

RESEARCH ARTICLE OPEN ACCESS

Bias in Peak Flood Discharges: Are Our Bridges and Culverts Undersized?

Steven E. Yochum¹  | Tyler Wible²

¹U.S. Forest Service, National Stream and Aquatic Ecology Center, Fort Collins, Colorado, USA | ²Colorado State University, One Water Solutions Institute, Fort Collins, Colorado, USA

Correspondence: Steven E. Yochum (steven.yochum@usda.gov)

Received: 16 June 2025 | **Revised:** 23 October 2025 | **Accepted:** 26 November 2025

Keywords: bias | bridges | culverts | flooding | floodplains | infrastructure | peak discharge

ABSTRACT

Reliable methods for peak discharge predictions at ungaged locations are required for infrastructure design and floodplain management. For decades, a standard practice in the United States has been to utilize US Geological Survey regional regression equations (StreamStats) as a singular method. However, implementation of multiple methods, such as streamgage-based transfers, rainfall-runoff modeling, and downscaled climate modeling, provides a more rigorous approach. The Flood Potential Portal (<https://floodpotential.erasms.com/>) Watershed Analysis module provides three streamgage-based methods for predictions, facilitating comparative evaluations of prediction bias across the United States. Our findings indicate that regional regression equations, overall, significantly underpredict 100-year (1% chance of exceedance; Q_{100}) peak discharges by -26.6% . In contrast, index flood frequency (Q_{100}) and expected flood potential discharge (Q_{efp}) show minimal bias of $+4.4\%$ and 0% , respectively. Within the 18 analyzed watersheds, regional regression equations underpredict in 15 (-55.1% to -1.8% , average: -27.8%), predict best in two, and overpredict in one. The other methods yield far less relative bias. Underprediction of -55.1% results in road-stream crossings having less than half the needed capacity, and an average bias of -27.8% indicates that crossings may be commonly undersized by more than a quarter—use of only regional regression equations for peak discharge predictions for bridges and culverts may result in infrastructure having insufficient flood resilience. A case study for a watershed impacted by the remnants of Hurricane Helene reveals underprediction of up to -53.2% ; peak discharge bias could result in undersized structures. Resilient recovery requires the use of multiple predictive methods.

1 | Introduction

Peak discharge magnitudes are required for the management and design of road-stream crossings and other stream valley infrastructure across the United States, to size bridge and culvert openings, as well as for floodplain and riparian ecosystem management. Most locations with such infrastructure do not have streamgage records—reliable prediction methods for ungaged locations are needed. Where streamgage data exists, a greater understanding of the reliability of standard statistical analyses for predicting appropriate design flood discharges is needed. For both ungaged and gaged locations, multiple predictions for

each site are preferred for decision making, for practitioners to inspect for precision and note if any methods are systematically biased relative to each other. From this information, professional judgment is used to select the most appropriate design flood discharge estimate.

For decades, a standard practice to quantify magnitudes at ungaged locations has been to utilize peak flow statistics from the results of regional regression equations (Miller 2003; Kenney et al. 2007; Waltemeyer 2008; Capesius and Stephens 2009; Kohn et al. 2016) developed by the US Geological Survey (USGS). These are typically generated within the StreamStats

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *River Research and Applications* published by John Wiley & Sons Ltd.



FIGURE 1 | Reconstruction of infrastructure along Rock Creek in Red Lodge, Montana, after the Yellowstone region floods of June 2022. Photograph collected on June 22, 2023. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/rra.7091)]

web application (Ries et al. 2024). These regressions are based on streamgauge flood-frequency analyses using the Bulletins 17B (IACWD 1982) and 17C (England et al. 2018) procedures. Are there systematic (relative) biases present with these predictions?

- *Underestimation* results in structures designed with undersized openings for flood passage, overtopped roads and structures, avoidable damage and failures, as well as underestimated base flood elevations and floodplains with insufficient extents.
- *Overestimation* leads to unnecessary additional infrastructure costs and excessive floodplain regulatory constraints.

Underestimated peak flood discharges pose the greatest risk to lives, property, and infrastructure. Flooding is a leading cause of casualties from natural disasters, with tens to hundreds of deaths each year in the United States (Han and Sharif 2021)—with underestimation, preventable deaths likely occur more frequently and in greater numbers. Lack of flood resilience from low biases in peak discharge estimates, if present, causes the impacts of catastrophic floods to be more extensive than necessary, instead of mitigating these impacts through stream corridor planning, engineered solutions, and flood insurance. Recent expansive riverine events, such as the Yellowstone region floods in June 2022 (Figure 1) and floods induced in the Southern Appalachians by powerful remnants of Hurricane Helene in September 2024 highlight the risks, with peak discharge prediction bias potentially increasing the severity of devastation.

Climate change is modifying flood hazards in some locations in a variety of ways. Observed changes in flood severity, or nonstationarity (Lins and Cohn 2011; Galloway 2011; Francois et al. 2019) as monitored by the USGS streamgaging network, have been quantified using the flood potential method (Yochum et al. 2019, 2024; USGCRP 2023) and presented in the Flood Potential Portal (<https://floodpotential.erams.com/>) Mapping

module, with tests for trends in 5 scales of flood magnitudes, annual and event flood frequencies, and flashiness. Increasing trends in large flood magnitudes have been quantified in 56 (of 206) zones of similar flood response across the contiguous United States, with an average change (where increasing trends are present) of +6.7%, and a maximum of +33.8% observed in a portion of Western North Carolina. Decreasing trends in large flood magnitudes have been quantified in 57 other zones, with an average change of −13.6%. Flood severity is increasing in some areas and decreasing in other areas, though these trend analyses are limited by the decreasing availability of streamgauge data in many portions of the United States.

Peak discharge estimation for design and planning of stream valley infrastructure and floodplain management requires a balance between rigor and pragmatism; interest in the use of complex methods and tools or strict adherence to singular approaches needs to be balanced by the demands of design practice as a licensed professional, the issue that no single method for flood discharge predictions is ideal in all circumstances, and given such situations as disaster recovery when thousands of estimates of design flood discharges may be needed for the most effective use of billions of dollars of infrastructure reconstruction funds. The use of multiple simple methods is an approach for addressing this issue, with replication needed to avoid the introduction of bias and future flood susceptibility. One simple approach, the rational method, has been documented to perform relatively well for predictions in small (headwater) watersheds (Mukherjee et al. 2024), but this method is limited to the smallest catchments. In regard to more complex approaches, advances are continuing to be made in rainfall-runoff modeling (Kumar et al. 2023), including the incorporation of future temporal trends in the precipitation-frequency estimates presented in NOAA Atlas 15 (NOAA 2025), in climatological modeling (Li et al. 2022; Keller et al. 2022), and with the use of a variety of machine learning and other artificial intelligence methodologies (Kumar et al. 2023; Jones et al. 2023), but these methods can take substantial resources and time to implement for site specific design and planning activities.



