood hazard of del Molle basin in Argentina. The study reported the probability of a serious flash flood hazard in the basin by analyzing different morphometric properties and suggested various mitigation measures.

Such studies have provided important insights into understanding floods in ungauged locations. [Castellarin \(2007\)](#) found that the reliability of probabilistic regional envelope curve (PREC) flood quantiles for ungauged sites is comparable with the reliability of regional estimates produced by the application of the index flood approach. [Gaume et al. \(2010\)](#) reduced the uncertainties in estimating regional flood quantiles by employing a Bayesian method on flash flood events occurring in ungauged catchments. [Ruiz-Villanueva et al. \(2013\)](#) characterized 41 flash flood events in small mountain basins of central Spain and analyzed their frequency, severity, seasonality, and synoptic meteorological causes, as well as the human impacts in terms of damage and fatalities. These methods can be used to reconstruct floods in ungauged basins.

Flash floods are generally related to localized, short, high-intensity rainfalls over basins with response times from minutes to a few hours. The nature of flash flood response is dictated by a complex interaction of various runoff-generating processes and conveyance of water out of the basin. An analysis of causative factors such as

geology, topography, pedology, and climatology is vital for revealing important aspects of the hydrological behavior. European studies have found that space and time scales of the most intense flash floods are smaller near the Mediterranean and decrease in intensity as one moves inland ([Gaume et al. 2009](#)). Furthermore, they find that the seasonality of the inland, continental flash floods tends to be in the warm season months, while those closer to the Mediterranean Sea typically occur in autumn months. Until recently, the lack of a comprehensive database that catalogs information related to flash flood timing, location, and severity such as the causative rainfall and basin geomorphology have hindered broad characterization studies. Some of the existing flooding databases include the International Flood Network (IFNET), which publishes a flood event database based on voluntary submission of events that caused 50 or more casualties between 2005 and 2007 (<http://www.internationalfloodnetwork.org/>). A global archive of large flood events (<http://www.dartmouth.edu/~floods/>) maintained by the Dartmouth Flood Observatory (DFO) curates one of the most comprehensive flood databases from a variety of sources such as remote sensing images and government reports. A digitized global flood inventory (1998–2008) with georeferenced events was also reported by [Adhikari et al. \(2010\)](#).

Building a comprehensive flash flood database is challenging because of the lack of availability of observations, and some studies have proposed alternative ways to describe floods. [Marchi et al. \(2009\)](#) indirectly reconstructed flood discharges in western Slovenia using postevent geomorphological surveys and interviews of eyewitnesses. [Wisser et al. \(2010\)](#) used the effects of human activities, irrigation, and reservoir operations in a flexible modeling framework to reconstruct twentieth-century global hydrography. [Smith et al. \(2014\)](#) employed the “structure from motion” technique to reconstruct flash flood magnitude using standard digital photographs and ground control points. [Casteller et al. \(2015\)](#) used dendrogeomorphic methods to reconstruct mean recurrence intervals of 21 flash flood events covering the period 1890–2009 along the Los Cipreses torrent (Neuquén, Argentina). To engage the public as citizen scientists and crowdsource data collection on natural hazards, the Meteorological Phenomena Identification Near the Ground (mPING) project released smartphone applications in which users can submit flooding reports based on different levels of severity ([Elmore et al. 2014](#)). The Severe Hazards Analysis and Verification Experiment (SHAVE), as described in [Gourley et al. \(2013\)](#), collected flooding and damage reports using public survey responses to a questionnaire with detailed impact classifications provided by [Calianno et al. \(2013\)](#).

Recently, [Gourley et al. \(2013\)](#) reported the building and online availability of the Unified Flash Flood Database that has been used in this study. The database sources flood events from USGS streamflow measurements, storm reports collected by the NWS, and public survey responses to a questionnaire developed for SHAVE. The data have been post processed extensively to harmonize data formats across the different sources. The long period of record and availability of flooding attributes such as flooding rise time, peak discharge, etc. makes it a suitable database for performing a spatially and temporally comprehensive flash flood characterization over the CONUS. Using a similar dataset, [Smith and Smith \(2015\)](#) identified the flashiest watersheds in the CONUS based on the frequency of discharge peaks exceeding $1 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$. Frequent flash flooding was observed in urban areas in the south-central United States (i.e., Tulsa, Oklahoma, and St. Louis, Missouri) up through the mid-Atlantic (i.e., Baltimore, Maryland) as well as the Pacific Northwest.

The purposes of this study are to 1) propose a new variable to describe the severity of flash flooding across the United States; 2) identify hotspots and evaluate their seasonal behavior; 3) extend the flashiness analysis beyond gauged basins to a continuous grid over the CONUS based on spatially distributed variables describing basin topography, hydroclimatology, underlying geology, and geomorphology; and 4) identify flash flood-prone areas that are not highlighted in the observation database. It is suggested that this analysis can be used for regional and community planning and mitigation purposes. The paper is organized as follows. [Section 2](#) describes the study region and datasets. [Section 3](#) proposes the flashiness variable, which is then evaluated spatially and seasonally in [section 4](#). The relationships of variables that describe flashiness are explored in [section 5](#). [Section 6](#) then regionalizes the flashiness variables to all grid points in the United States, permitting an analysis of the potential for severe flash flooding in ungauged basins. This section evaluates the flashiness maps through comparisons to related databases. The summary and conclusions are provided in [section 7](#).

2. Study area and data used

The Unified Flash Flood Database released by the Hydrometeorology and Remote Sensing (HyDROS) group at the University of Oklahoma is a curated database of flooding information from a variety of sources such as gauge measurements of streamflow by USGS, flash flooding reports in the NWS Storm Events Database, and public survey responses on flash flood impacts collected during the Severe Hazards Analysis and

Verification Experiment ([Ortega et al. 2009](#); [Gourley et al. 2010](#)). The high-resolution information provided by SHAVE, spatial coverage of NWS reports, and automated data collection mechanism of USGS streamflow records makes it one of the most representative flash flood databases in the United States. It is publicly available for free (<https://blog.nssl.noaa.gov/flash/database/>).

This study used automated streamflow measurements from the USGS. USGS collects instantaneous streamflow data at intervals ranging from 5 to 60 min for 10 106 gauges in the database. The NWS coordinates with local stakeholders and the USGS to define stages corresponding to action stage and minor, moderate, and major flooding for 3490 stream gauge locations. This subset of gauges from the USGS network has defined flooding thresholds, which is useful information for many applications including modeling. Action stage is defined as the stage at which NWS forecasters take “mitigation action for possible significant hydrologic activity,” and it often corresponds to bankfull conditions ([Hydrologic Services Program 2012](#), p. 2). In fact, 41% of USGS stations have identical action stage and bankfull stages, differing on average by 1.3%.

USGS also supplies regulation codes for these gauges, which is used to further screen out the gauges that have some amount of anthropogenic influence from regulation or diversion. After removing gauges with anthropogenic influences and no defined action stage, we are finally left with a data sample of 70 596 flooding events from 1649 stations. Flood events are defined when streamflow exceeds the defined action stage for that gauge. There must be a 24-h difference between when streamflow drops below action stage to the next rise for it to be counted as a separate event. The primary database comes with the following information for each gauge: the USGS gauge identifier (ID), latitude (decimal degrees), longitude (decimal degrees), start time (UTC) at which the flow first exceeded the action stage threshold, end time (UTC) when the flow dropped below the threshold, peak flow magnitude ($\text{m}^3 \text{ s}^{-1}$), peak time (UTC) at which peak flow occurred, and the difference between the time at which the discharge first exceeded action stage and reached its maximum value, defined as the flood rise time (h).

This event database was further enhanced with geomorphologic and climatological attributes derived for each basin in this study. A natural flood generally begins with snowmelt or intense rainfall. The characteristics of the underlying basin then dictate the speed at which water is conveyed through the basin and the magnitude of the maximum discharge. We concentrate on those floods with faster rise times and higher peak flows because of their devastating nature and lack of time to take mitigating actions. We introduce spatially

TABLE 1. Geomorphologic parameters included in this study.

