Received: 25 March 2018 | Revised: 14 August 2018 | Accepted: 26 October 2018

DOI: 10.1111/jfr3.12512

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management | WILEY

ORIGINAL ARTICLE

The weight of the flood-of-record in flood frequency analysis

Scott St. George[1,2] | Manfred Mudelsee[3,4]

[1]Department of Geography, Environment and Society, University of Minnesota, Minneapolis, Minnesota

[2]Department of Geography, Johannes Gutenberg University, Mainz, Germany

[3]Climate Risk Analysis, Heckenbeck, Bad Gandersheim, Germany

[4]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

Correspondence
Scott St. George, Department of Geography, Environment and Society, University of Minnesota, MN.
Email: stgeorge@umn.edu

Funding information
Alexander von Humboldt Foundation

Abstract

The standard approach to flood frequency analysis (FFA) fits mathematical functions to sequences of historic flood data and extrapolates the tails of the distribution to estimate the magnitude and likelihood of extreme floods. Here, we identify the most exceptional floods in the United States as compared against other major floods at the same location, and evaluate how the flood-of-record (Q_{max}) influences FFA estimates. On average, floods-of-record are 20% larger by discharge than their second-place counterparts (Q_2), and 212 gages (7.3%) have Q_{max}:Q_2 ratios greater than two. There is no clear correspondence between the Q_{max}:Q_2 ratio and median instantaneous discharge, and exceptional floods do not become less likely with time. Excluding Q_{max} from the FFA causes the median 100-year flood to decline by −10.5%, the 200-year flood by −11.8%, and the 500-year flood by −13.4%. Even when floods are modelled using a heavy tail distribution, the removal of Q_{max} yields significantly "lighter" tails and underestimates the risk of large floods. Despite the temporal extension of systematic hydrological observations in the United States, FFA is still sensitive to the presence of extreme events within the sample used to calculate the frequency curve.

KEYWORDS

flood frequency analysis, floods, heavy tail analysis, record floods, United States

1 | INTRODUCTION

Worldwide, flooding is the leading natural hazard that affects humanity (Kellens, Terpstra, & De Maeyer, 2012). Floods have been the costliest peril for the past 4 years running (2013–2016), and in 2016 economic losses due to flooding were $62 billion (Aon Benfield, 2016). During the past decade, on average floods have affected 87 million people and caused nearly 6,000 deaths every year (Guha-Sapir, Hoyois, & Below, 2016). Nearly all potential responses to floods—including the construction of dams, levees, or diversions, the implementation of insurance systems to compensate victims, and land-use management—depend at least partly on our judgement of the risk posed by future floods. The most common tool used to evaluate those risks is flood frequency analysis (FFA), which attempts to answer questions related to flood problems through the application of probability principles. The standard approach to FFA fits mathematical functions to sequences of historic flood data and extrapolates the tails of the distribution to estimate the magnitude and likelihood of extreme floods (England Jr. et al., 2018; Klemeš, 1989; Mertz & Blöschl, 2008). Estimates obtained from FFA are used to support decisions regarding the design of individual flood mitigation projects (Brooks & St. George, 2015; Clark, 1996) and are central to most national and international schemes for hazard risk assessments (Michel-Kerjan & Kunreuther, 2011; Porter & Demeritt, 2012) and land-use planning (Dewan, Islam, Kumamoto, & Nishigaki, 2007; Ganoulis, 2003).

The efficacy of the standard approach is usually tested on the ability of its statistical model to fit the distribution of observations (Kochanek et al., 2014; Rahman, Zaman, Haddad, El Adlouni, & Zhang, 2015) or to reproduce the probability of synthetic flood data generated by a known distribution (Rahman et al., 2015). But since their inception, FFA methods have been known to struggle when used to assess the risks of high-magnitude, low-frequency floods. In one of the earliest attempts to apply objective estimation methods to flood hydrology, Gumbel (1941) pointed out that "[f]or the two or three extreme floods, the return periods are based on a few observations, and consequently, the agreement [between statistical theory and hydrological observation] is not very good." The regional frequency approach (Dalrymple, 1960; Hosking & Wallis, 1997) was developed to circumvent the problem of estimating rare events from short observational records, effectively trading space for time by pooling streamflow data at the target site with data from hydrologically similar gages elsewhere. But regional flood frequency analysis (RFFA) can still perform poorly when used to predict large, infrequent floods; across the globe, discharge estimates of the 100-year flood derived from RFFA have errors greater than 50% (A. Smith, Sampson, & Bates, 2015). And since such events are rare in practice, the uncertainty in estimated recurrence intervals still increases substantially towards the upper end of the flood frequency curve (Eychaner, 2015; Parkes & Demeritt, 2016). These results imply that, despite the progressive spatial expansion and temporal extension of systematic hydrological observations globally, the probability estimates generated by FFA are still sensitive to the presence of specific extreme events within the sample used to calculate the frequency curve (Parkes & Demeritt, 2016). Furthermore, because FFA fits a parametric probability distribution to the log-transformed observations (United States Geological Survey, 1982), the validity of that approach depends upon the assumed distribution being a suitable match for the underlying hydrological processes. A heavy-tail distribution—where the probability of observing an extreme value equal to or greater than a certain value, x, is proportional to $x - \alpha$ (Resnick, 2007)—may therefore be a useful alternative to standard FFA methods because its form is parametrically less restricted and offers more distributional robustness. Instead of assuming that floods follow a specific distribution, this alternative approach presupposes only that the tail probability behaves as a power law, which can be described by the heavy-tail index (α). Because heavy-tail distributions also encompass other shapes applied commonly to extremes, including the Generalised Extreme Value and Generalised Pareto functions (Resnick, 2007), this test allows us to examine whether our results are sensitive to the particular choice of distribution. The α value can be applied to derive return periods and other risk measures (Anderson & Meerschaert, 1998; El Adlouni, Bobée, & Ouarda, 2008), but its sensitivity to the presence of specific extreme events has, however, not been analysed previously.

