Received: 28 February 2024 | Revised: 11 July 2024 | Accepted: 12 July 2024
DOI: 10.1002/rra.4354

RESEARCH ARTICLE

WILEY

Flood Potential Portal: A web tool for understanding flood variability and predicting peak discharges

Steven E. Yochum[1] | **Tyler Wible**[2] | **Matthew Korsa**[2] | **Mahshid Ghanbari**[2] | **Mazdak Arabi**[2]

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

[2]Colorado State University, Civil and Environmental Engineering, One Water Solutions Institute, Fort Collins, Colorado, USA

Correspondence
Steven E. Yochum, U.S. Forest Service, National Stream and Aquatic Ecology Center, Fort Collins, CO, USA
Email: steven.yochum@usda.gov

Abstract

The Flood Potential Portal (https://floodpotential.erams.com/) has been developed for the contiguous United States, as a practitioner-focused tool that uses observational data (streamgages) to enhance understanding of how floods vary in space and time, and assist users in making more informed peak discharge predictions for infrastructure design and floodplain management. This capability is presented through several modules. The Mapping module provides tools to explore variability using multiple indices, and provides detailed information, figures, and algorithms describing and comparing flooding characteristics. The Cross-Section Analysis module allows users to cut regional-scale sections to interpret the role of topography in driving flood variability. The Watershed Analysis module provides multiple methods for quantifying expected peak discharge magnitudes and flood frequency relationships at user-selected locations, including the integration of observed trends in flood magnitudes due to climate change and other sources of nonstationarity into decision making. The Streamgage Analysis module performs streamgage flood-frequency analyses. These modules are based in part on the flood potential method, through the use of 207 zones of similar flood response defined using more than 8200 streamgages with watershed areas <10,000 km^2. Regression models that define each zone had high explained variance (average $R^2 = 0.93$). An example is provided to illustrate use of the Flood Potential Portal for the design of a hypothetical bridge replacement.

KEYWORDS
bridge, climate change, culvert, design, extreme, floods, hydrology

1 | INTRODUCTION

Riverine floods are a leading environmental threat to life, infrastructure, and property. Climate change is likely shifting these threats in some areas, as one of the mechanisms inducing nonstationarity (Francois et al., 2019; Galloway, 2011; Lins & Cohn, 2011). To maximize infrastructure resilience and knowledge used for stream and riparian ecosystem management, it is essential to develop enhanced understanding of riverine flood hazards and make this knowledge readily available to hydrologic professionals, to be incorporated into decision-making. The observational (streamgage) network is fundamental for this resilience. Streamgaging data are highly informative for understanding past events and detecting current trends, to quantify the status of riverine flooding and how these events may be changing.