Geomorphologic parameter	Details
Basin area	Total upstream area that contributes runoff
Shape factor	A dimensionless number that is given by drainage area divided by square of the main channel length ($K = \text{drainage area}/\text{channel length}^2$)
River length	Measured along a line centered from the basin outlet to the intersection of the extended main channel and the basin boundary
Relief ratio	Relief is the difference in elevation between the outlet and the highest point in the basin and relief ratio is relief divided by the basin length (a measure of the basinwide river slope); the higher the relief ratio, the higher the runoff and the shorter the flooding rise time
Slope index	Slope between two points along the main channel upstream from the mouth of the basin at distances equal to 10% and 85% of the total main channel length (Costa 1987a)
Slope to outlet	Local slope computed at a distance of 1 km over the basin outlet
Basin curve number	Soil Conservation Service Curve Number (SCS-CN) is an empirical parameter that characterizes the runoff properties for a particular soil and ground cover (U.S. Soil Conservation Service 1972)
K factor	Relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall
Rock depth	Depth to bedrock at the outlet
Soil texture (b parameter)	A proxy for soil texture, derived from the STATSGO database (Miller and White 1998)

distributed parameters to elucidate and quantify how the underlying, static basin characteristics influence flood response. Each of the attributes that are included in this study are provided in Table 1.

Several geomorphological attributes were extracted from the digital elevation model (DEM) data of the National Elevation Dataset (NED; <http://ned.usgs.gov/>) as potential explanatory variables of flash flood severity. Flow accumulation and flow direction information were extracted by delineating basins with USGS stations. The 30-m DEM was resampled to a 1-km grid using the National Hydrography Dataset (NHD; <http://nhd.usgs.gov/>) to ensure that DEM-based flow accumulation computations agree with the actual river network across the CONUS. The geomorphologic parameters were derived from the grid-based delineated catchments using custom libraries developed using MATLAB. Soil datasets from the State Soil Geographic (STATSGO) database (Miller and White 1998) were utilized to derive variables such as mean depth to bedrock and K factor (erodability). Land-cover and land-use data from the National Land Cover Dataset (Fry et al. 2011) were used to estimate the runoff curve number. Last, in addition to the geomorphological variables presented in Table 1, the hydroclimatic variables of mean annual precipitation and temperature were extracted from the 30-yr datasets (for period 1981–2010) prepared by the PRISM Climate Group of Oregon State University (<http://www.prism.oregonstate.edu/normals/>).

3. Flashiness

Flash floods are naturally occurring events that have been characterized by the NWS as “a rapid and extreme flow of high water into a normally dry area, or a rapid water

level rise in a stream or creek above a predetermined flood level, beginning within six hours of the causative event (e.g., intense rainfall, dam failure, ice jam)” (Hydrologic Services Program 2012, p. 3). Several approaches have been attempted to identify and characterize flash floods, such as geomorphology based (Costa 1987b; Gaume et al. 2009; Marchi et al. 2010), frequency based (Reed et al. 2007; Vogel et al. 2001), and flash flood guidance (Georgakakos 2006), among others. A new variable called “flashiness” is introduced in this paper as a measure of flood severity. It is defined as the difference between the peak discharge and action stage discharge normalized by the flooding rise time and basin area as given in Eq. (1) and visualized in Fig. 1. The flashiness metric gives the rate of rise of the hydrograph during flooding conditions and thus captures both the magnitude and timing aspects with higher values corresponding to more severe floods. Let ϕ be the flashiness, S be the number of gauging stations, and N_i be the number of events for a given gauge i , where $i = 1, \dots, S$. Thus, the flashiness for a given event j , where $j = 1, \dots, N_i$, at a given location i is

$$\phi_{ij} = \frac{Q_{ij}^{(p)} - Q_{ij}^{(a)}}{A_i T_{ij}}, \quad (1)$$

where $Q^{(p)}$ denotes the peak discharge, $Q^{(a)}$ is the action stage discharge, A is the basin area, and T is the flooding rise time. An empirical cumulative distribution function (ecdf) was then used to scale the values between 0 and 1 [Eq. (2)]. The standardized version of ϕ is

$$\tilde{\phi}_{ij} = \frac{1}{\sum_{i=1}^S N_i} \sum_{i=1}^S \sum_{j=1}^{N_i} I_{(\phi_{ij} \leq t)}, \quad (2)$$

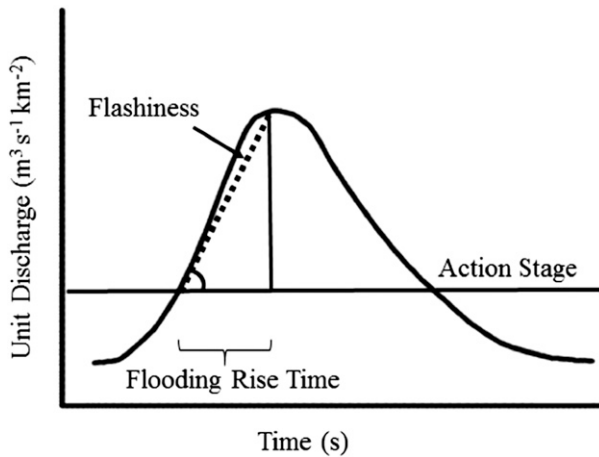


FIG. 1. Graphical representation of the definition of event-level flashiness.

where $I_{(E)}$ is the indicator function yielding 1 if the condition E is true and 0 otherwise. The flashiness information is available at two levels: event and basin. Event-level flashiness given in Eq. (2) is computed for all 70 596 flooding events. The characteristic scaled flashiness variable for a given basin i is summarized by the median value computed from all flooding events N_i observed at that station

$$[\tilde{\phi}_i: P(\tilde{\phi}_{ij} \leq \tilde{\phi}_i)] = \frac{1}{2}, \text{ for } j \in \{1, \dots, N_i\}, \quad (3)$$

where P represents the probability that the random variable ϕ_{ij} takes on a value less than or equal to ϕ_i .

Often, frequency-based approaches such as the discharge peak-over-threshold occurrences in Smith and Smith (2015) are used to quantify basin response as being flashy or not. The frequency of peaks-over-threshold highlights small basins, typically in urban areas, that are situated in hydroclimatic regimes with persistent and heavy annual rainfall such as the south-central United States, mid-Atlantic, and Pacific Northwest. However, this frequency-based definition of flashiness fails to identify those regions that may not flood on a frequent basis, but when they do flood, it can be catastrophic. Some dramatic examples include Big Thompson Canyon in July 1976 (which killed 145 people) and the Great Colorado Flood of September 2013 (Gochis et al. 2015), both of which occurred in locations situated along the Front Range of the Rocky Mountains, and the Arizona–Utah border canyon flash flood that killed up to 20 people in September 2015. It is postulated that the paucity of these flash floods are even more devastating than they would be if they were persistent because the occupants are less prepared and often unaware

of the danger. The flashiness variable used in this study differs from the frequency-based approaches in that it identifies those basins that have a high conditional probability of having a large-magnitude discharge in a short period of time. Flashiness is conditioned on the occurrence of heavy rainfall; thus, it represents the potential for a flashy response to input rainfall. The scaled flashiness variable for a given basin is the median value computed from all flooding events observed at that station.

Figure 2 shows the observed flashiness across the CONUS. At this point, the true spatial distribution of flashiness is limited by the density of the USGS stations with defined flooding thresholds. However, several regions emerge as being prone to flash flooding: 1) the West Coast, 2) Arizona, 3) the Front Range, 4) Flash Flood Alley, 5) the Missouri Valley, and 6) the Appalachians. High flashiness in the West Coast region is restricted to the coastal basins and the upslope region of the Sierra Nevada near Lake Tahoe. Arizona hosts a large number of flashy basins that range from the low deserts in southeastern Arizona all the way up to the Mogollon Rim and the higher-terrain plateau in the northern part of the state. Several flashy basins are apparent just to the east of the Rocky Mountains in the Front Range region. In Texas, several flashy basins are clustered around San Antonio, Austin, and Waco along the Balcones Escarpment in what is locally known as Flash Flood Alley (Flood Safety Education Project 2005). There is a secondary cluster closer to the Gulf Coast near Houston. Moving farther to the northeast, flashiness increases and appears to maximize in the center of Missouri. The Appalachian Mountains in the eastern United States exhibit high values of flashiness from Georgia all the way to Maine.

Flashiness, a continuous variable, may prove to be useful in the definition of a flash flood. To date, many definitions exist and often refer to the stream response to causative rainfall on the order of a few minutes to hours, typically less than six (U.S. Department of Commerce 2016). In the NWS, the time scale of 6 h is used to divide operational responsibility between local weather forecast offices that issue flash flood warnings and regional river forecast centers that issue river flood warnings. Definitions also refer to the basin catchment scale, which is linked to the basin’s response time. The European flash flood database described in Gaume et al. (2009) uses a catchment area threshold of 500 km². Marchi et al. (2010) used the European flash flood database to examine the characteristics of extreme events. They refer to a maximum basin scale associated with flash flooding of 1000 km². A limitation of a basin scale threshold to define flash flooding is that the effective basin area can be quite small for a localized convective storm near the basin outlet, which can produce a rapid response for a relatively large catchment.