Here, we evaluate how a singular flood—the flood-of-record (or maximum flood; Crippen & Bue, 1977; Vogel, -Zafirakou-Koulouris, & Matalas, 2001)—affects the estimates produced by the FFA approach recommended by Bulletin 17B (United States Geological Survey, 1982). In part, we focus on the flood-of-record because affected residents often identify the largest event they have experienced as their most pressing concern, even if that flood occurred decades ago (Lave & Lave, 1991), and as a result, the largest known flood can have an outsized influence on flood mitigation decisions (St. George & Rannie, 2003). Drawing upon a large set of long-term hydrological records from the United States, we compare the flood-of-record at each gage against other floods at the same location and identify those events are most exceptional in comparison to high flows that came either before or after. Next, we illustrate how these singular events affect quantitative estimates of flood risk across the country by conducting paired FFA and heavy-tail analyses that either include or exclude the flood-of-record. Finally, we show that, even though hydrological records from the United States now span several decades or more, the issue raised by Gumbel in the 1940s remains a challenge to flood-risk assessment today.

2 | DATA AND METHODS

We obtained annual peak streamflow data from the U.S. Geological Survey's National Water Information System (U.S. Geological Survey, 2016) from all streamgages in the continental United States with more than 50 years of observations, as well as the associated metadata for each gage. For each streamflow record, any observations made after the river was affected by regulation or diversion were omitted from the analysis; this screening step also eliminated all floods caused by upstream dam failure (including, e.g., the South Fork Dam failure in 1889, and the Teton Dam failure in 1976; Seed & Duncan, 1987; Katkins, Davis Todd, Wojno, & Coleman, 2013). The length criterion was applied a second time to ensure that each peak flow record still retained at least 50 observations after the screening. Overall, the final data set had peak flow records from 2,790 gages, had a median record length of 70 years, and included flood observations made between CE 1773 and 2016.

For each gage, we identified the flood-of-record (the largest flood by discharge; Q_{max}) and then computed the ratio between the discharge of Q_{max} and that of the second largest observed flood (Q_2), which represents the "degree of exceptionalness" exhibited by the largest known flood. In order to determine the influence of the flood-of-record on quantitative estimates of flood probability, for all records, we conducted FFA twice, first using all flow data from that gage and then repeating it after Q_{max} was excluded. In both cases, the magnitude for three recurrence intervals (the 100-, 200-, and 500-year floods) were estimated after fitting the flood observations with a Log Pearson III distribution (United States Geological Survey, 1982). Flood frequencies were calculated in the MATLAB® computing environment using code written by Jeff Burkey of King County's Department of Natural Resources and Parks (Burkey, 2009). We estimated the heavy-tail index for each gage record in the same manner, first using all annual maxima and then repeating it after Q_{max} was excluded. Our heavy-tail index estimator (Mudelsee & Bermejo, 2017) has been proven to be accurate due to its usage of an optimal selector of the order (i.e., the fraction of largest values to utilise for the estimation). But because achieving a satisfactory level of accuracy requires a large number of observations (100 or more years of data), the heavy-tail index was estimated only on those 146 gages that satisfied the stricter length criterion.

3 | RESULTS AND DISCUSSION

Because the National Water Information System includes flood information from historical evidence, there are 44 record floods that pre-date the installation of the nation's first stream gage in 1889 (Frazer & Heckler, 1989). The earliest flood-of-record occurred in 1791, when the Swannanoa River in western North Carolina rose to a level that has still not been equalled more than two centuries later (Tennessee Valley Authority, 1963). Most record floods have been observed during the last 50–60 years (Figure 1a), but because this increase has followed the progressive growth of the national hydrological monitoring network (Supporting Information Figure S1), it should be regarded as the by-product of expanded monitoring rather than reflecting a trend towards bigger floods.

Record floods naturally cluster together geographically (Figure 1b). During the top flood year of 1964, rivers throughout the Central Pacific Coast set records, including streams in Oregon, Idaho, northern California, southern Washington, and small parts of western and northern Nevada. Most of these floods were triggered by an atmospheric river (also known as a "Pineapple Express") that made landfall on December 21 and 22, 1964 (Dettinger, Ralph, Das, Neiman, & Cayan, 2011; Waananen, Harris, & Williams, 1971), and the synchrony and magnitude of flooding across the region caused this event to become known locally as the "Thousand Year Flood" (Lucia, 1965). Earlier that year, records were also set at gages in the southeastern United States, including several exceptional floods in northern Florida caused by Hurricane Dora (Frank, 1965). In 1972, nearly 100 gages measured record floods, most located in Pennsylvania, New Jersey, and New York and