FIGURE 2 | Eighteen test watersheds with peak flood discharges computed using the Flood Potential Portal Watershed Analysis module, for a census of road-stream crossings with a Strahler order ≥ 3 . [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/rra.70991)]

While it is important to incorporate observed increasing flood trends or prognostication of future changes in floods into peak discharge estimates, how do these changes compare to biases induced by analytical methodologies? And if low biases do exist, is there a false narrative linking causation with climate change when such biases may instead be the result of data eccentricities and analysis aberrations? To test, the Flood Potential Portal (<https://floodpotential.erasm.com/>) Watershed Analysis module was used to estimate 100-year scale (1% annual exceedance probability) peak discharges for a range of watershed sizes using three methods that transfer streamgage-based data to ungaged locations: flood potential, index flood frequency, and USGS regional regression equations flood frequency (Yochum et al. 2024). Consistently having three sets of results at ungaged locations, using one tool, enables practitioners to more thoroughly evaluate historical streamgage data and therefore make more informed decisions for greater flood resiliency, while working consistently and efficiently. Estimates of design peak discharges at road-stream crossings were developed for 18 watersheds distributed across the contiguous United States. All bridge, culvert, and ford crossings for streams with a Strahler stream order ≥ 3 were included in the analyses, for each watershed. Of special interest is to understand if predictions made using the commonly implemented results of USGS regional regression equations frequently underpredict flood hazards; anecdotal experiences of practitioners have raised concerns regarding such underprediction, but this requires systematic evaluation. If such predictions are biased low, undersizing of bridge and culvert openings as well as the underestimation of floodplain risk results in insufficient flood resiliency. Furthermore, systematically undersized stream valley infrastructure and floodplain extents enhance exposure to increasing flood

magnitudes, worsening the impacts of climate variability and change where floods are becoming more severe.

2 | Methods

To systematically assess bias across the contiguous United States, a watershed should be sampled in each of the 206 flood potential zones (Yochum et al. 2024). For this initial analysis, 18 watersheds (Figure 2) were selected for a census of road-stream crossings (Table 1) in 18 spatially distributed zones, within the framework of the flood potential methodology (Yochum et al. 2019). These watersheds were selected to maximize diversity in precipitation, relief, flood potential, flashiness, and location. Each watershed was entirely within a single zone (flood potential method analysis area), and did not have major reservoir regulation. It was unknown prior to each watershed selection whether there was any systematic bias for any of the three approaches for predictions of peak discharges (a priori).

The Flood Potential Portal (<https://floodpotential.erasm.com/>) Watershed Analysis module was used to generate peak discharge magnitudes for each site. This module provides predictions of peak flood discharge using three streamgage-based methods: flood potential (Yochum et al. 2019), index flood frequency (Dalrymple 1960; Hosking 1990; Hosking and Wallis 1993; Yochum et al. 2024), and USGS regional regression equations flood frequency (Miller 2003; Kenney et al. 2007; Waltemeyer 2008; Capesius and Stephens 2009; Kohn et al. 2016) as computed in StreamStats (Ries et al. 2024) and presented in the Flood Potential Portal. The flood potential and index flood-frequency methods use watershed area as the primary predictor

TABLE 1 | Analyzed watersheds, with characteristics.

Watershed	Area		Ave P		Ave S	Zone	P_f		R-B	
	(km ²)	(mi ²)	(mm)	(in)	(deg)		Value	%	Value	%
Big Creek, KS	1422	549	569	22	1.6	62	15.6	64th	0.55	81st
Broad River, GA	1963	758	1336	53	6.8	74S	14.4	60th	0.44	66th
Clark Fork, MT	2557	987	480	19	11.8	44	1.5	6th	0.07	2nd
S.F. Crow River, MN	2889	1115	744	29	1.1	67S	2.8	16th	0.1	8th
Eagle River, CO	2446	945	650	26	17.7	3	2.3	13th	0.10	8th
M.F. Eel River, CA	1928	745	1503	59	18.6	25	31.4	91st	0.42	64th
Homochitto River, MS	2050	791	1585	62	4.2	83SW	45.0	96th	0.61	86th
Kaweah River, CA	1350	521	956	38	23.3	23	16.6	68th	0.26	39th
Little Miami River, OH	950	367	1064	42	2.2	93	15.6	64th	0.48	72nd
Meramec River, MO	2023	781	1180	46	5.6	59N	30.9	90th	0.51	75th
Mora River, NM	2848	1100	554	22	8.1	4	15.1	63rd	0.59	84th
Pemigewasset River, NH	1614	623	1398	55	12.6	106	23.5	80th	0.36	55th
S.B. Potomac River, WV	3782	1460	1054	41	16.7	72N	34.8	93rd	0.37	57th
Redwater River, MT	1432	553	313	12	3.5	49	7.8	41st	0.51	75th
Sauk River, WA	1855	716	2690	106	27.1	38	25.9	86th	0.25	37th
Shoal Creek, TN	900	348	1534	60	7.8	81	37.1	94th	0.54	77th
Spanish Fork, UT	1628	628	591	23	18.3	9	1.8	8th	0.08	3rd
Tonto Creek, AZ	1740	672	611	24	15.6	16	16.8	68th	0.64	87th

Abbreviations: %, percentile rank, compared to 207 zones across the contiguous United States; P, average watershed annual precipitation; P_f , zone flood potential index; R-B, zone-average Richards-Baker flashiness index; S, average watershed slope; Zone, flood potential zone ID.

and a secondary predictor (if significant) of average watershed slope, elevation, aspect, or average annual or monthly precipitation. The equations and analysis periods for each zone are provided as summary information within the Flood Potential Portal Mapping module. A list of datasets used in these analyses is provided in Table 2. The USGS regional regression equations and index methods provide flood-frequency estimates based on annual peak discharge estimates at each streamgage within regions and flood potential zones, respectively. The flood potential method is a relatively simple analysis that identifies the central tendency of large floods across each of 206 zones delineated across the contiguous United States, using regressions of record peak discharges collected at more than 8200 streamgages. Explained variances of these regressions are high, with an average $R^2 = 0.93$. Considering the complex meteorological and hydrological processes that generate large floods, and the observed high spatial variability in large flood magnitudes, the flood potential method assumes that a reasonable design flood discharge for many applications is a value that has an even chance of both falling short and exceeding a flood magnitude—this central tendency of large floods is the expected flood potential discharge.

The 100-year peak discharge (Q_{100}) predictions for index flood and regional regression equations were compared to the expected flood potential discharge (Q_{efp}), which has been found to be, generally, insignificantly different from Q_{100} discharges (Yochum et al. 2019). While the “true” flood hazard at any site

is unknown, may be changing, and may be unknowable, results from multiple analyses can indicate a range of what a true flood hazard may be at each site (with a larger number of predictions providing greater confidence in this range, using an ensemble prediction strategy). The intent of this analysis is to initially explore possible bias using three methods that are repeatable using the Flood Potential Portal Watershed Analysis module, as an exploration of issues in flood magnitude prediction techniques that may be causing infrastructure to be undersized and flood-plain risk to be understated.