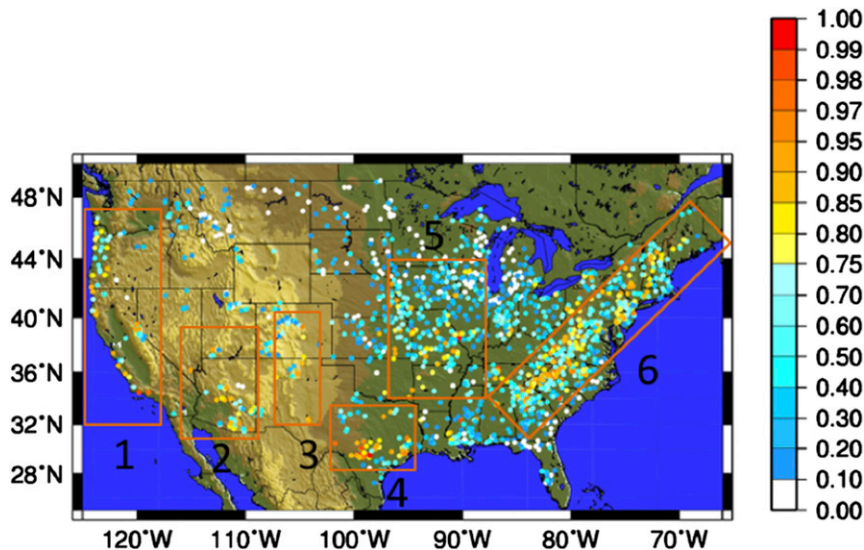


FIG. 2. Distribution of observed flashiness (0–1) over CONUS. The bounding boxes highlight known flash flood hotspots: 1) West Coast, 2) Arizona, 3) Front Range, 4) Flash Flood Alley, 5) Missouri Valley, and 6) the Appalachians.

In this study, we apply a subjective 75% quantile threshold on the flashiness variable (corresponding nonstandardized flashiness index is $0.0279 \text{ m}^{-3} \text{ s}^{-2}$) to separate basins that have “flashy” and “nonflashy” responses. This cutoff was later verified using actual storm data from the NWS. It is interesting to note that there are no flashy gauged basins (according to our definition) in the state of Florida or in a contiguous area stretching from the north central plains westward into the intermountain region of the Rockies. Some caution must be exercised at this point because flashiness can only be assessed in USGS-gauged basins that have flood stage definitions. The state of Wyoming, for example, only has four of these candidates.

4. Monthly variability of flash floods by region

A variety of meteorological processes, such as convective thunderstorms, tropical cyclones, and orographically enhanced precipitation in complex terrain, causes precipitation of varying characteristics and intensity at different times of the year to cause flash flooding (Saharia et al. 2016, manuscript submitted to *J. Hydrol.*). A better understanding of the monthly variation of flash flooding is necessary for assessing vulnerabilities and developing flood mitigation strategies. The locations of the regional hotspots identified using basin-level flashiness as shown in Fig. 2 can be attributed to specific conditions of topography and climate. Figure 3 shows the monthly frequency of flash floods using our 75th quantile of flashiness definition for each of the regions.

Flash floods on the West Coast (Fig. 3a) are clearly a cool season phenomenon that is related to the position of the jet stream, which directs extratropical cyclones with moisture from the Pacific Ocean into the mountains. Orographic enhancement by the topography increases precipitation amounts from the storms that can last several days. These flash flooding events begin to increase in November, reach their maximum frequency in December, and are essentially finished for the season by April. Very few flash floods occur in this region during the warm season months from May through August. Despite its geographical proximity to the West Coast, the frequency of flash floods in Arizona has a bimodal character (Fig. 3b). There is a reflection of the same cool season phenomenon experienced on the West Coast, but a smaller, secondary peak occurs from July through September. This region is impacted by the North American monsoon that transports moisture from the Gulf of California northward into the semiarid and hot deserts. The monsoon-forced flash floods do not occur as frequently as the larger-scale storms in the cool season, but they can be particularly catastrophic with intense, localized rainfall. As we move farther inland to the Front Range region, the frequency of flash floods increases at the beginning of the warm season and peaks during August (Fig. 3c). These storms are also related to the larger-scale circulation patterns with the North American monsoon. They differ from the Arizona storms in that their moisture fetch tends to be from the Gulf of Mexico up to the upslope region of the Front Range. The shift of the flood season from winter to

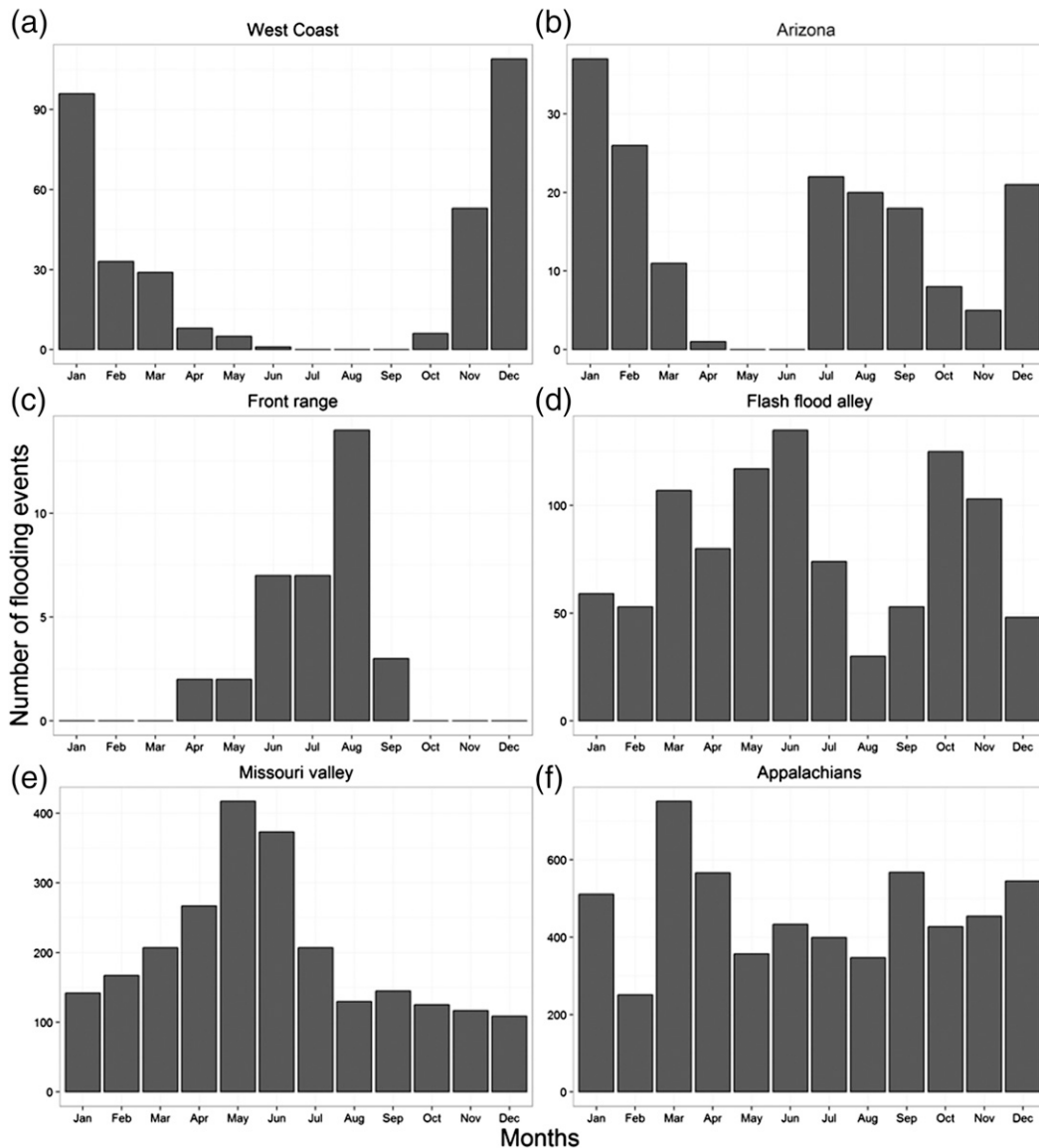


FIG. 3. Regionwise monthly frequency of events for floods exceeding 75th quantile flashiness, that is, flash floods, for known hotspots: (a) West Coast, (b) Arizona, (c) Front Range, (d) Flash Flood Alley, (e) Missouri Valley, and (f) the Appalachians.

summer as one moves inland was discussed by [Michaud et al. \(2001\)](#) and is further confirmed here.