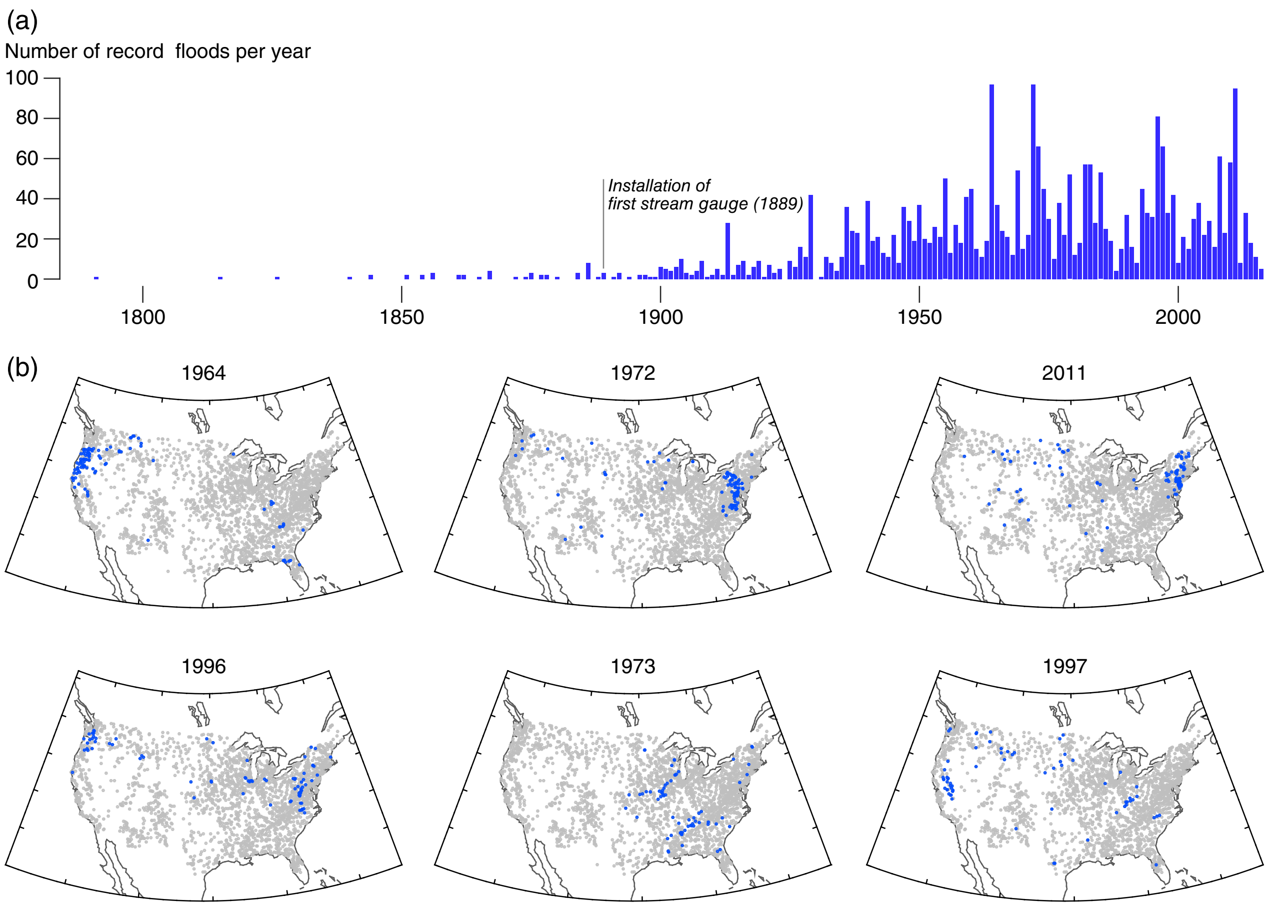

FIGURE 1 Record floods in the continental United States. (a) Total number of record floods, by year, in the screened National Water Information System database. (b) Maps showing the location of record floods for the six most exceptional years (highest number of record floods). Blue symbols represent gages where record floods occurred in the specified year, while the grey symbols denote gages that were active but did not observe their flood-of-record. The six annual maps are arranged from the greatest (upper left) to least number of record floods (lower right)

caused by heavy rains associated with Hurricane Agnes (Namais, 1972). The unprecedented 2011 floods were split between the northeastern United States and the central Plains. The former was a product of northward track of Hurricane Irene along the Atlantic Seaboard (Avilia & Cangialosi, 2011), while the latter were part of extensive flooding in the Mississippi basin during April and May (Goodwell et al., 2014), and included a record crest at Vicksburg, Mississippi (where the gage record extends back to 1897). Record-setting floods in 1996 were part of widespread high water in Pennsylvania, Virginia, and West Virginia due to intense rain and snowmelt runoff in January (Leathers, Kluck, & Kroczynski, 1998), while at the other end of the country, the Pacific Northwest saw flooding in many of the region's major rivers (Colle & Mass, 2000), most notably the Willamette River in Portland. Although the Mississippi Flood of 1973 did reach a then-record stage at St. Louis (Belt Jr., 1975) that was later eclipsed in 1993, more than a dozen tributaries in the Illinois and Ohio river basins had peak flows that still remain unequalled today. Finally, in 1997, records were set in northern California during the so-called New Year's Flooding (Galewsky & Sobel, 2005), in the northern Central Plains (including tributaries of the Red River of the North; Todhunter, 2001), and in the upper Ohio Basin (mainly in Kentucky; Hughey & Tobin, 2006). On the opposite end of the spectrum, since 1900, there have only been 2 years without any record floods—1924 and 1930—and both coincided with widespread drought conditions (respectively, affecting the western Pacific Coast and the southeastern United States; Shelton, 1977; Cook & Krusic, 2004).