All road-stream crossings with a Strahler stream order ≥ 3 in each test watershed had a watershed analysis performed in the Flood Potential Portal, as a census of crossings within the transportation system. An example watershed boundary and the analyzed road-stream crossings are provided for the Eagle River Watershed, Colorado (Figure 3). The National Hydrology Dataset (NHD) was used for the crossing determinations (ArcGIS source: <https://basemap.nationalmap.gov/arcgis/rest/services/USGSHydroCached/MapServer/0>). Each of these locations is a road crossing (bridge, culvert, or ford); this approach quantifies bias in each watershed at locations where such estimates are actually needed, to maintain, repair, or replace existing structures to maintain transportation networks.

Modifications to this general procedure were occasionally utilized. Specifically, stream order 3 or 4 (and higher) channels

TABLE 2 | Summary of utilized datasets, with sources and version year or period of record.

Dataset description	Source	Version/ period
Flood Potential Zones and Algorithms	US Forest Service	2024
Index Flood-Frequency Algorithms	One Water Solutions Institute	2024
Regional Regression Equations Results	US Geological Survey, StreamStats	2024
Streamgage Annual Peak Discharges	US Geological Survey	1815–2024
Annual and Monthly Precipitation, 2010	PRISM Group	1980–2010
Annual and Monthly Precipitation, 2020	PRISM Group	1990–2020
DEM Elevation Data	US Geological Survey	2010
Surface Water Extent	US Geological Survey	2022

Abbreviation: DEM, digital elevation model.

in roadless portions of the Tonto Creek, Kaweah River, and the Middle Fork Eel River watersheds had analysis points added immediately above confluences despite road crossings not being present; this was necessary to ensure a consistent point density in these roadless areas. Due to the high stream channel density represented by NHD in Mississippi, a Strahler stream order of 4 was used as the analysis threshold for the Homochitto River watershed. Additionally, redundant crossings and home driveways, artificial channels, sites below flood-control reservoirs, and locations (on small streams) where the watershed was too small for application of the flood potential method or where StreamStats repeatedly failed to provide results within the flood potential portal were excluded.

It is most reasonable to assume that the median result is preferred for design and planning purposes (Yochum et al. 2024), to avoid both underestimation and overestimation, though some settings may justify the use of a higher value to avoid the risk of underestimation and to account for associated flood risks, such as large wood (debris) conveyance. In this analysis, it was assumed that the median result is most appropriate for use, and dividing the three individual results at each location by this median value provides a relative bias ratio (β). If β is < 1 , the individual estimate underpredicts, and if β is > 1 , the individual estimate overpredicts. If relative errors are random and unbiased, each analytical technique should have a similar number of sites with β values > 1 , < 1 , and $= 1$. However, if there is a disproportionate result for any technique with $\beta < 1$, then that technique is biased low and underpredicts peak flood discharges. The percent (bias) underprediction or overprediction was computed as $(\text{median } \beta - 1) * 100$.

Flood-frequency analyses for streamgages with ≥ 10 years of data in each of the 18 watersheds were also performed, using the Flood Potential Portal Streamgage Analysis module. The log-Pearson distribution and methods detailed in Bulletin 17B (IACWD 1982) and Bulletin 17C (England et al. 2018) were used, using both station and weighted generalized skewness, yielding four sets of flood-frequency results at each streamgaged location. The largest Q_{100} discharge of these four was used for comparison to the median watershed analysis module result, for the computation of β ratios. This relative bias was compared to the bimodality index (B_p), computed for each streamgage as the maximum annual peak discharge divided by the median (Yochum et al. 2024).

3 | Results

Across the entire dataset of 1517 points in 18 watersheds, regional regression equations yield peak discharge Q_{100} values that significantly underpredict peak discharges (Figure 4; Table 3), with a relative bias of -26.6% . If sized using this common (StreamStats-based) method, these findings indicate that bridge and culvert openings could have commonly been undersized by 26.6% . The index method Q_{100} had an overprediction bias of 4.4% , though this method had the least variability (standard deviation $= 0.184$). The flood potential discharge Q_{eff} had no systematic bias but the greatest variability (standard deviation $= 0.343$), with this method substantially overpredicting peak discharge at some locations.

Grouping these results by watershed, 15 of the 18 assessed watersheds indicate significant Q_{100} underprediction in regional regression equations results (Table 3 and Figure 5). The relative bias ranges from -55.1% (Big Creek, Kansas) and -44.7% (South Branch Potomac River, West Virginia) to -1.8% (Middle Fork Eel River, California); use of such values could result in undersizing of road-stream infrastructure openings by half in some locations, and more than 25% in 10 of the 18 watersheds. Regional regression equations provided median (generally unbiased) predictions for two of the 18 watersheds (Broad River, Georgia; Little Miami River, Ohio)—USGS regional regression equations provide the best peak discharge estimates in some zones and watershed sizes. Finally, in the Tonto Creek, Arizona watershed, regional regression equations results are biased high ($+50.4\%$).

Relative bias often varies by watershed size (Figure 5). For example, the Clark Fork, Montana regional regression results overpredict peak discharges for larger watersheds ($> 300 \text{ km}^2$) but underpredict in smaller watersheds, while these results in the Sauk River, Washington watershed predict similarly to the index flood and expected flood potential results in larger watersheds but underpredict in smaller watersheds.

The flood potential (Q_{eff}) and index flood (Q_{100}) methods yield predictions with less relative bias (Figures 4 and 5), generally without systematic underprediction. Results from the flood potential method (Q_{eff}) underpredict in three watersheds (-0.7% , -2.3% , -17.2%), overpredict in 9 ($+1.1\%$ to $+105.4\%$), and are the median in 6 of the 18 watersheds. The flood potential method provided the most overpredicted peak discharges in this study