The urban corridor spanning from Dallas to San Antonio in south-central Texas, also known as Flash Flood Alley, experiences some of the most dangerous floods in the country. This is caused by a combination of climatic and geomorphologic factors. Tropical air from the Gulf of Mexico, tropical cyclones, extratropical cyclones, and orographic uplift over the Balcones Escarpment result in very high precipitation efficiencies in this area, which produces flash floods ([Sharif et al. 2010](#)). This area shows a unique variation in flash flooding, with only a single month (August) where the frequencies are

very close to zero ([Fig. 3d](#)). The approximate bimodal distribution of flash flood peaks in spring and autumn can be attributed to the synoptic-scale patterns that drive the climatological rainfall peaks in these two seasons. Tropical cyclones also contribute in the latter, autumn peak.

Similar to Flash Flood Alley, the Missouri Valley region has a strong peak in flash flooding in the late spring months of May and June ([Fig. 3e](#)). This region is also known to have a secondary rainfall peak in the autumn months like Flash Flood Alley. However, this secondary rainfall maximum is not reflected in the monthly frequency of flash flooding. There are additional factors probably related to the characteristics of the rainfall (i.e., intensity) that apparently

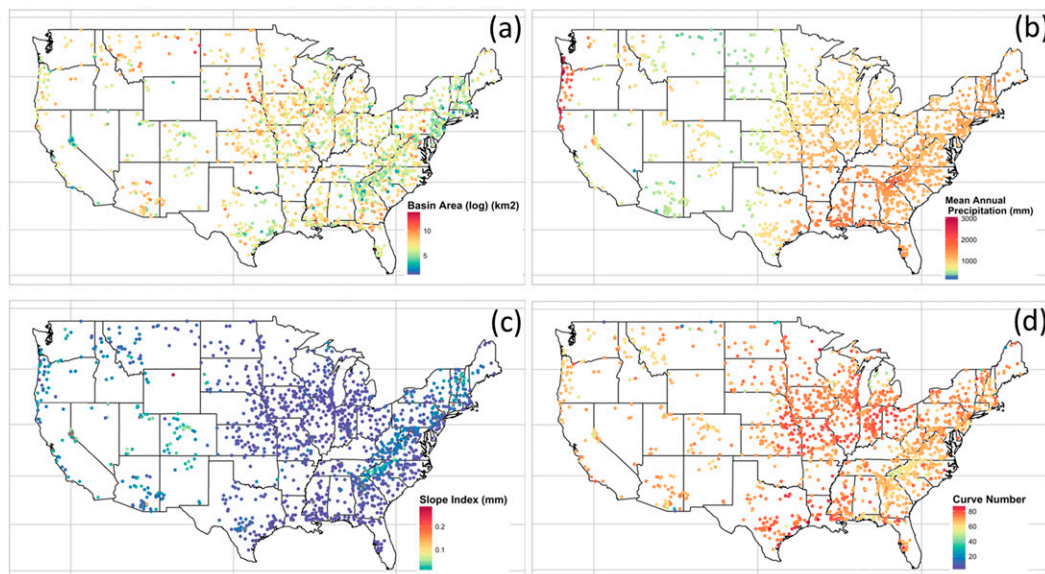


FIG. 4. Distribution of (a) basin area, (b) mean annual precipitation, (c) slope index, and (d) curve number over CONUS.

are not sufficient to cause many flash floods during the autumn months. The frequency of flash flooding in the Appalachians differs from the other regions, with a peak occurring in early spring (Fig. 3f). Moisture-laden air from both the Gulf of Mexico and the Atlantic Ocean is forced up the slopes of the Appalachian Mountains and causes rapid formation of runoff. Villarini and Smith (2010) explore the role of tropical cyclones in controlling the upper tail of flood distributions in eastern United States. Villarini et al. (2014) indicated that North Atlantic tropical cyclones are responsible for large-scale flooding over the eastern United States from Florida to Vermont and Maine, along with a secondary swath of enhanced flooding in the central United States. Tropical cyclones are more common in autumn, and we see a relative maximum in flash flooding in September. Sturdevant-Rees et al. (2001) also noted the large concentrations of unit peak discharges along the Atlantic seaboard and southeastern United States. The lack of flash floods in Florida and the coastal plains of the eastern seaboard states (see Fig. 2) further highlights the importance of the combination of moisture-rich air being forced upward by the terrain in causing flash floods. Konrad (2001) found that these comparatively flat and permeable areas do not produce large unit discharges, despite their proximity to moisture sources and being subject to frequent hurricanes.

5. Association of flashiness with basin geomorphology and climatology

Figures 4a–d show the spatial distributions of potentially important geomorphologic and climatological

characteristics that could influence a basin's response. In Fig. 4a, we see that there are concentrations of relatively small gauged catchments near Lake Tahoe on the California–Nevada border, in some Midwest cities such as St. Louis and Indianapolis, and along the Appalachians extending into the more populated regions of the Northeast. In comparing Fig. 4a to Fig. 2, we see that many of these basins are deemed as flashy, which is largely driven by the fact that they are small catchments. The climatological rainfall seems to influence flashiness in the Southeast near the Appalachians as well as in the Pacific Northwest, but there are some notable exceptions (Fig. 4b). Arizona, for example, hosts a number of flashy basins, but it is much more arid than other flashy regions. The slope index is the DEM-derived slope along the main channel length of a basin and is shown in Fig. 4c (Costa 1987a). Higher slope indices are associated with flashy responses in the Appalachians, the Sierra Nevada of California, and some basins in Arizona. But, again, there are numerous flashy basins that are relatively flat. Finally, the curve number is an empirical parameter that characterizes the runoff response to excess rainfall. It includes many factors such as hydrologic soil group and land cover in order to approximate infiltration, vegetative interception, and soil moisture retention processes on runoff generation. Figure 4d indicates higher runoff potential and thus some correlation with flashiness in Missouri and Flash Flood Alley. In reality, the behavior of a basin's response to rainfall is a result of a complex interaction between a large number of geomorphologic and climatological factors.

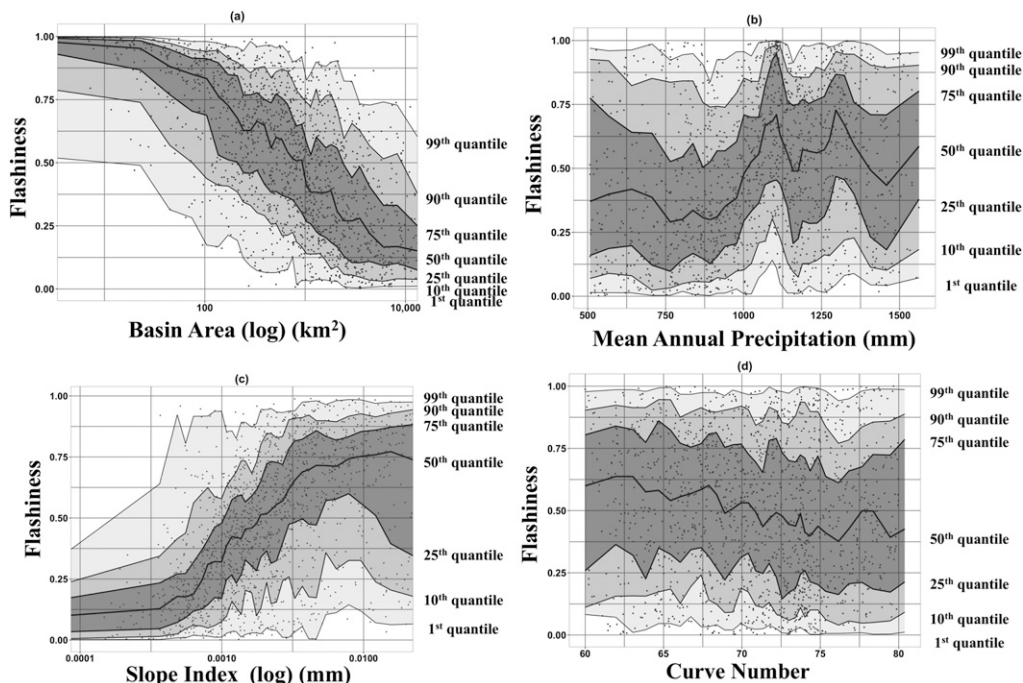


FIG. 5. First through 99th quantiles of flashiness vs (a) basin area, (b) mean annual precipitation, (c) slope index, and (d) curve number. Dots represent the actual data.