Overall, floods-of-record are 20% larger by discharge than their second-place counterparts (the median Q_{max}:Q_2 ratio is 1.2; Figure S2). More than one fifth of all gages (21.8%) have maximum floods that are more than one and a half times the magnitude of Q_2, and 212 gages (7.3%) have Q_{max}:Q_2 ratios greater than two. Although we might expect "surprise" floods (i.e., floods that are much larger than all others ever observed at the same gauge) to be more common on small rivers, there is no clear correspondence between the Q_{max}:Q_2 ratio and median instantaneous discharge (Figure S3a), although there are few instances of truly exceptional floods (Q_{max}:Q_2 ratios above 1.5) in larger rivers (1,000 m^3/s and up). Furthermore, having a long observational record does not mean that surprise floods are not possible. Prior analyses of individual gages in the United States had shown that, as hydrological records become longer, the uncertainty in the magnitude of the 50- or 100-year flood decreases (Benson & Carter, 1973; Feaster, 2010). But exceptional floods do not become less likely with time (Figure S3b); even after a century of observations, it is still possible to experience a flood that is substantially larger than all others that have occurred on a given reach of river. Those floods that have been least paralleled by other high flows at the same location (high Q_{max}:Q_2 ratios) are found mainly in small watersheds with drainage areas below 10,000 km^2 (Table 1). The most unprecedented flood in our network was generated by Rayado Creek, near Cimarron, New Mexico, on June 17, 1965. That flood, which had a discharge more than 10 times larger than the second-biggest event in the century-long record (255 m^3/s compared to 24 m^3/s), destroyed several campsites but did not cause any injuries or deaths. The remarkable 1969 Reedy Creek Flood in eastern Virginia was caused by Hurricane Camille, a Category Five hurricane (Schwartz, 1970), while exceptional floods on the White River (Nebraska) and Prairie Dog Creek (Kansas) were both the product of convective thunderstorms in early summer. For bigger watersheds (drainage areas greater than 10,000 km^2), most of the rivers with extremely high Q_{max}:Q_2 ratios (Table S1) are located in the Great Plains (including gages in Iowa, Kansas, Montana, Nebraska, North Dakota, and South Dakota). The 2008 Cedar River Flood (J. A. Smith, Baeck, Villarini, Wright, & Krajewski, 2013) happened 105 years after the gage at Cedar Rapids (Iowa) was installed. Even with knowledge of a historical flood in 1851, the 2008 flood was still nearly twice as large as any other flood on record, illustrating that even after more than a century and a half of observations the river still held some surprise.

How much do the results of FFA change when the flood-of-record is omitted? As would be expected, in practically all cases (8,367 out of 8,370), excluding Q_{max} causes the

TABLE 1 The most unparalleled floods on record in the United States

Site identification number	Site name	Drainage area (km^2)	Length of record (years)	Date of record flood	Q_{max}:Q_2
7208500	Rayado Creek near Cimarron, NM	168	93	June 17, 1965	10.6
1674200	Reedy Creek near Dawn, VA	45	53	August 20, 1969	8.1
6444000	White River at Crawford, NE	811	62	May 10, 1991	7.8
6847900	Prairie Dog Creek above Keith Sebelius Lake, KS	1,528	54	May 23, 1953	7.4
10322980	Cole Creek near Palisade, NV	30	53	June 1983 (no date)	6.4
1475000	Mantua Creek at Pitman, NJ	16	62	September 1, 1940	6.3
6792000	Cedar River near Fullerton, NE	3,160	56	July 19, 1950	6.2
6847500	Sappa Creek near Stamford, NE	9,946	71	June 24, 1966	5.8
11084500	Fish Creek near Duarte, CA	16	59	January 25, 1969	5.8
2197190	McBean Creek at US 25, near McBean, GA	107	52	October 12, 1990	5.4

estimated magnitude of floods with long recurrence intervals to go down (Figure 2a). The effect is more pronounced for larger and more rare events, with the median 100-year flood declining by −10.5%, the 200-year flood by −11.8%, and the 500-year flood by −13.4%. For roughly a quarter of all gages, the estimated magnitude of the 500-year event is reduced by 20% or more compared to the FFA estimate obtained when Q_{max} is included. And, the more exceptional the flood, the larger its weight in the FFA, especially upon the higher portions of the curve (Figure 2b). To return again to the case of Rayado Creek, if the 1965 flood-of-record (which has an estimated return period of 2,550 years) was not known, the estimated magnitude of the 100-year flood would drop by 45%, the 200-year flood by 51%, and the 500-year flood by 58%). But even when the flood-of-record is not quite so exceptional, the very largest flood still has substantial weight within the FFA. For those more typical gages where Q_{max} is "only" 20% larger than Q_2, adding the flood-of-record to the peak flow sequence still raises the estimated magnitude of the 100-year flood by more than 10%. Finally, what recurrence interval would be estimated for Q_{max} itself? Because the 500-year flood often serves as an upper limit for risk assessments in the United States and other jurisdictions (Bell & Tobin, 2007; Ludy & Kondolf, 2012; Sauer, Thomas Jr., Stricker, & Wilson, 1983), it can be used here as a benchmark to evaluate the likelihood of a future flood just as large as the flood-of-record. In a quarter of all cases, if Q_{max} was unknown, a flood with the same