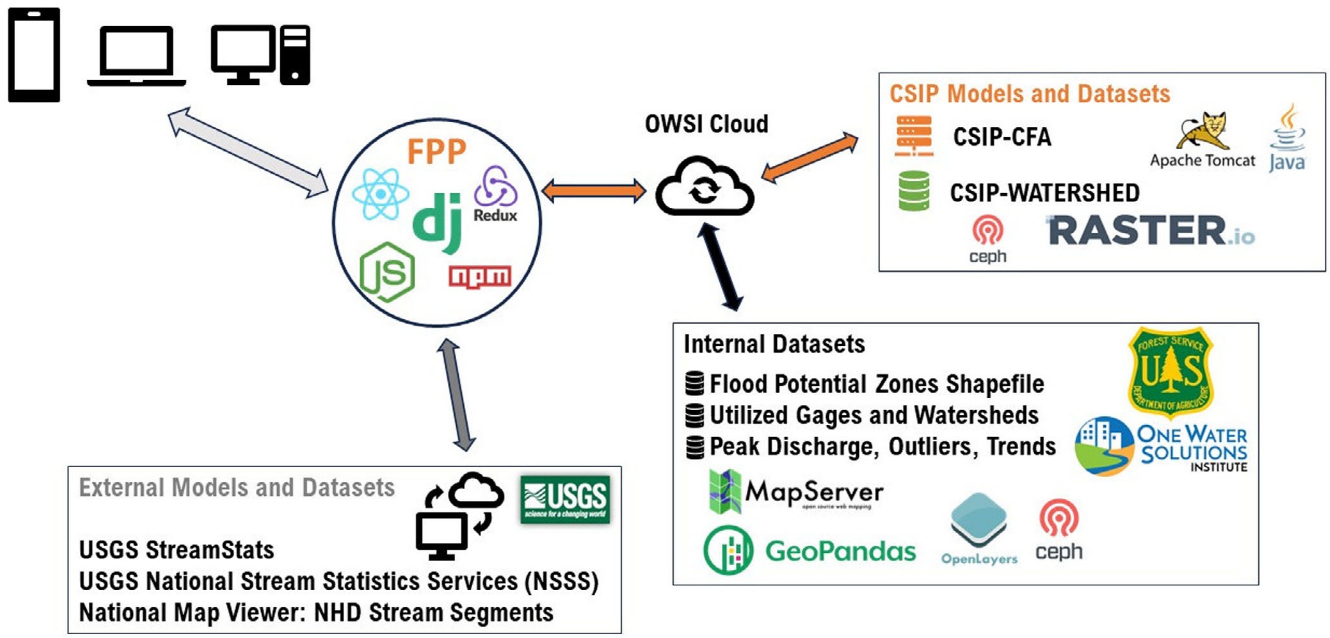

FIGURE 1 Flowchart describing the architecture of the Flood Potential Portal. This includes the relationships between internal and external datasets and modeling services, with logos of key architecture and processing software dependencies for internal components. [Color figure can be viewed at wileyonlinelibrary.com]

The Flood Potential Portal (FPP; https://floodpotential.erams.com/) is a web-based geospatial decision support system for enhancing the understanding of riverine flood hazards in the United States. The FPP reconciles more than a century of stream discharge monitoring efforts using both traditional and new flood analysis techniques, to assist practitioners with enhanced understanding of flood variability and expected peak discharge magnitudes. The tool helps users understand the observed spatial and temporal variability of floods and quantify flood magnitudes for determining design flood and base flood discharges. The FPP assists professionals with understanding how floods vary in space and time (from continental to catchment scales), exploring how floods differ across regions, and predicting flood magnitudes at points of interest using multiple methodologies, as a form of ensemble decision-making.

The stationarity assumption for streamgage analyses is an important consideration when using common flood-frequency techniques. The FPP tests for nonstationarity of observed flood time series due to climate change, reservoir storage, and other mechanisms, and provides applicable percentage change adjustments where trends in flood magnitudes are detected. The FPP also provides capabilities for testing trends in flood frequency and flashiness. This suite of analyses within the FPP serves to test and account for observed changes in flooding. The FPP combines the results of the flood potential method (Yochum et al., 2019), an approach for predicting, comparing, and communicating expected flood magnitudes and variability, with traditional flood-frequency analysis methods for predictions. The use of ensemble results provides redundancies that mitigate shortcomings or limitations present within each approach.

The FPP is deployed as a web analysis tool using the Catena Analytics software (Figure 1). Catena Analytics offers powerful cloud-based capabilities for building computationally scalable and platform-independent tools that can be accessed through web browsers on desktop or mobile devices. The software supports content management, geospatial mapping and processing, data analysis and modeling services, and applications. Services in Catena Analytics incorporate the Cloud Services Implementation Platform (David et al., 2014; Lloyd et al., 2013, 2015) RESTful web services for datasets and analyses.

Features presented within the FPP include:

- Mapping tools to explore characteristics of floods, including scale of floods experienced within a zone of interest, and in comparison to other zones; flashiness and bimodality of the experienced floods; flood variability; and dominant and secondary flood seasonality
- Mapping and time series analyses reporting trends in flood magnitudes, frequency, and flashiness
- Tabular and graphical tools to assess flooding characteristics within each flood potential zone, and quantifiably compare characteristics between zones
- Cross-section analysis, to understand the role of topography in flood variability across regions
- Flood magnitude prediction at ungaged locations using multiple techniques, including the flood potential method
- Streamgage flood-frequency analysis