The influence of each of the factors described above (i.e., basin area, mean annual precipitation, slope index, and curve number) on flashiness is further analyzed using quantile plots in Figs. 5a–d. Information regarding the variability of the dependency can be extracted from the quantiles (1st, 10th, 25th, 50th, 75th, 90th, and 99th) of basin flashiness conditioned on the evaluated variables. The conditional median provides the first-order information of the dependency, while the interquartile area estimates the uncertainty in the relationship and the 10th and 90th quantiles describe the variation of extreme values of flashiness.

Figure 5a confirms the anticipated result that flashiness is more common in small catchments. The basin area associated with a median flashiness value of 0.75 (i.e., our subjective threshold for flash flooding) is 145 km². Mean annual precipitation is likely to be correlated to the frequency of flooding. It is noted that the flashiness variable describes the potential for fast and extreme runoff generation conditioned on heavy rainfall. In other words, it is not dependent on the frequency of flash flooding. Figure 5b shows the quantiles of flashiness with mean annual precipitation across the study region of the CONUS. The correlation of flashiness to mean annual precipitation is much weaker than that with basin area, but there is a slight

increase in flashiness with increasing climatological rainfall amounts.

Basins with steeper topography generally experience flashier floods with higher unit peak discharges and faster concentration times. The relationship between flashiness and slope index in Fig. 5c shows how flashiness of a basin increases as slope index increases, that is,

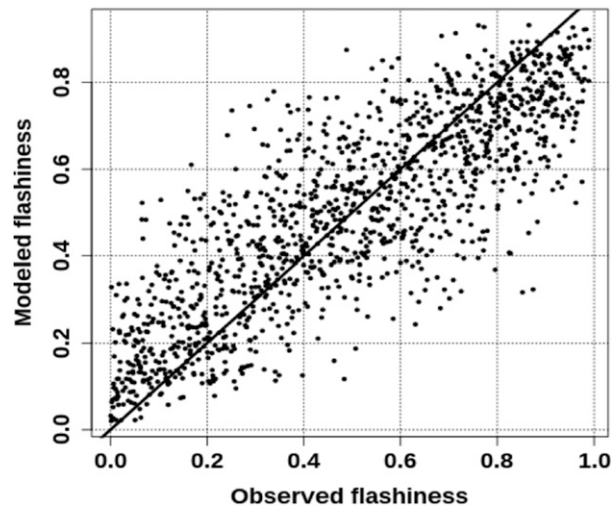


FIG. 6. Scatterplot of predicted vs observed flashiness. Bias is 0.6% and correlation $R = 0.83$.

basins become steeper. This variable influences flashiness approximately equal to the basin's catchment area. These two plots confirm that small catchments in steep terrain are generally expected to have a flashy response, according to our definition. Basin curve number is a widely used empirical parameter in hydrology that is based on soil and ground cover of an area and is used to approximate direct runoff from a rainfall event. It has a range of 30–100, with higher numbers indicating higher runoff potential. Figure 5d shows how flashiness of a basin depends on curve number. When grouping all the basins together in a single plot, there is no apparent trend in the median flashiness with increasing curve number.

The quantile plots provide valuable information about the variation of flashiness with individual geomorphologic and climatological variables. But in reality, flood processes are influenced by complex interactions between a large number of variables. Thus, this technique is extended into a multidimensional approach where the collective influence of a large number of explanatory variables on basin median flashiness can be understood. This can be used to not only predict flashiness in ungauged locations, but detect which explanatory variables have the greatest impact on floods in any particular location as well as ascribe a band of uncertainty to predicted flashiness.

6. Regionalization of flashiness

The dependence between the various explanatory variables and flashiness is examined using the generalized additive models for location, scale, and shape (GAMLSS; Rigby and Stasinopoulos 2005) technique. GAMLSS was proposed as an extension of the classical generalized additive models (Hastie and Tibshirani 1990), generalized linear models (McCullagh and Nelder 1989), and generalized additive mixed models (Fahrmeir and Lang 2001). The underlying assumption of all such models is that the variable we want to explain (flashiness, in this case) is a response variable whose distribution function varies according to the value assumed by the explanatory variables listed in Table 1. GAMLSS offers advantages over the previously mentioned approaches, such as 1) higher flexibility, as the response variable can follow a general distribution function and is not restricted to follow a distribution from the exponential family; and 2) it allows for modeling of not only the location parameter (related to the mean), but also scale and shape parameters (related to dispersion, skewness, and kurtosis). Because of its flexibility, GAMLSS has been used to model various hydrometeorological variables (e.g., Kirstetter et al. 2015).

Two main assumptions are made: 1) the response variable flashiness is a random variable following a known parametric distribution with density f conditional on the parameters μ (mu) and σ (sigma) and 2) the observed α (alpha) values are mutually independent given the parameter vectors μ and σ . Each distribution parameter is modeled as a function of the explanatory variables using monotonic (linear/nonlinear or smooth) link functions. More details are provided by Rigby and Stasinopoulos (2001, 2005), Akantziliotou et al. (2002), and Stasinopoulos and Rigby (2007), particularly on the model fitting and selection. It involves identifying a suitable distribution of flashiness, the explanatory variables, and the link functions. The estimation method is based on the maximum likelihood principle and the model selection is carried out by checking the significance of the fitting improvement in terms of information criteria such as the Akaike information criterion (AIC), the Schwarz Bayesian criterion (SBC), and the generalized AIC (Stasinopoulos and Rigby 2007). Forward, backward, and stepwise procedures were applied to select the meaningful explanatory variables, supervised by diagnostic plots to check the fitting performance, as discussed in Stasinopoulos and Rigby (2007). The GAMLSS modeling has been performed using the gamlss package developed for the R language.

A number of conditional two-parameter density functions (lognormal, normal, reverse Gumbel, logistic, gamma, etc.) were tested to fit the data, and the goodness of fit on the dataset was checked with the AIC for each of the semiparametric density fits as well as by checking the Gaussianity and independence of residuals. The beta distribution was found to be the most appropriate to examine the dependence of flashiness on various geomorphological variables. The original beta distribution with beta function B is given by

$$f(y | \alpha, \beta) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}, \quad (4)$$

for $y = (0, 1)$, $\alpha > 0$, and $\beta > 0$. In the GAMLSS implementation, $\alpha = \mu$ and $\beta > \sigma$. The function above was used to model the conditional flashiness distributions, where the location μ is linked to the expected flashiness value, and the scale σ is representative of prediction uncertainty. After selecting the distribution family, the structure of the model was refined through an iterative procedure by trying several combinations of explanatory variables. The trends for each parameter are fitted using penalized splines, which are more flexible than polynomials or fractional polynomials for modeling complex nonlinear relationships. The geophysical variables retained after analysis are presented in Table 2

TABLE 2. Statistical significance of explanatory variables in GAMLSS model. Not retained or not considered variables are marked with an em dash. Significance is expressed as a p value.

Variable	p value
Basin area (km ²)	$<2.2 \times 10^{-16}$
Shape factor	$<2.2 \times 10^{-16}$
River length (km)	—
Relief ratio	7.828×10^{-6}
Slope index	$<2.2 \times 10^{-16}$
Slope to outlet	7.473×10^{-11}
Annual precipitation (mm yr ⁻¹)	8.785×10^{-10}
Mean temperature (°C)	$<2.2 \times 10^{-16}$
Curve number	$<2.2 \times 10^{-16}$
K factor (erodability)	—
Depth to rock (cm)	8.796×10^{-8}
Soil texture (b parameter)	2.487×10^{-13}

along with their corresponding statistical significance values.

Consistency can be observed in the identification of the most important factors with what the spatial analysis suggested as discussed in section 5. Drainage area, slope index, the curve number, and the hydroclimatic variables of mean annual precipitation and temperature are

highlighted by their significance levels. This can be interpreted as a sign of robustness for the GAMLSS model. For the purpose of validation, the dataset is separated into two randomly selected samples, and the model is trained over a 75% random sample of the observations while 25% is used for validation. The expected values yielded by the GAMLSS model are compared to the observations and exhibit a correlation of 0.82 (67% of the variance of the data explained) and a negligible bias (0.4%). A similar result (correlation 0.83) is obtained with the validation dataset. The GAMLSS model is then recalibrated using geomorphological and climatological variables for the entire USGS observation dataset. Figure 6 shows a scatterplot of predicted versus observed flashiness. The two populations exhibit a correlation of 0.83 and bias as small as 0.6%. The model displays significant skill to predict the flashiness values; thus, we have confidence in the results as they are regionalized to ungauged basins.