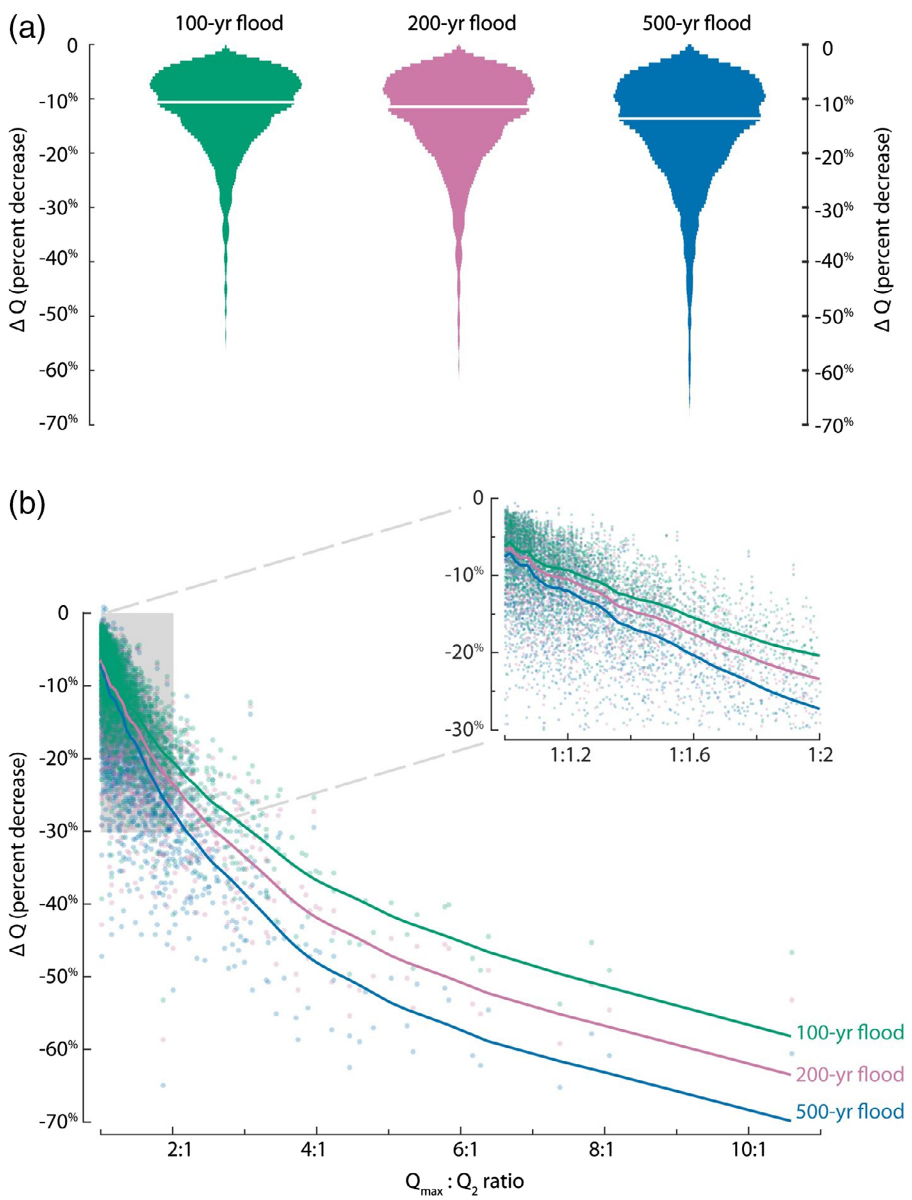

FIGURE 2 Relative effect of excluding the flood-of-record from flood frequency analysis (FFA) conducted on long-term streamflow records from the United States. (a) Violin plots showing the percent decrease (relative to FFA conducted on the complete flood sequence) in the magnitude of the 100-, 200-, and 500-year flood at all gages. The horizontal white lines mark the median of each distribution. (b) Scatterplot comparing the decrease in flood magnitude to the degree of exceptionalness of the flood-of-record (the Q_{max}:Q_2 ratio). The lines are Lowess smoothing curves that follow the 100- (green), 200- (violet), and 500-year floods (blue)

magnitude would be assigned a recurrence interval greater than 500 years. On the other hand, if Q_{max} is included in the FFA, more than four times out of five, that event falls below the 500-year threshold and would be considered as a rare but plausible future threat.

We find that the results of the heavy-tail index estimation are also sensitive to the omission of the flood-of-record. For the 146 long gages (Figure 3), if the flood-of-record is retained, the estimated values of α lie between a minimum of 0.25 (strong heavy tail) and a maximum of 1.99 (close to a normal distribution). The average equals 1.58 ± 0.02 (one-sigma *SE* of the mean), which is in close agreement to the value of 1.48 ± 0.13 (one-sigma root-mean squared error based on simulations) found for the 207-year long record from the Elbe River, Germany (Mudelsee & Bermejo, 2017). When Q_{max} is omitted, its removal yields significantly "lighter" tails (that is, larger α values) in nearly all cases (143 out of 146), and the mean difference in alpha (calculated in a paired manner) is 0.12 ± 0.02 (one-sigma *SE* of the mean of differences). Overall our α-tests demonstrate that the smaller the tail probability, the larger the effect. For example, consider an idealised example where the true underlying distribution of the annual maxima is a Generalised Extreme Value distribution with parameters location 0, scale 1, and shape 1/α (Mudelsee, 2014) and the gage record includes 100 annual flood observations. In that case, removing Q_{max} would cause a true 100-year event to have an estimated return period of 121 years, the 200-year flood to shift to a 252-year event, and the 500-year flood to be described as a 670-year flood. These findings of a significant effect on the heavy-tail index is corroborated by a Monte Carlo experiment, where 10,000 random series with a stable distribution (Mudelsee & Bermejo, 2017), a prescribed $\alpha = 1.5$, and a sample size of 100 were generated.