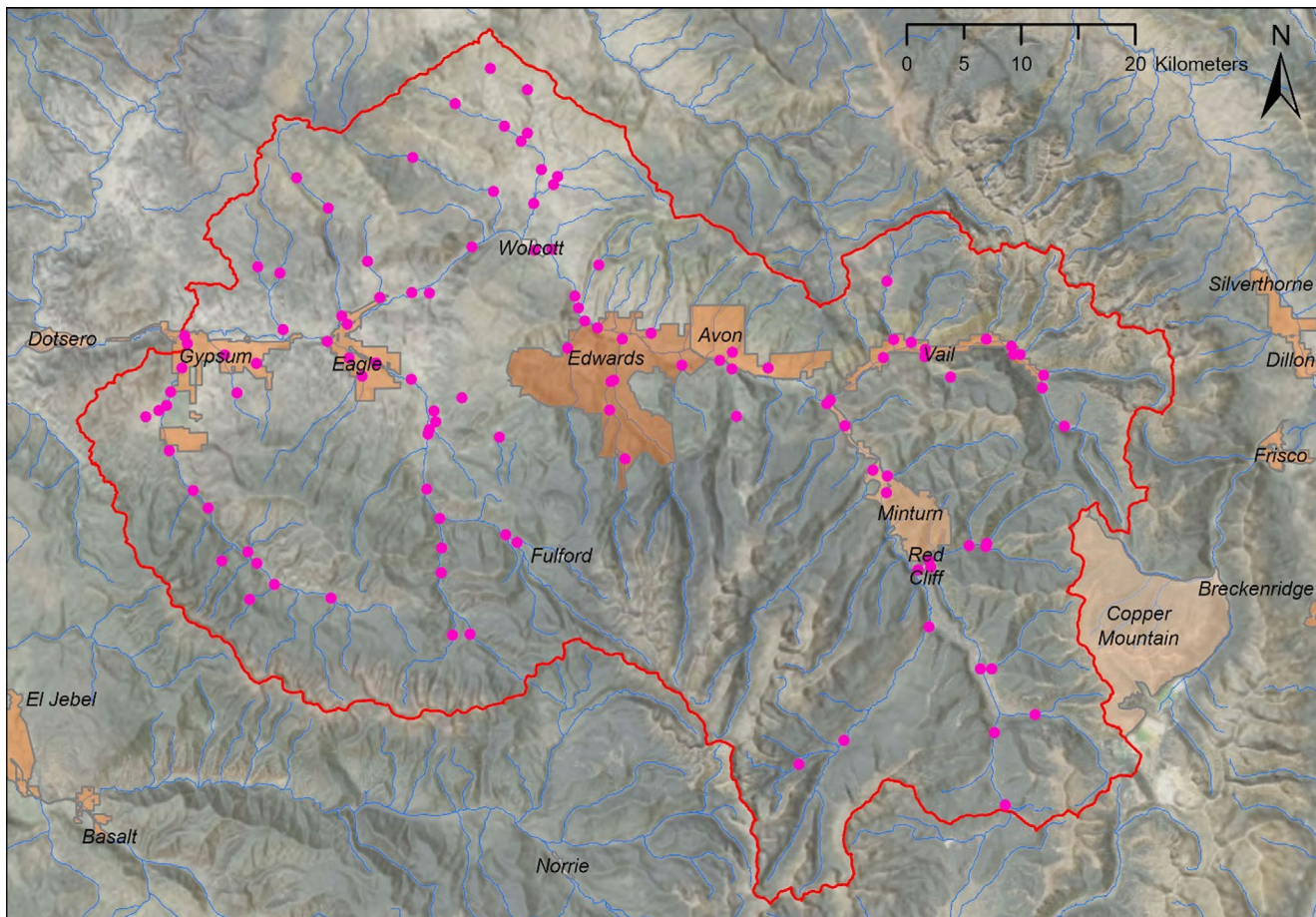


FIGURE 3 | Boundary for the Eagle River Watershed, with 116 analysis points, National Hydrography Dataset streams, and populated places. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

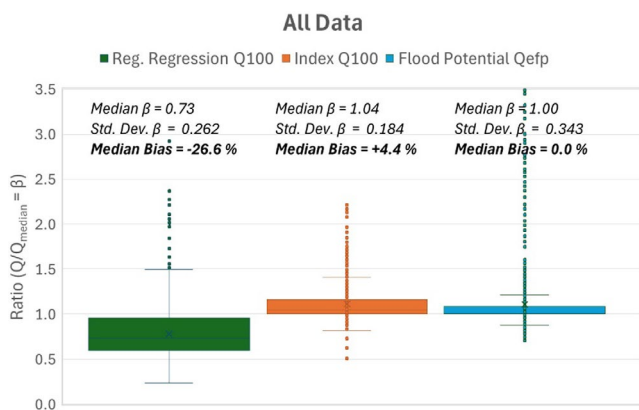


FIGURE 4 | Box plots of relative bias ratios ($\beta = Q/Q_{\text{median}}$), by method (USGS regional regression equations, index flood frequency, expected flood potential discharge), for the entire dataset. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

(+105.4%), in the Mora River, New Mexico watershed, with a β standard deviation of 0.844. The index flood method (Q_{100}) underpredicts in none of the watersheds, overpredicts in seven watersheds (+6.6% to +55.3%), and provides the most appropriate results in 11 of the watersheds—in these tested watersheds using these three prediction methods, the index flood method

is least susceptible to peak discharge underprediction and most commonly provides appropriate peak discharges for design. However, even in watersheds where, overall, the index flood method predicts without bias, for some watershed sizes, these predictions may not be the best. In the S.B. Potomac River watershed, for example, despite the zero overall bias for index flood results, the index flood method overpredicts in some smaller watersheds (Figure 5), and flood potential results are most appropriate.

4 | Discussion

Systematic underestimation of riverine peak flood discharges, resulting in undersized bridge and culvert openings for flood passage and insufficient floodplain extents, can represent an existential threat to transportation networks, property, and life. Accordingly, the selection of design flood and base flood discharges needs to be done with considerable care, using multiple methods to avoid biases that can be inadvertently introduced by any single approach. An available tool for achieving such multiple analyses is the Flood Potential Portal Watershed Analysis module (Yochum et al. 2024). This tool predicts peak flood discharges at ungaged locations using three streamgage-based methods, providing results for designs and the data used in this assessment.

TABLE 3 | Bias ratio comparisons using paired two-sided *t*-tests and median percent bias (median $\beta-1$)*100, for the overall analysis and the individual watershed analyses.

	Number of points	Paired two-sided <i>t</i> -test <i>p</i> -value			Median bias (percent)		
		Regional regression			Regional regression		
		Q_{100} vs. Q_{100} Index	Q_{eff} vs. regional regression Q_{100}	Q_{100} Index vs. Q_{eff}	Q_{100}	Q_{eff}	Q_{100} Index
Overall	1517	<2.2E-16	<2.2E-16	0.5040	-26.6	0.0	4.4
Big Creek, KS	79	<2.2E-16	<2.2E-16	6.3E-07	-55.1	6.3	0.0
Broad River, GA	136	<2.2E-16	0.0148	<2.2E-16	0.0	-2.3	7.5
Clark Fork, MT	73	1.57E-05	7.36E-06	0.0517	-26.7	2.4	0.0
S.F. Crow River, MN	71	0.0001	0.0016	0.0473	-4.0	5.4	0.0
Eagle River, CO	116	<2.2E-16	<2.2E-16	0.0120	-30.9	0.0	6.6
M.F. Eel River, CA	78	<2.2E-16	0.0052	<2.2E-16	-1.8	0.0	19.1
Homochitto River, MS	86	<2.2E-16	<2.2E-16	<2.2E-16	-39.6	0.0	55.3
Kaweah River, CA	60	<2.2E-16	<2.2E-16	1.15E-05	-29.6	7.7	0.0
Little Miami River, OH	44	0.716	0.0003	5.37E-07	0.0	-0.7	0.0
Meramec River, MO	119	<2.2E-16	<2.2E-16	2.48E-13	-21.2	7.6	0.0
Mora River, NM	92	<2.2E-16	<2.2E-16	3.85E-16	-40.0	105.4	0.0
Pemigewasset River, NH	80	<2.2E-16	<2.2E-16	2.89E-07	-28.4	9.0	0.0
S.B. Potomac River, WV	207	<2.2E-16	<2.2E-16	0.5546	-44.7	1.1	0.0
Redwater River, MT	69	<2.2E-16	<2.2E-16	4.59E-11	-28.9	0.0	8.1
Sauk River, WA	55	<2.2E-16	<2.2E-16	0.0002	-28.7	0.0	7.9
Shoal Creek, TN	76	<2.2E-16	1.25E-11	<2.2E-16	-14.9	0.0	33.3
Spanish Fork, UT	33	1.13E-05	<2.2E-16	9.91E-15	-22.7	27.3	0.0
Tonto Creek, AZ	42	3.56E-13	<2.2E-16	<2.2E-16	50.4	-17.2	0.0

Abbreviation: vs., versus.