The model is used to make predictions of flashiness at every grid point over the CONUS with a spatial resolution of 1 km. Figure 7a is the expected value of the predicted flashiness values between 0 and 1, and Fig. 7b shows the standard deviation of predicted flashiness.

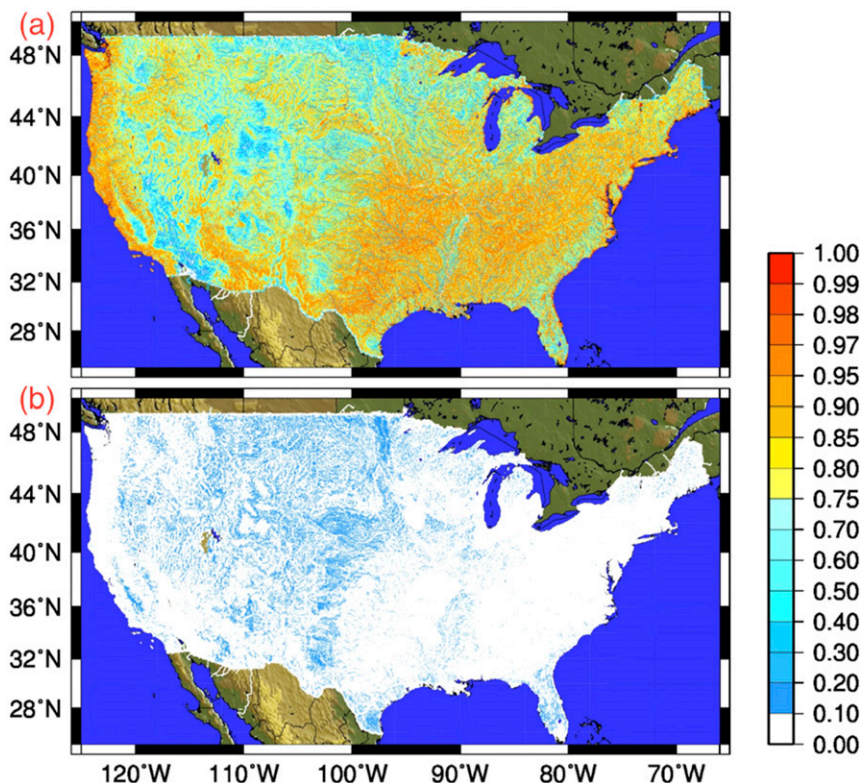


FIG. 7. Distribution of (a) expectation and (b) standard deviation of predicted flashiness values over CONUS.

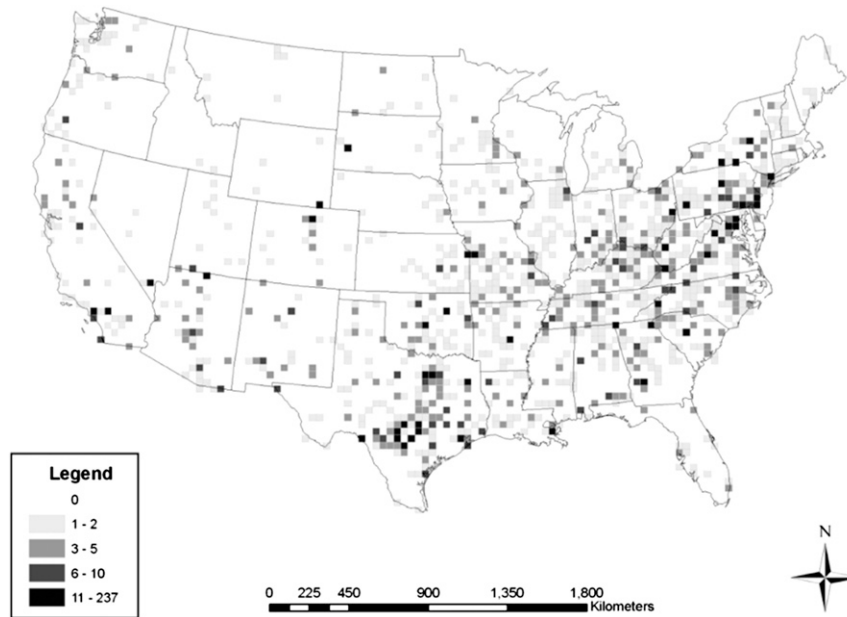


FIG. 8. Map of flash flood fatalities from 1959 to 2005 (source: Ashley and Ashley 2008).

The real value of this approach is the potential to identify flash flood hotspots in ungauged areas. In the predicted map, we see that the flashy basins on the West Coast are confined to the coastal areas and the upslope region of the inland Sierra Nevada extending northward. Arizona and the Front Range areas are also better highlighted in the predicted map. In Arizona, the flash flood-prone basins are located where there are steep slopes extending from southeastern Arizona and along the Mogollon Rim that separate the lower deserts from the higher plateau region in the northern part of the state. The predicted flashiness along the Front Range extends southward from north-central Colorado to southeastern parts of the state and then continues southward along the front range of the Rockies in New Mexico. In the observation database, Flash Flood Alley is largely concentrated in central Texas hill country and the Houston area. But the predicted map highlights a band of flood-prone areas all the way from southwestern Texas to Oklahoma, Arkansas, Kansas, and Missouri. Villarini et al. (2014) also highlighted the same area on spatial interpolation of maximum and 90th percentile of flood ratios associated with tropical cyclones. The predicted flashiness map highlights a number of regions that were not identified in the observed flashiness map in Fig. 2. Several localized hotspots are revealed, such as the western slopes of the Appalachians (Tennessee, Kentucky, and West Virginia) and a contiguous area in the western Dakotas, eastern Montana, and north-eastern Wyoming.

We introduce additional maps related to flash flooding in order to evaluate the predicted flashiness values in

ungauged regions. Ashley and Ashley (2008) compiled a national database of all flood fatalities in the CONUS between 1959 and 2005 along with their coordinates, which is shown in Fig. 8. This fatalities map shows good qualitative agreement with the predicted flashiness map of Fig. 7a.

The entire West Coast and Sierra Nevada mountain range has been highlighted in the predicted flashiness map, but it does not experience as many flood fatalities as the rest of the country. This may be due to better infrastructure, lower population densities in mountain communities, and better community resilience to disasters. The fatalities in the Arizona area are situated geographically similar to the predicted flashiness map. Fatalities in the Front Range tend to occur farther north rather than in southeastern Colorado and New Mexico, as highlighted in the flashiness map. Low population densities are likely the culprit for the mismatch. Flash Flood Alley and the populated Northeast are the most devastating regions in the country in terms of flood fatalities. Though our observation database in Fig. 2 identifies a localized region, the fatalities in Flash Flood Alley are spread over a wider area, as shown in Fig. 8. The extension of Flash Flood Alley by the predicted flashiness map up to the north and east correlates better with the spatial distribution of flood fatalities, attesting to the predictive power of the model in ungauged locations. The Missouri Valley with high flashiness values also experiences large numbers of flood fatalities. The predicted flashiness map also points to the flood-prone nature of the entire Appalachians, which is

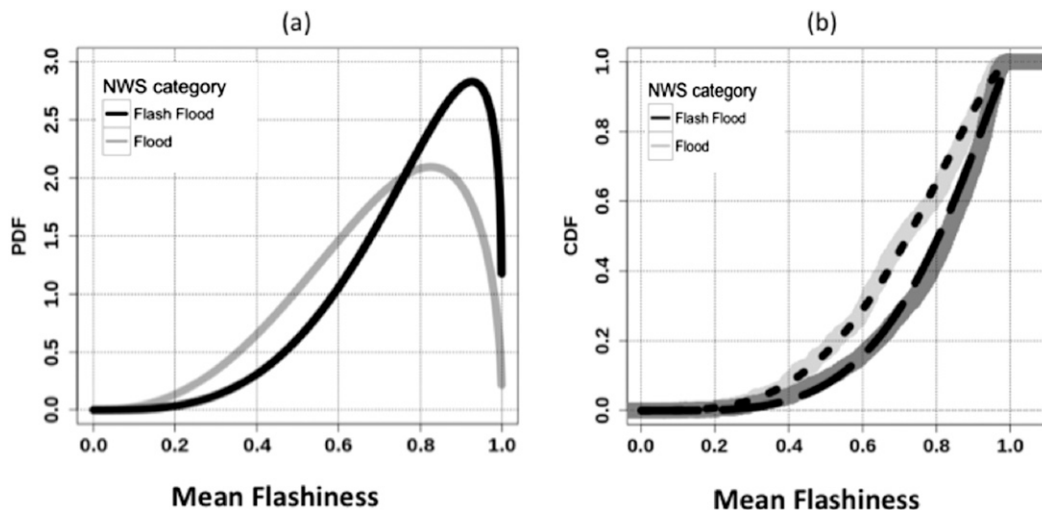


FIG. 9. (a) PDF of fitted beta models and (b) CDF of empirical mean flashiness by NWS categories of floods and flash floods. The fitted beta distribution models for floods (dotted line) and flash floods (dashed lines) are superimposed in (b).

observed in the number of casualties in the whole belt. It must be kept in mind that this is a very indirect way of validating our model and that flood fatalities are highly correlated with population density, infrastructure, and societal vulnerability.