This article introduces the FPP (version 2.1) to practitioners and researchers, documenting key features and underlying methodologies, and provides an example. Use of this tool is expected to enhance resilience of stream and floodplain infrastructure to floods, and aid in the management of stream and riparian systems. Due to the complex

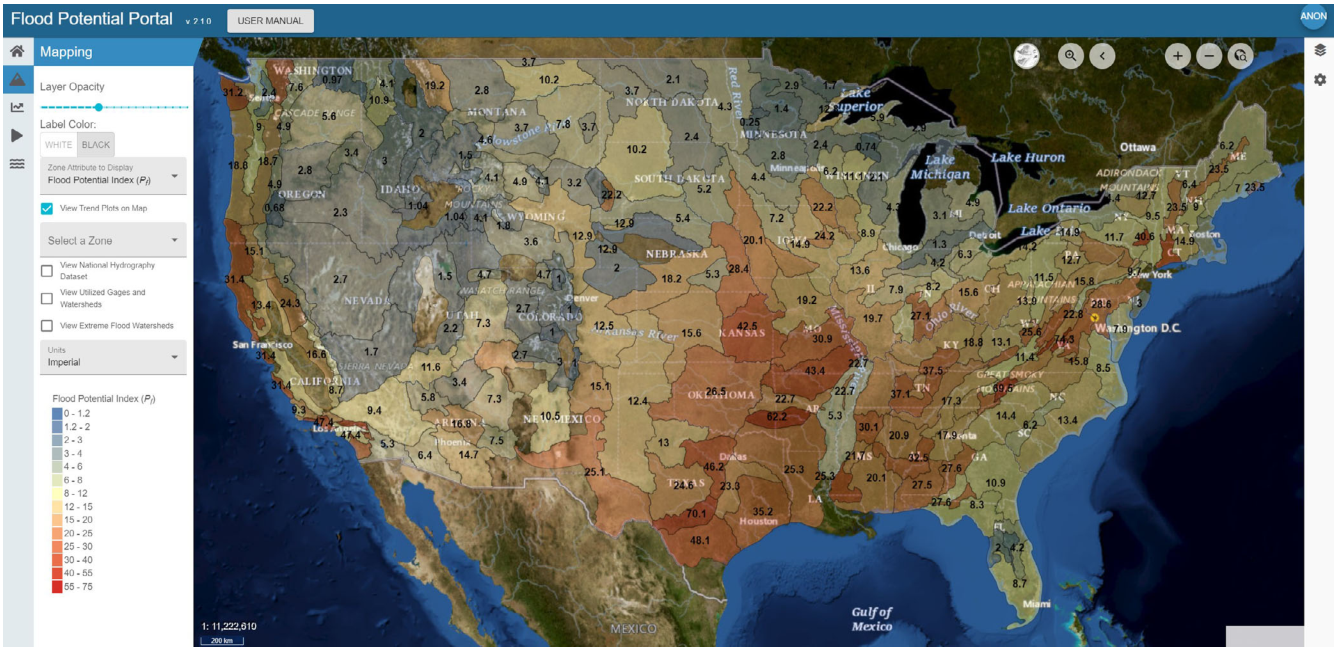

FIGURE 2 Zones and flood potential index (P_f) values across the contiguous United States, as viewed within the Flood Potential Portal. Developed using 8233 streamgaged watersheds across 207 zones. [Color figure can be viewed at wileyonlinelibrary.com]

considerations involved in interpretations regarding expected flood magnitudes, professional expertise in the hydrologic sciences and flood hydrology is required for the proper use of results.

2 | FLOOD POTENTIAL PORTAL

The FPP (https://floodpotential.erams.com/) assists with visualizing, quantifying, and communicating about flood hazards, using the flood potential method and traditional flood-frequency analyses. The flood potential method has been presented in Yochum et al. (2019), Yochum (2019), and Yochum and Levinson (2023). The User Manual (Yochum et al., 2024) contains detailed information on the use of the decision support system, the flood potential method, and example applications. This tool has been developed as a resource for hydrologists, civil engineers, and floodplain managers for understanding how floods vary in space and time, and for determining peak magnitudes using multiple methods that utilize data collected by the nation's streamgaging network. General information on the FPP is provided, as well as selected details that support the analytical methods.