Relative bias induced by USGS regional regression equations, a standard methodological approach for using streamgage data for predictions at unengaged locations, was found to be significant with the average bias (where present) of -27.8% being much greater (in absolute magnitude) than the average increasing trends in flood magnitudes of +6.7% (Yochum et al. 2024) due to climate variability and change, as well as other sources of non-stationarity. Moreover, the largest bias (-55.1%) is substantially greater than the magnitude of the largest observed increasing trend (+33.8%; zone 76: Southern Blue Ridge Mountains in North Carolina). Twelve of the 18 watersheds underpredict peak discharges by more than 20% (Figure 6). At this time, underprediction bias, where regional regression equations' results are used for design and planning, has much greater impacts on flood resilience than observed flood severity enhancements. This previously undocumented exposure to flood risk has always existed, but has been obscured within the nuance of data and analyses. This finding counters the common attribution of poor flood resilience to climate change. Moreover, where trends in flood magnitudes are increasing, this analytical bias compounds the effects of climate change and other sources of non-stationarity.

Systematic underprediction bias from regional regression equations may be induced by a number of causes, including: variable streamgage record lengths and periods, combined with the presence or absence of large floods in each record; the nature of the log-Pearson distribution utilized in Bulletins 17B and 17C methods, given the presence of bimodal and multimodal data (Yochum et al. 2024), also referred to as mixed population floods (England et al. 2018) and heavy tail behavior (Merz et al. 2022); and region extents that do not well represent the flood risk quantified by the streamgaging network. Comparison of the median Watershed Analysis module results with streamgage analysis results illustrates the influence of data variability and analytical methods (Figure 7). Bias in streamgage analysis results (compared to the median watershed analysis module results) generally increases with increasing bimodality ($r=0.697$)—at sites where the largest annual peak discharges are much greater than typical annual peaks (the median), streamgage analysis results may overpredict the true flood risk as computed using neighboring and similarly responding watersheds (flood potential zones). Though there is substantial variability in this general trend, with nuance by record length.

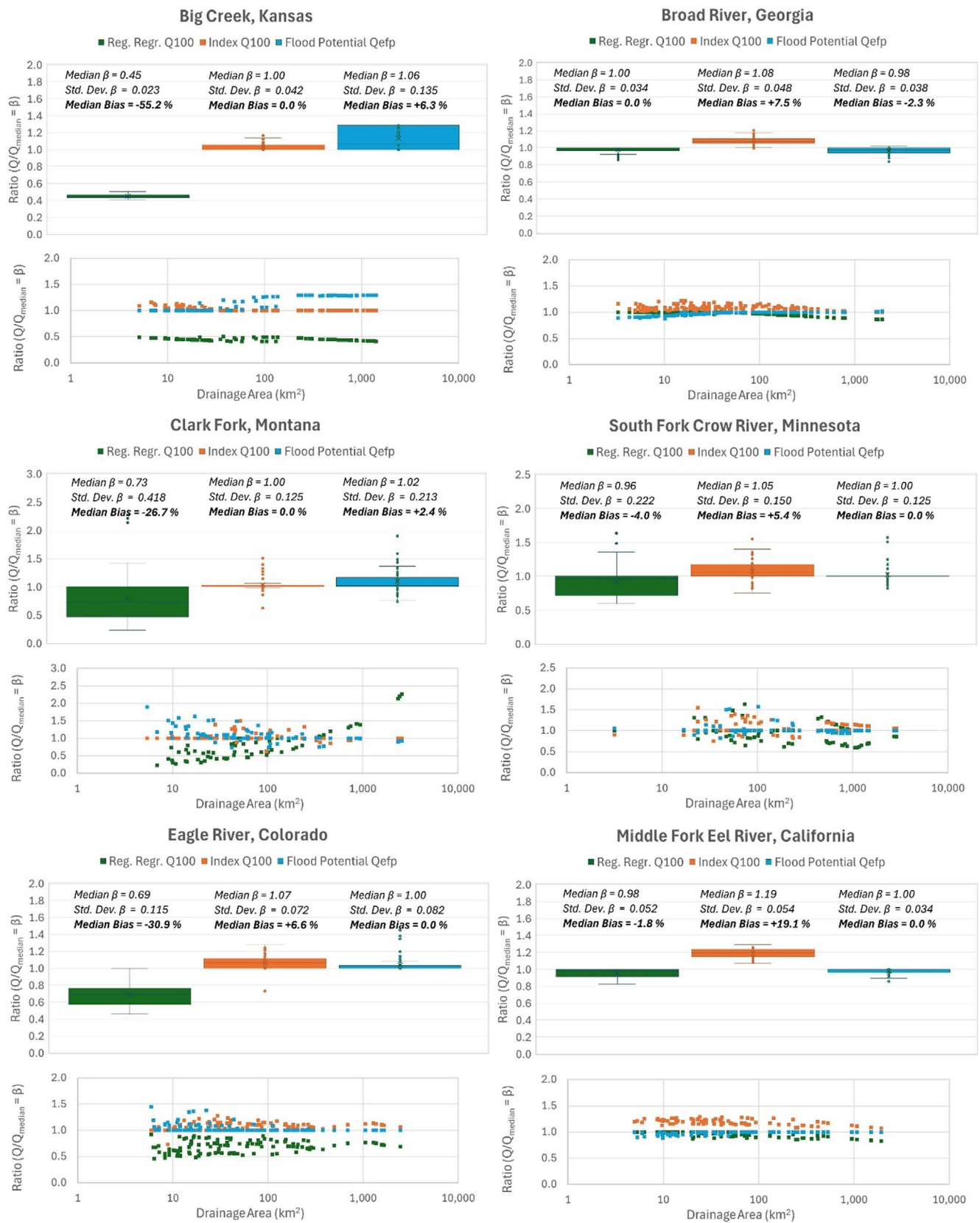


FIGURE 5 | Box plots and variability plots (by watershed size) of bias ratios ($\beta = Q/Q_{median}$) computed for each point, by method. Utilized methods: USGS regional regression equations, from StreamStats (Reg. Regr., Q_{100}); index flood frequency (Q_{100}); expected flood potential (Q_{efp}). (A) Big Creek – Middle Fork Eel River; (B) Homochitto River – Pemigewasset River; (C) South Branch Potomac River – Tonto Creek. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]



FIGURE 5 | (Continued).