The flashiness predictions are also compared to the storm database of flooding and flash flooding reports from 2007 to 2013. This dataset is included in the flash flood observation database described in Gourley et al. (2013). NWS forecasters report locations of flooding impacts using bounding polygons defined by as many as eight vertices. Currently, there are around 35 000 events in the database with an ID mentioning whether the event was a flood or a flash flood according the NWS definitions. Using GIS software, the mean flashiness in all such polygons was calculated and beta distributions were fitted to the populations of flashiness values in the flood and flash flood categories. Figure 9a gives the fitted probability distribution functions (PDFs) of mean flashiness for floods and flash floods according to actual NWS reports. The PDF shows a clear distinction between NWS-reported floods and flash floods between flashiness values of 0.75, which is the value we used initially to define flash flooding. The empirical and fitted cumulative distributions are shown in Fig. 9b. A Kolmogorov–Smirnov test performed on the two distributions yielded a p value of less than 2.2×10^{-16} and a D value of 0.22. Here, D is the maximum absolute difference between the two cumulative distribution functions (CDFs), which is maximum near the mean flashiness value of 0.75. The p value being small indicates that the distributions are significantly different,

accrediting the usefulness of flashiness to characterize the flash flood severity.

7. Conclusions

A long flood database spanning 78 years over the CONUS was used to explore the dependency of flood severity on geomorphological variables and climatology. A new variable called flashiness was introduced in this paper as a measure of flood severity. Flashiness is not dependent on the annual likelihood of flash flooding, but rather gives the potential of a basin to produce a rapid and significant response to heavy rainfall. Complex relationships between observed flashiness and a large number of geomorphologic and climatological variables were modeled using GAMLSS to predict flashiness at every location. The findings are summarized below:

- The spatial patterns of flood severity correlate well with regions that have been previously reported. Six flash flood hotspots were identified across the country: the West Coast, Arizona, the Front Range, Flash Flood Alley, the Missouri Valley, and the Appalachian Mountains.
- A monthly analysis of flash flooding in each of the hotspots revealed very different behavior in each region. The West Coast had the maximum frequency in flash flooding during the cool season while the interior regions were more commonly impacted during the warm season. Bimodal distributions in the monthly frequency of flash flooding were noted in both Arizona and in Flash Flood Alley in Texas.

- Several variables were used to model flashiness, and the most influential ones were the basin area and the basin's slope index. Small, steep basins had the flashiest responses.
- The predicted flashiness values were trained on observed values with a correlation of 0.82; the same correlation was met with stations that were independent from the training dataset.
- Additional spatial datasets related to flash flooding including fatalities and NWS reports showed good correspondence with the predicted flashiness map. The flashiness variable was shown to discriminate between NWS reports of flood and flash floods.
- Though the observation database showed Flash Flood Alley as being largely concentrated in central Texas hill country and the Houston area, the model-predicted flashiness extends this flash flood-prone area from southwestern Texas through the hill country and continuing northeastward into adjoining states.
- Localized hotspots were identified within the broad flash flood-prone areas as well as some of those outside of the originally defined regions, including the western slopes of the Appalachians in Tennessee, Kentucky, and West Virginia.

This study proposes an overview of how flood severity varies across the United States using a model that is able to highlight flash flood-prone areas in ungauged locations. As an extension of this study, we will combine the existing variables with event-level precipitation variability indices for improved modeling of flood severity. This analysis framework will serve as a baseline for evaluating distributed hydrologic model simulations such as the Flooded Locations and Simulated Hydrographs (FLASH) project (Gourley et al. 2016) under a variety of conditions. Since the geomorphological and climatological variables included in this paper can be computed globally and the data available over the United States is geographically diverse, this model can be extended to predict flashiness globally. The goal is to augment the existing database and to use sophisticated modeling techniques to improve predictions in gauged and ungauged locations.

Acknowledgments. This work was supported by the Disaster Relief Appropriations Act of 2013 (P.L. 113-2), which funded NOAA Research Grant NA14OAR4830100. The authors also thank the two anonymous reviewers, Francesco Serinaldi, and the editors for providing useful comments, which greatly improved the manuscript.

REFERENCES

- Adhikari, P., Y. Hong, K. R. Douglas, D. B. Kirschbaum, J. Gourley, R. Adler, and G. R. Brakenridge, 2010: A digitized global flood inventory (1998–2008): Compilation and preliminary results. *Nat. Hazards*, **55**, 405–422, doi:10.1007/s11069-010-9537-2.
- Akantziliotou, C., R. Rigby, and D. Stasinopoulos, 2002: The R implementation of generalized additive models for location, scale and shape. *Statistical Modelling in Society: Proceedings of the 17th International Workshop on Statistical Modelling*, Statistical Modelling Society, 75–83.
- Ashley, S. T., and W. S. Ashley, 2008: Flood fatalities in the United States. *J. Appl. Meteor. Climatol.*, **47**, 805–818, doi:10.1175/2007JAMC1611.1.
- Bhaskar, N., M. French, and G. Kyiamah, 2000: Characterization of flash floods in eastern Kentucky. *J. Hydrol. Eng.*, **5**, 327–331, doi:10.1061/(ASCE)1084-0699(2000)5:3(327).
- Calianno, M., I. Ruin, and J. J. Gourley, 2013: Supplementing flash flood reports with impact classifications. *J. Hydrol.*, **477**, 1–16, doi:10.1016/j.jhydrol.2012.09.036.
- Castellarin, A., 2007: Probabilistic envelope curves for design flood estimation at ungauged sites. *Water Resour. Res.*, **43**, W04406, doi:10.1029/2005WR004384.
- Casteller, A., M. Stoffel, S. Crespo, R. Villalba, C. Corona, and E. Bianchi, 2015: Dendrogeomorphic reconstruction of flash floods in the Patagonian Andes. *Geomorphology*, **228**, 116–123, doi:10.1016/j.geomorph.2014.08.022.
- Costa, J. E., 1987a: A comparison of the largest rainfall-runoff floods in the United States with those of the People's Republic of China and the world. *J. Hydrol.*, **96**, 101–115, doi:10.1016/0022-1694(87)90146-6.
- , 1987b: Hydraulics and basin morphometry of the largest flash floods in the conterminous United States. *J. Hydrol.*, **93**, 313–338, doi:10.1016/0022-1694(87)90102-8.
- Elmore, K. L., Z. L. Flamig, V. Lakshmanan, B. T. Kaney, V. Farmer, H. D. Reeves, and L. P. Rothfusz, 2014: MPING: Crowdsourcing weather reports for research. *Bull. Amer. Meteor. Soc.*, **95**, 1335–1342, doi:10.1175/BAMS-D-13-00014.1.
- Fahrmeir, L., and S. Lang, 2001: Bayesian semiparametric regression analysis of multicategorical time-space data. *Ann. Inst. Stat. Math.*, **53**, 11–30, doi:10.1023/A:1017904118167.
- Flood Safety Education Project, 2005: *Flash Flood Alley*. Flood Safety Education Project, Boulder, CO, DVD.
- Fry, J. A., and Coauthors, 2011: Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Remote Sensing*, **77**, 858–864.
- Gaume, E., and Coauthors, 2009: A compilation of data on European flash floods. *J. Hydrol.*, **367**, 70–78, doi:10.1016/j.jhydrol.2008.12.028.
- , L. Gaál, A. Viglione, J. Szolgay, S. Kohnová, and G. Blöschl, 2010: Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites. *J. Hydrol.*, **394**, 101–117, doi:10.1016/j.jhydrol.2010.01.008.
- Georgakakos, K. P., 2006: Analytical results for operational flash flood guidance. *J. Hydrol.*, **317**, 81–103, doi:10.1016/j.jhydrol.2005.05.009.
- Gochis, D., and Coauthors, 2015: The Great Colorado Flood of September 2013. *Bull. Amer. Meteor. Soc.*, **96**, 1461–1487, doi:10.1175/BAMS-D-13-00241.1.
- Gourley, J. J., J. M. Erlingis, T. M. Smith, K. L. Ortega, and Y. Hong, 2010: Remote collection and analysis of witness reports on flash floods. *J. Hydrol.*, **394**, 53–62, doi:10.1016/j.jhydrol.2010.05.042.
- , and Coauthors, 2013: A Unified Flash Flood Database across the United States. *Bull. Amer. Meteor. Soc.*, **94**, 799–805, doi:10.1175/BAMS-D-12-00198.1.