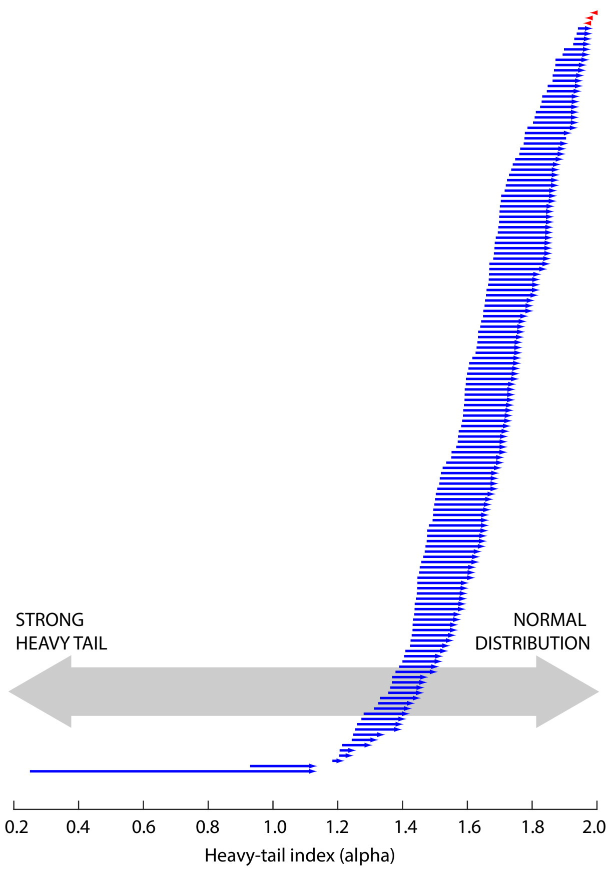

FIGURE 3 Floating bar chart illustrating the heavy-tail index values (α) for gage records longer than 100 years ($n = 146$). For each gage, the left-most position of the horizontal bar marks the α calculated from the full hydrological record, while the right-most position shows the same metric after Q_{max} is excluded from the analysis. The horizontal length represents the α − difference for each gage, with either values either more normal (blue) or having a heavier tail (red) if Q_{max} is removed

4 | CONCLUSION

FFA is without question the most widely applied and influential tool used to make predictions about how often floods will occur in the future and how severe they might be. Hydrologists are well aware that FFA techniques are most accurate when applied to floods with recurrence intervals that are shorter than the duration of the available hydrological record (Dalrymple, 1960; Eychaner, 2015). And because extreme floods are rare and are usually too few to constitute a sample adequate for statistical analysis (Kochanek et al., 2014), at the upper end of the frequency curve, as the recurrence interval increases so too does the uncertainty of the estimate (Eychaner, 2015). Systematic flood observations on American rivers now span multiple decades (and in many cases, more than a century), but the results produced by FFA are still sensitive to the influence of the single largest flood on record. The flood-of-record has little influence on risk estimates for smaller floods with short return periods. But if decisions related to flood mitigation demand information about the 100-, 200-, or 500-year flood, then it is crucial to understand the largest or most rare events that have been generated by the river. And even if the FFA adopts statistical functions that are better able to model extreme values, such as a heavy tailed distribution, our results show it is still very difficult to gage the probability of floods that are more severe than those previously observed.

Even several centuries of hydrological observations can sometimes not be sufficient to anticipate the most exceptional floods. Eychaner (2015) described the 500-year sequence of flood stage on the Danube River recorded at Passau, Germany, and pointed out that, despite the extraordinary length of these data, it is still difficult to estimate accurately the magnitude of those floods with recurrence intervals greater than 100 years. For that reason, he argued that, for the very largest floods, knowing their maximum elevation is more important than relying on a highly uncertain estimate of their recurrence interval. In another study conducted on the Elbe River in central Europe, Mudelsee, Börngen, Tetzlaff, and Grünewald (2003) concluded that, despite having more than two centuries of uninterrupted daily streamflow measurements, it was still not possible to quantify the magnitude of floods with return periods greater than 100 years due to the inherent uncertainties. The same limitation also appears to hold in general for gaged rivers in the United States. For that reason, we recommend assessments of flood risk should incorporate observation and/or documentation by non-hydrologists, or natural recordings of past floods interpreted by paleohydrologists whenever those sources are available. Recurrence estimates for rare floods are highly uncertain and are likely to remain so, even as instrumental hydrological records become gradually longer. Because of that limitation, we recommend that knowledge of truly exceptional floods, whether obtained from direct hydrological measurements, observation and/or documentation by non-hydrologists, or natural recordings of past floods interpreted by paleohydrologists, must remain a priority for research on hydrological extremes.

ACKNOWLEDGEMENTS

This work was conducted during a Humboldt Research Fellowship for Experienced Researchers awarded to S.St.G., and is the product of a collaboration fostered by the PAGES Floods Working Group. Additional support was provided by the Goethe-Institut. We thank Victor Baker, Lin Ji, Tao Liu, and an anonymous reviewer for comments on a prior version of the manuscript that improved the final product substantially.