FIGURE 5 | (Continued).



FIGURE 6 | Relative bias (%) in US Geological Survey (USGS) regional regression equations results in 18 test watersheds. Peak discharges at the 100-year (1% annual exceedance probability) scale are being underpredicted by regional regression equations (StreamStats) in most of the analyzed watersheds. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/raa.70991)]

The bottom-left quadrant of Figure 7 includes sites where streamgauge flood-frequency analyses are underpredicting flood magnitudes (in places by 70% to 80%); these sites typically have fewer annual peak discharge values in each record. These are streamgages with shorter records that do not include a larger flood magnitude than other methods (and neighboring streamgages) indicate as being expected in that area. However, some sites in both this quadrant and the bottom-right quadrant do have longer records, indicating complexity in streamgauge analysis results using the log-Pearson distribution, where both lower and higher B_i sites can induce what may be underestimated design peak discharges. The upper right quadrant, where bimodality is high and streamgauge analyses frequently appear to be overpredicting peak discharges (in places, dramatically so), streamgages predominantly have shorter record lengths but have experienced large or extreme floods compared to the records of neighboring streamgages. Other unknown mechanisms may also be at play.

USGS regional regression equations (as delivered through StreamStats) have been utilized for decades as a standard practice for the design of road-stream crossings and other applications at ungaged locations, with underprediction leading to undersized bridges and culverts that have insufficient capacity to safely convey expected flood magnitudes. If the relative bias assessments in these 18 watersheds are representative of the overall situation across the nation, undersized structures and insufficient floodplain capacity may be impairing flood resiliency in much of the contiguous United States. From an infrastructure perspective, addressing this issue as

structures exceed their effective life or fail can be expected to be expensive and take decades. Importantly, it is not suggested that regional regression equations not be used in analyses for peak flood discharge estimation; in some areas (such as the Little Miami River, Ohio and Broad River, Georgia watersheds) these estimates provide the best estimates for design and planning. In other areas (see the Cane Creek case study below), regional regression equations may underpredict for some watershed sizes but provide the best predictions in other watershed sizes. Potential low bias by any method for predicting design flood discharges needs to be assessed on a site-by-site basis, rather than a general deference to or disregard of any individual method.

The risk of incorporating bias into design and planning efforts can be minimized by utilizing multiple methods—a best practice is to utilize at least three methods for each site to select a more informed and resilient design flood discharge. A single method should not be preferred—the median result is commonly most appropriate. Such analyses can be quickly performed using the Flood Potential Portal (<https://floodpotential.erams.com/>), which can also be used to account for observed increasing trends in floods to adjust for nonstationarity due to climate variability, change, and other mechanisms (using a *monitor and adjust* philosophy). For example, for Rock Creek at Red Lodge, Montana (Figure 1), use of the watershed analysis module in the Flood Potential Portal yields results of $Q_{\text{efp}} = 214$ cms (7560 cfs); $Q_{100 \text{ index}} = 181$ cms (6400 cfs); and $Q_{100 \text{ regional regression}} = 86$ cms (3020 cfs), with an observed change in large flood magnitudes of +9.8%. From these results, the

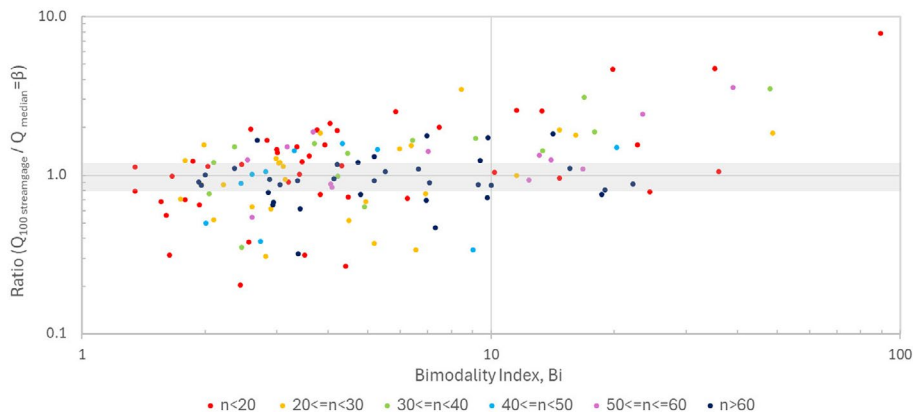


FIGURE 7 | Bias ratios (β) for 144 streamgage flood-frequency analysis results compared to the bimodality index, binned by record length (n). The gray area represents β values between 0.8 and 1.2, where $\beta=1$ indicates exact agreement between the streamgage analyses and median Flood Potential Portal Watershed Analysis module results. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