Adhikari, P., Y. Hong, K. R. Douglas, D. B. Kirschbaum, J. Gourley, R. Adler, and G. R. Brakenridge, 2010: A digitized

- , and Coauthors, 2016: The Flooded Locations And Simulated Hydrographs (FLASH) project: Improving the tools for flash flood monitoring and prediction across the United States. *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS-D-15-00247.1, in press.
- Groisman, P. Ya., R. W. Knight, T. R. Karl, D. R. Easterling, B. Sun, and J. H. Lawrimore, 2004: Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. *J. Hydrometeorol.*, **5**, 64–85, doi:10.1175/1525-7541(2004)005<0064:CCOTHG>2.0.CO;2.
- , —, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. N. Razuvaev, 2005: Trends in intense precipitation in the climate record. *J. Climate*, **18**, 1326–1350, doi:10.1175/JCLI3339.1.
- Hastie, T. J., and R. J. Tibshirani, 1990: *Generalized Additive Models*. CRC Press, 352 pp.
- Hong, Y., P. Adhikari, and J. J. Gourley, 2012: Flash flood. *Encyclopedia of Natural Hazards*, P. Bobrowsky, Ed., Encyclopedia of Earth Science Series, Springer, 324–325, doi:10.1007/978-1-4020-4399-4_136.
- Hydrologic Services Program, 2012: National Weather Service manual 10-950. National Weather Service, 5 pp. [Available online at <http://www.nws.noaa.gov/directives/sym/pd01009050curr.pdf>.]
- Kirstetter, P.-E., J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur, 2015: Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resour. Res.*, **51**, 1422–1442, doi:10.1002/2014WR015672.
- Konrad, C. E., 2001: The most extreme precipitation events over the eastern United States from 1950 to 1996: Considerations of scale. *J. Hydrometeorol.*, **2**, 309–325, doi:10.1175/1525-7541(2001)002<0309:TMEPEO>2.0.CO;2.
- Marchi, L., and Coauthors, 2009: Comprehensive post-event survey of a flash flood in western Slovenia: Observation strategy and lessons learned. *Hydrol. Processes*, **23**, 3761–3770, doi:10.1002/hyp.7542.
- , M. Borga, E. Preciso, and E. Gaume, 2010: Characterisation of selected extreme flash floods in Europe and implications for flood risk management. *J. Hydrol.*, **394**, 118–133, doi:10.1016/j.jhydrol.2010.07.017.
- McCullagh, P., and J. A. Nelder, 1989: *Generalized Linear Models*. Chapman and Hall, 511 pp.
- Merz, R., and G. Blöschl, 2003: A process typology of regional floods. *Water Resour. Res.*, **39**, 1340, doi:10.1029/2002WR001952.
- Michaud, J. D., K. K. Hirschboeck, and M. Winchell, 2001: Regional variations in small-basin floods in the United States. *Water Resour. Res.*, **37**, 1405–1416, doi:10.1029/2000WR900283.
- Miller, D. A., and R. A. White, 1998: A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interact.*, **2**, doi:10.1175/1087-3562(1998)002<0001:ACUSMS>2.3.CO;2.
- NWS, 2014: Annual flood loss summary reports. Accessed 1 April 2016. [Available online at <http://www.nws.noaa.gov/hic/summaries/>.]
- Ortega, K. L., T. M. Smith, K. L. Manross, A. G. Kolodziej, K. A. Scharfenberg, A. Witt, and J. J. Gourley, 2009: The Severe Hazards Analysis and Verification Experiment. *Bull. Amer. Meteor. Soc.*, **90**, 1519–1530, doi:10.1175/2009BAMS2815.1.
- Perucca, L. P., and Y. E. Angilieri, 2011: Morphometric characterization of del Molle basin applied to the evaluation of flash floods hazard, Iglesia Department, San Juan, Argentina. *Quat. Int.*, **233**, 81–86, doi:10.1016/j.quaint.2010.08.007.
- Pielke, R. A., and M. W. Downton, 2000: Precipitation and damaging floods: Trends in the United States, 1932–97. *J. Climate*, **13**, 3625–3637, doi:10.1175/1520-0442(2000)013<3625:PADFTI>2.0.CO;2.
- Reed, S., J. Schaake, and Z. Zhang, 2007: A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *J. Hydrol.*, **337**, 402–420, doi:10.1016/j.jhydrol.2007.02.015.
- Rigby, R. A., and D. M. Stasinopoulos, 2001: The GAMLSS project: A flexible approach to statistical modelling. *New Trends in Statistical Modelling: Proceedings of the 16th International Workshop on Statistical Modelling*, B. Klein and L. Korsholm, Eds., University of Southern Denmark, 337–345.
- , and —, 2005: Generalized additive models for location, scale and shape. *J. Roy. Stat. Soc. Ser. C Appl. Stat.*, **54**, 507–554, doi:10.1111/j.1467-9876.2005.00510.x.
- Ruiz-Villanueva, V., A. Díez-Herrero, J. M. Bodoque, J. A. Ballesteros Cánovas, and M. Stoffel, 2013: Characterisation of flash floods in small ungauged mountain basins of central Spain using an integrated approach. *Catena*, **110**, 32–43, doi:10.1016/j.catena.2013.06.015.
- Sharif, H. O., A. A. Hassan, S. Bin-Shafique, H. Xie, and J. Zeitler, 2010: Hydrologic modeling of an extreme flood in the Guadalupe River in Texas. *J. Amer. Water Resour. Assoc.*, **46**, 881–891, doi:10.1111/j.1752-1688.2010.00459.x.
- Smith, B. K., and J. A. Smith, 2015: The flashiest watersheds in the contiguous United States. *J. Hydrometeorol.*, **16**, 2365–2381, doi:10.1175/JHM-D-14-0217.1.
- Smith, M. W., J. L. Carrivick, J. Hooke, and M. J. Kirkby, 2014: Reconstructing flash flood magnitudes using ‘Structure-from-Motion’: A rapid assessment tool. *J. Hydrol.*, **519B**, 1914–1927, doi:10.1016/j.jhydrol.2014.09.078.
- Stasinopoulos, D. M., and R. A. Rigby, 2007: Generalized additive models for location scale and shape (GAMLSS) in R. *J. Stat. Software*, **23**, 1–46, doi:10.18637/jss.v023.i07.
- Sturdevant-Rees, P., J. A. Smith, J. Morrison, and M. L. Baeck, 2001: Tropical storms and the flood hydrology of the central Appalachians. *Water Resour. Res.*, **37**, 2143–2168, doi:10.1029/2000WR900310.
- U.S. Department of Commerce, 2016: Flash flooding definition. Accessed 20 December 2016. [Available online at <http://www.weather.gov/phi/FlashFloodingDefinition>.]
- U.S. Soil Conservation Service, 1972: Hydrology. *SCS National Engineering Handbook*, Supplement A, section 4, chapter 10, USDA.
- Villarini, G., and J. A. Smith, 2010: Flood peak distributions for the eastern United States. *Water Resour. Res.*, **46**, W06504, doi:10.1029/2009WR008395.
- , R. Goska, J. A. Smith, and G. A. Vecchi, 2014: North Atlantic tropical cyclones and U.S. flooding. *Bull. Amer. Meteor. Soc.*, **95**, 1381–1388, doi:10.1175/BAMS-D-13-00060.1.
- Vogel, R. M., A. Zafirakou-Koulouris, and N. C. Matalas, 2001: Frequency of record-breaking floods in the United States. *Water Resour. Res.*, **37**, 1723–1731, doi:10.1029/2001WR900019.
- Wisser, D., B. M. Fekete, C. J. Vörösmarty, and A. H. Schumann, 2010: Reconstructing 20th century global hydrography: A contribution to the Global Terrestrial Network–Hydrology (GTN-H). *Hydrol. Earth Syst. Sci.*, **14**, 1–24, doi:10.5194/hess-14-1-2010.