ORCID

Scott St. George https://orcid.org/0000-0002-0945-4944

REFERENCES

Anderson, P. L., & Meerschaert, M. M. (1998). Modeling river flows with heavy tails. *Water Resources Research*, *34*, 2271–2280.

Avilia, L. A., & Cangialosi, J. (2011). *Tropical cyclone report. Hurricane Irene* (p. 45). Miami, FL: National Hurricane Center.

Bell, H. M., & Tobin, G. A. (2007). Efficient and effective? The 100-year flood in the communication and perception of flood risk. *Environmental Hazards*, *7*, 302–311.

Belt, C. B., Jr. (1975). The 1973 flood and man's constriction of the Mississippi River. *Science*, *189*, 681–684.

Aon Benfield (2016). *2016 Annual global climate and catastrophe report.* Retrieved from http://thoughtleadership.aonbenfield.com/sitepages/display.aspx?tl=638.

Benson, M. A., & Carter, R. W. (1973). *A national study of the streamflow data-collection program* (U.S. Geological Survey Water-Supply Paper 2208), 44 p.

Brooks, G. R., & St. George, S. (2015). Flooding, structural flood control measures, and a geomorphic context for the flood problem along the Red River, Manitoba, Canada. In P. Hudson & H. Middelkoop (Eds.), (pp. 87–117). Springer.

Burkey, J. (2009). *Log-Pearson flood flow frequency using USGS 17B, Mathworks® File Exchange*. Retrieved from https://www.mathworks.com/matlabcentral/fileexchange/.

Clark, A. O. (1996). Estimating probable maximum floods in the Upper Santa Ana Basin, Southern California, from stream boulder size. *Environmental & Engineering Geoscience*, *II*(2), 165–182.

Colle, B. A., & Mass, C. F. (2000). The 5–9 February 1996 flooding event over the Pacific northwest: Sensitivity studies and evaluation of the MM5 precipitation forecasts. *Journal of Climate*, *128*, 593–617.

Cook, E. R., & Krusic, P. J. (2004). *The north American drought atlas*. Palisades, NY: Lamont-Doherty earth observatory and the National Science Foundation.

Crippen, J. R., & Bue, C. D. (1977). *Maximum flood flows in the conterminous United States* (*United States Geological Survey Water Supply Paper 1887*). *52* p.

Dalrymple, T. (1960). *Flood frequency analysis* (U.S. Geological Survey Water Supply Paper, 1543A). 80 pp.

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

Dewan, A. M., Islam, M. M., Kumamoto, T., & Nishigaki, M. (2007). Evaluating flood hazard for land-use planning in Greater Dhaka of Bangladesh using remote sensing and GIS techniques. *Water Resources Management*, *21*, 1601–1612.

El Adlouni, S., Bobée, B., & Ouarda, T. B. M. J. (2008). On the tails of extreme event distributions in hydrology. *Journal of Hydrology*, *355*, 16–33.

England, J. F., Jr., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas, W. O., Jr., Veilleux, A. G., … Mason, R. R., Jr. (2018). *Guidelines for determining flood flow frequency—Bulletin 17C* (p. 148). Reston, VA: United States Geological Survey Techniques and Methods, Book 4, Chapter B5.

Eychaner, J. H. (2015). *Lessons from a 500-year record of flood elevations*. Madison, WI: Association of State Floodplain Managers, Technical Report 7.

Feaster, T. D. (2010). *Importance of record length with respect to estimating the 1-percent chance flood.* Proceedings of the 2010 South Carolina Water Resources Conference, Columbia, South Carolina.

Frank, N. L. (1965). The 1964 hurricane season. *Weatherwise*, *18*(1), 18–25.

Frazer, A.H., & Heckler, W. (1989), *Embudo, New Mexico, birthplace of systematic stream gaging* (Professional Paper 778). United States Geological Survey, 23 p.

Galewsky, J., & Sobel, A. (2005). Moist dynamics and orographic precipitation in northern and Central California during the New Year's Flood of 1997. *Journal of Climate*, *133*, 1594–1612.

Ganoulis, J. (2003). Risk-based floodplain management: A case study from Greece. *International Journal of River Basin Management*, *1*(1), 41–47.

Goodwell, A. E., Zhu, Z., Dutta, D., Greenberg, J. A., Kumar, P., Garcia, M. H., … Jacobson, R. B. (2014). Assessment of floodplain vulnerability during extreme Mississippi River Flood 2011. *Environmental Science and Technology*, *48*, 2619–2625.

Guha-Sapir, D., Hoyois, P., & Below, R. (2016). *Annual disaster statistical review 2015: The numbers and trends*. Brussels: Centre for Research on the Epidemiology of Disasters, Catholic University of Louvain.

Gumbel, E. J. (1941). The return period of flood flows. *The Annals of Mathematical Statistics*, *12*, 163–190.

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis—An approach based on L moments*. Cambridge, England: Cambridge University Press.

Hughey, E. P., & Tobin, G. A. (2006). Hazard response capabilities of a small community: A case study of Falmouth, Kentucky and the 1997 flood. *Southeastern Geographer*, *46*, 66–78.

Katkins, U., Davis Todd, C., Wojno, S., & Coleman, N. (2013). Revisiting the timing and events leading to and causing the Johnstown flood of 1889. *Pennsylvania History: A Journal of Mid-Atlantic Studies*, *80*, 335–363.

Kellens, W., Terpstra, T., & De Maeyer, P. (2012). Perception and communication of flood risks: A systematic review of empirical research. *Risk Analysis*, *33*, 24–49.