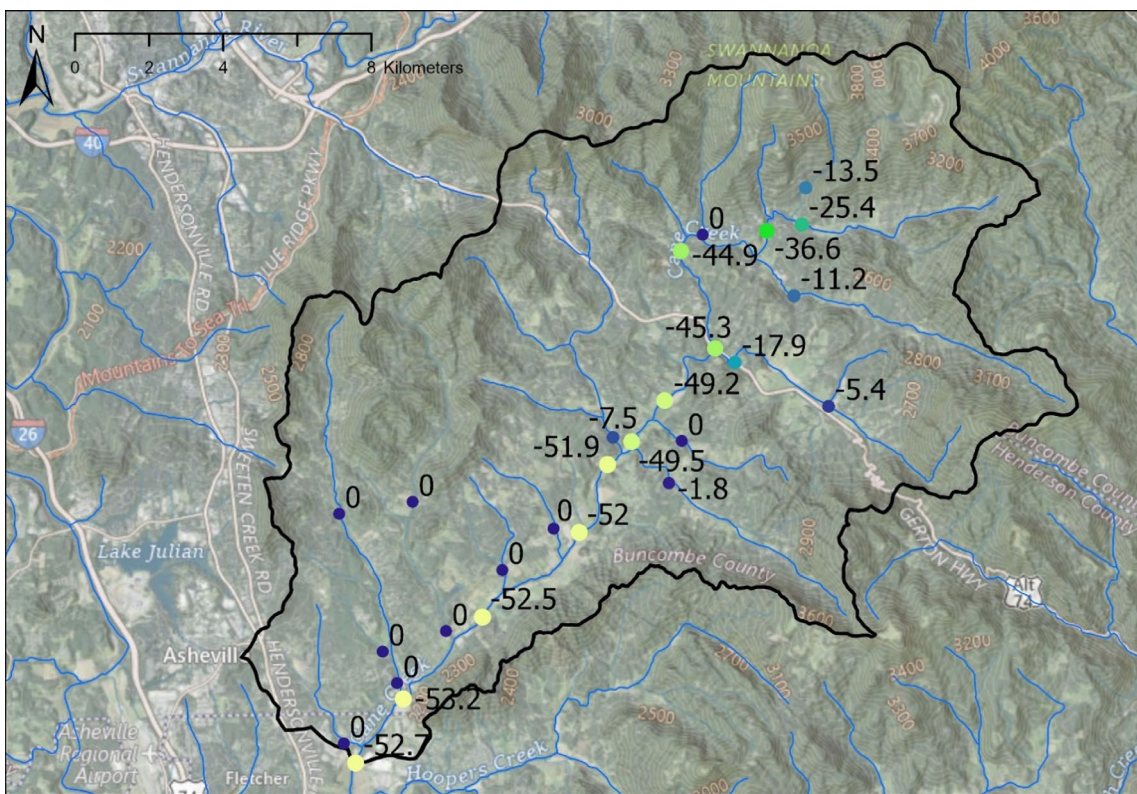


FIGURE 8 | Percent bias in USGS Q_{100} regional regression equations results at road-stream crossings in the Cane Creek Watershed near Asheville, North Carolina. The point labeled as “-53.2%” indicates a bridge over Cane Creek where, if regional regression equations results were used for sizing (without freeboard), the opening would have been undersized by more than half. Zero values indicate points (with smaller watersheds) where the regional regression equations’ results are the median values and most appropriate for design. Pre-Helene data and analyses were utilized for peak discharge predictions in the Flood Potential Portal. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

most appropriate design flood discharge is the median value adjusted for the observed increasing magnitudes, specifically $Q_{design} = 181 \text{ cms} \cdot 1.098 = 199 \text{ cms}$ (7030 cfs). The regional regression equation’s result is biased low, with a value only 42% of the recommended design value. The bridge opening shown in Figure 1 appears to be small; the utilized design flood discharge may be similar to the regional regression equation’s result, resulting in poor flood resilience when the next large flood occurs.

4.1 | Case Study: Cane Creek in Western North Carolina

Powerful remnants of Hurricane Helene induced catastrophic floods in the Southern Appalachians in September 2024, especially in Western North Carolina. Transportation infrastructure was damaged or destroyed, and floodplains inundated, resulting in 103 lives lost in North Carolina (NCDHHS 2024). Data and analyses indicate that this hurricane-induced flood event was

not unprecedented (Flood Potential Portal Mapping module, zone 73S; Amorim et al. 2025), with expansive floods experienced in 1940 and 1916. If USGS regional regression equations results were used to quantify peak flood discharges at ungaged locations for design, would this result in infrastructure being undersized? The 164 km² Cane Creek Watershed southeast of Asheville, NC (Figure 8) provides an example of transportation infrastructure vulnerability due to peak discharge prediction bias. Compared to the index flood frequency and Flood Potential results, regional regression equations substantially underpredicted peak discharges (Figure 8). Underprediction bias of up to −53.2% was quantified. However, relative bias varies by watershed size, with no or little bias in smaller watersheds and the greatest bias in the largest watersheds—regional regression equations results are most appropriate (the median result) for design in many of the smallest watersheds, while flood potential or index results are most appropriate for larger watersheds. If regional regression equations were exclusively used to size road-stream crossings, many primary bridges and culverts would be undersized and susceptible to damage and failure.

5 | Summary and Conclusions

Using a dataset of peak discharge estimates at 1517 points collected as a census of road-stream crossings in 18 sampled watersheds distributed across the United States, relative bias was assessed in the three methods used for prediction in the Flood Potential Portal (<https://floodpotential.erams.com/>) Watershed Analysis module. These methods are the flood potential, index flood frequency, and USGS regional regression equations flood frequency, with the median result at each point assumed to be most appropriate for use in design. Overall, variation from the median value for each site indicates significantly different results between regional regression equations and both index flood and expected flood potential, but insignificant differences between index flood and expected flood potential. Fifteen of the 18 watersheds indicated significant regional regression equations underprediction, with an average percent bias of −27.8% (ranging from −55.1% to −1.8%). Use of only this method for design would result in bridge and culvert openings being undersized by up to about half. If these sampled watersheds are representative of the overall contiguous United States, use of USGS regional regression equations at ungaged sites for design and planning may have inadvertently resulted in poor flood resilience across much of the United States. However, these 18 watersheds are relatively few compared to the vastness of the contiguous United States; additional analyses in many more watersheds, in the other 188 (of 206) flood potential zones, are needed to better understand bias. Despite this limitation, it is clear that a best practice in addressing bias in any technique for peak discharge prediction at ungaged locations is to use multiple (at least three) methods for prediction, with the median value being most reasonable in many situations for use (to avoid both underestimation and overestimation). Some settings may justify the use of a higher value. To ensure due diligence during design and planning activities, use of this best practice is recommended.

A case study of a watershed in Western North Carolina impacted by remnants of Hurricane Helene was provided to illustrate the vulnerability of transportation infrastructure to large floods. Peak flood discharges computed using regional regression equations,

while appearing to perform without bias for numerous small watersheds (compared to this suite of prediction methods), underpredict peak flood magnitudes by half for primary streams and larger watersheds—if this method was used for sizing bridges and culverts in this watershed, flood openings in many structures would be half the size needed to safely convey large floods.

Additional work is required to enhance the knowledge and data available for the field of flood hydrology. Specifically, research is needed to quantify how relative bias varies at much higher resolution than this preliminary analysis using 18 watersheds, and to better elucidate mechanisms inducing bias in streamgage flood-frequency analyses and the impacts of such bias on peak discharge estimation at ungaged locations. Bias at return intervals other than the 100-year scale is also of interest. Dependence on a single statistical distribution in the United States, log-Pearson, may need to be reconsidered—consistent results should be a priority over consistent methods. Additional research into the physical basis of large flood production, such as spatial precipitation variability, the influence of antecedent soil moisture, the importance of saturation-excess overland flow, and other physical processes, is additionally needed to increase the reliability of rainfall-runoff and downscaled climate modeling.

Acknowledgments

Review of this manuscript by David Levinson and three anonymous reviewers has been greatly appreciated; their suggestions improved the article. Mazdak Arabi provided valuable insight on this research and the article presentation, increasing the impact of this contribution. Additional staff at the One Water Solutions Institute at Colorado State University were integral to the development of the Flood Potential Portal, specifically Matthew Korsak, Mahshid Ghanbari, Lucas Yaege, Dave Patterson, and Holm Kipka. David Levinson, Heidi Klingel, Joel Murray, and Julian Scott, of the U.S. Forest Service National Stream and Aquatic Ecology Center, are appreciated for their contributions to the development of the flood potential method and its implementation across the contiguous United States. Additionally, appreciation is also expressed to the many hydrologists, hydraulic engineers, and technicians with the US Geological Survey and other agencies who have collected and analyzed peak stage and discharge data, as well as the StreamStats staff for the regional regression equations results presented within the Flood Potential Portal.