Klemeš, V. (1989). The improbable probabilities of extreme floods and droughts. In O. Starosolszky & O. M. Melder (Eds.), *Hydrology of disasters* (pp. 43–51). London, England: James & James.

Kochanek, K., Renard, B., Arnaud, P., Aubert, Y., Lang, M., Cipriani, T., & Sauquet, E. (2014). A data-based comparison of flood frequency analysis methods used in France. *Natural Hazards and Earth System Sciences*, *14*, 295–308.

Lave, T. R., & Lave, L. N. (1991). Public perception of the risks of floods: Implications for communication. *Risk Analysis*, *11*, 255–267.

Leathers, D. J., Kluck, D. R., & Kroczynski, S. (1998). The severe flooding event of January 1996 across north-Central Pennsylvania. *Bulletin of the American Meteorological Society*, *79*, 785–797.

Lucia, E. (1965). *Wild water. The story of the far West's great Christmas week floods* (Vol. 72). Portland, OR: Overland West Press.

Ludy, J., & Kondolf, G. M. (2012). Flood risk perception in lands "protected" by 100-year levees. *Natural Hazards*, *61*, 829–842.

Mertz, R., & Blöschl, G. (2008). Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resources Research*, *44*, W08432.

Michel-Kerjan, E., & Kunreuther, H. (2011). Redesigning flood insurance. *Science*, *333*, 408–409.

Mudelsee, M. (2014). *Climate time series analysis: Classical statistical and bootstrap methods* (2nd ed., p. 454). Berlin, Germany: Springer.

Mudelsee, M., & Bermejo, M. A. (2017). Optimal heavy tail estimation, part I: Order selection. *Nonlinear Processes in Geophysics Discussions*, *24*, 737–744.

Mudelsee, M., Börngen, M., Tetzlaff, G., & Grünewald, U. (2003). No upward trends in the occurrence of extreme floods in Central Europe. *Nature*, *425*, 166–169.

Namais, J. (1972). Hurricane Agnes – an event shaped by large-scale air-sea systems generated during antecedent months. *Quarterly Journal of the Royal Meteorological Society*, *99*, 506–519.

Parkes, B., & Demeritt, D. (2016). Defining the hundred year flood: A Bayesian approach for using historic data to reduce uncertainty in flood frequency estimates. *Journal of Hydrology*, *540*, 1189–1208.

Porter, J., & Demeritt, D. (2012). Flood risk management, mapping and planning: The institutional politics of decision-support in England. *Environment and Planning A*, *44*, 2359–2378.

Rahman, A., Zaman, M. A., Haddad, K., El Adlouni, S., & Zhang, C. (2015). Applicability of Wakeby distribution in flood frequency analysis: A case study for eastern Australia. *Hydrological Processes*, *29*, 602–614.

Resnick, S. I. (2007). *Heavy-tail phenomena: Probabilistic and statistical modeling* (p. 404). Berlin, Germany: Springer.

Sauer, V. B., Thomas, W. O., Jr., Stricker, V. A., & Wilson, K. V. (1983). *Flood characteristics of urban watersheds in the United States* (United States Geological Survey Water Supply Paper 2207).

Schwartz, F. K. (1970). The unprecedented rains in Virginia associated with the remnants of Hurricane Camille. *Monthly Weather Review*, *98*, 851–859.

Seed, H. B., & Duncan, J. M. (1987). The failure of Teton Dam. *Engineering Geology*, *24*, 173–205.

Shelton, M. L. (1977). The 1976 and 1977 drought in California: Extent and severity. *Weatherwise*, *30*, 139–153.

Smith, A., Sampson, C., & Bates, P. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, *51*, 539–553.

Smith, J. A., Baeck, M., Villarini, G., Wright, D. B., & Krajewski, W. (2013). Extreme flood response: The June 2008 flooding in Iowa. *Journal of Hydrometeorology*, *14*, 1810–1825.

St. George, S., & Rannie, B. (2003). The causes, progression and magnitude of the 1826 Red River flood in Manitoba. *Canadian Water Resources Journal*, *28*, 99–120.

Tennessee Valley Authority. (1963). *Floods on Swannanoa River and Beetree Creek in vicinity of Swannanoa, North Carolina.* Report No. 0-6198, Knoxville, Tennessee, 118 p.

Todhunter, P. E. (2001). A hydroclimalogical analysis of the Red River of the North snowmelt flood catastrophe of 1997. *Journal of the American Water Resources Association*, *37*, 1263–1278.

United States Geological Survey. (1982). *Guidelines for determining flood flow frequency*. Bulletin 17B. Reston, Virginia: Hydrology Committee.

United States Geological Survey. (2016). *National Water Information System data available on the World Wide Web (USGS Water Data for the Nation).* Retrieved from http://waterdata.usgs.gov/nwis/.

Vogel, R. M., Zafirakou-Koulouris, A., & Matalas, N. C. (2001). Frequency of record-breaking floods in the United States. *Water Resources Research*, *37*, 1723–1731.

Waananen, A. O., Harris, D. D., & Williams, R. C. (1971). *Floods of December 1964 and January 1965 in the far western states; part 1 description* (Water Supply Paper 1866-A). United States Geological Survey, 265 p.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: St. George S, Mudelsee M. The weight of the flood-of-record in flood frequency analysis. *J Flood Risk Management*. 2019;12 (Suppl. 1):e12512. https://doi.org/10.1111/jfr3.12512