Funding

This research was supported through the US Forest Service National Stream and Aquatic Ecology Center by two challenge cost share agreements (20-CS-11132422-274), with the second agreement funded in part by a grant provided by the Federal Highway Administration Innovation and Research Council (24-CS-11132422-254).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in the Flood Potential Project Page at <https://y-water.com/floodpotential/>.

References

Amorim, R., G. Villarini, J. Czajkowski, and J. Smith. 2025. “Flooding From Hurricane Helene and Associated Impacts: A Historical

- Perspective.” *Journal of Hydrology X* 27: 100204. <https://doi.org/10.1016/j.hydroa.2025.100204>.
- Capesius, J. P., and V. C. Stephens. 2009. Regional Regression Equations for Estimation of Natural Streamflow Statistics in Colorado. U.S. Geological Survey Scientific Investigations Report 2009–5136.
- Dalrymple, T. 1960. Flood-Frequency Analyses. Manual of Hydrology Part 3. Flood-Flow techniques. Geological Survey Water-Supply Paper, 1543-A.
- England, J. F., T. C. Cohn, B. A. Faber, et al. 2018. Guidelines for Determining Flood Flow Frequency—Bulletin 17C. U.S. Geological Survey Techniques and Methods, Book 4, Chap. B5, 148 p.
- Francois, B., K. E. Schlef, S. Wi, and C. M. Brown. 2019. “Design Considerations for Riverine Floods in a Changing Climate: A Review.” *Journal of Hydrology* 574: 557–573. <https://doi.org/10.1016/j.jhydrol.2019.04.068>.
- Galloway, G. E. 2011. “If Stationarity Is Dead, What Do We Do Now?” *Journal of the American Water Resources Association* 47, no. 3: 563–570.
- Han, Z., and H. O. Sharif. 2021. “Analysis of Flood Fatalities in the United States, 1959–2019.” *Water* 13: 1871. <https://doi.org/10.3390/w13131871>.
- Hosking, J. R. M. 1990. “L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics.” *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 52, no. 1: 105–124.
- Hosking, J. R. M., and J. R. Wallis. 1993. “Some Statistics Useful in Regional Frequency Analysis.” *Water Resources Research* 29, no. 92: 271–281.
- Interagency Advisory Committee on Water Data. 1982. *Guidelines for Determining Flood Flow Frequency: Bulletin #17B for the Hydrology Subcommittee*. U.S. Department of the Interior, Geological Survey, Office of Water Data Coordination.
- Jones, A., J. Kuehnert, P. Fraccaro, et al. 2023. “AI for Climate Impacts: Applications in Flood Risk.” *npj Climate and Atmospheric Science* 6: 63. <https://doi.org/10.1038/s41612-023-00388-1>.
- Keller, A. A., K. L. Garner, N. Rao, E. Knipping, and J. Thomas. 2022. “Downscaling Approaches of Climate Change Projections for Watershed Modeling: Review of Theoretical and Practical Considerations.” *PLOS Water* 1, no. 9: e0000046. <https://doi.org/10.1371/journal.pwat.0000046>.
- Kenney, T. A., C. D. Wilkowske, and S. J. Wright. 2007. Methods for Estimating Magnitude and Frequency of Peak Flows for Natural Streams in Utah. U.S Geological Survey Scientific Investigations Report 2007–5158.
- Kohn, M. S., M. R. Stevens, T. M. Harden, J. E. Godaire, R. E. Klinger, and A. Mommandi. 2016. Paleoflood Investigations to Improve Peak-Streamflow Regional-Regression Equations for Natural Streamflow in Eastern Colorado, 2015. U.S. Geological Survey Scientific Investigations Report 2016–5099.
- Kumar, V., K. V. Sharma, T. Caloiero, D. J. Mehta, and K. Singh. 2023. “Comprehensive Overview of Flood Modeling Approaches: A Review of Recent Advances.” *Hydrology* 10: 141. <https://doi.org/10.3390/hydrology10070141>.
- Li, Z., S. Gao, M. Chen, J. J. Gourley, and Y. Hong. 2022. “Spatiotemporal Characteristics of US Floods: Current Status and Forecast Under a Future Warmer Climate.” *Earth’s Future* 10: 1–9. <https://doi.org/10.1029/2022EF002700>.
- Lins, H. F., and T. A. Cohn. 2011. “Stationarity: Wanted Dead or Alive?” *Journal of the American Water Resources Association* 47, no. 3: 475–480.
- Merz, B., S. Basso, S. Fischer, et al. 2022. “Understanding Heavy Tails of Flood Peak Distributions.” *Water Resources Research* 58: 1–37. <https://doi.org/10.1029/2021WR030506>.
- Miller, K. A. 2003. Peak-Flow Characteristics of Wyoming Streams. U.S Geological Survey Water-Resources Investigations Report 03–4107.
- Mukherjee, S., D. M. Amatya, J. L. Campbell, et al. 2024. “A Watershed-Scale Multi-Approach Assessment of Design Flood Discharge Estimates Used in Hydrologic Risk Analyses for Forest Road Stream Crossings and Culverts.” *Journal of Hydrology* 632: 130698. <https://doi.org/10.1016/j.jhydrol.2024.130698>.
- NCDHHS. 2024. *Hurricane Helene Storm Related Fatalities*. North Carolina Department of Health and Human Services. <https://www.ncdhhs.gov/assistance/hurricane-helene-recovery-resources/hurricane-helene-storm-related-fatalities>.
- NOAA (National Oceanic and Atmospheric Administration, National Water Prediction Service). 2025. NOAA Atlas 15: Update to the National Precipitation Frequency Standard. <https://water.noaa.gov/about/atlas15>.
- Ries, K. G., P. A. Steeves, and P. McCarthy. 2024. “StreamStats—A Quarter Century of Delivering Web-Based Geospatial and Hydrologic Information to the Public, and Lessons Learned: U.S.” *Geological Survey Circular* 1514: 40. <https://doi.org/10.3133/cir1514>.
- USGCRP. 2023. *Fifth National Climate Assessment*, edited by A. R. Crimmins, C. W. Avery, D. R. Easterling, K. E. Kunkel, B. C. Stewart, and T. K. Maycock. U.S. Global Change Research Program. <https://doi.org/10.7930/NCA5.2023>.
- Waltemeyer, S. D. 2008. Analysis of the Magnitude and Frequency of Peak Discharge and Maximum Observed Peak Discharge in New Mexico and Surrounding Areas. U.S Geological Survey Scientific Investigations Report 2008–5119.
- Yochum, S. E., J. A. Scott, and D. H. Levinson. 2019. “Methods for Assessing Expected Flood Potential and Variability: Southern Rocky Mountains Region.” *Water Resources Research* 55: 6392–6416. <https://doi.org/10.1029/2018WR024604>.
- Yochum, S. E., T. Wible, M. Korsas, M. Ghanbari, and M. Arabi. 2024. “Flood Potential Portal: A Web Tool for Understanding Flood Variability and Predicting Peak Discharges.” *River Research and Applications* 40: 1855–1867. <https://doi.org/10.1002/rra.4354>.