The FPP provides access to flood information and assessments for 207 flood potential zones across the contiguous United States (Figure 2). The zones, ranging in area from 3100 to 202,000 km^2, were developed by clustering 8233 streamgaged watersheds (with watershed areas <10,000 km^2) to obtain relatively consistent flood characteristics within each zone. These zones are analogous to the regions used in USGS regional regression analyses. Regression models defining each zone typically had high explained variance, with an average $R^2 = 0.93$. The technique used to develop these zones is presented in Yochum et al. (2019) and the user manual (Yochum et al., 2024).

The FPP is a web tool that can be accessed by commonly used browsers (e.g., Microsoft Edge, Google Chrome, and Mozilla Firefox). A token is generated for each unique session; this token can be used as a custom URL to return to previous work or share with collaborators. A variety of base map layers are available as background imagery, for varying preferences by user, physical setting, and objective. Additional user layers, tables, and geoprocessing tools can be uploaded by using the mapping tools in the mapping canvas. Access to numerous public datasets is provided, and users can also upload their own spatial datasets. Features of the FPP are deployed in five modules: Overview, Mapping, Cross-Section Analysis, Watershed Analysis, and Streamgage Analysis.

2.1 | Overview

The Overview module describes the tool and provides links to the user manual and quick start guide. The user manual is also available through a link at the top of the page, in all modules; this extensive reference should be regularly consulted by users.

2.2 | Mapping

The Mapping module includes tools for users to explore flood variability across the United States as quantified using the flood potential method, at a full range of scales. This module provides options for users to map zone attributes (Table 1) and overlay supplemental layers (National Hydrography Dataset, zone gages and watersheds, watersheds that have experienced extreme floods). Clicking on any zone provides detailed information and figures describing zone flooding

TABLE 1 Zone attributes available for mapping within the Flood Potential Portal.

Description
Flood potential index (P_f)
Zone ID
Zone name
Watershed scale ratio (R_f)
Beard flashiness index (*F*)
Richards-Baker flashiness index (*R-B*)
Bimodality index (B_i)
Flood variability index (V_f)
Flood hazard index (H_f)
Explained variance (R^2)
Dominant flooding month
Secondary flooding month
Trends in largest (5%) flood magnitudes
Q4: Trends in largest quarter magnitudes (>4 year RI)
Q3: Trends in moderate quarter magnitudes (4 to 2 years RI)
Q2: Trends in ~Bankfull magnitudes (2 to 1.33 years RI)
Q1: Trends in < Bankfull magnitudes (<1.33 year RI)
Trends in annual flood frequency
Trends in event flood frequency
Trends in flashiness (*R-B*)

Note: Attributes are described in the user manual.
Abbreviation: RI: Recurrence interval.

characteristics, prediction algorithms, and trend analysis results to assess stationarity, analysis periods, as well as flood potential, seasonality, and regional comparison plots.

Many practitioners will focus their use on the Watershed Analysis and Streamgage Analysis modules, however, it is important to initially explore the status and variability of floods in the area of interest. Professionals should have a general understanding of floods that have been experienced before they make design flood discharge predictions. The Mapping and Cross-Section Analysis modules provide such insight.

Flood variability across the contiguous United States, as quantified using the flood potential index (P_f), is the default zone attribute initiated when activating the Mapping module. Zooming in to a region of interest (Figure 3) illuminates the inherent spatial variably of peak discharge magnitudes as measured by the streamgaging network. In this example, the variability within the Appalachians Mountains, Piedmont, and Coastal Plain of the Virginias and Carolinas are illustrated. P_f allow users to quantify flood variability between areas of interest. Flood potential varies substantially within this extent, from $P_f = 74.3$ (zone 73F: Appalachian Front in vicinity of Charlottesville, Virginia) to $P_f = 6.2$ (zone 75 W: Carolina Sandhills, at the boundary between the Atlantic Coastal Plain and the Piedmont). This means that, on average, flood magnitudes are 74.3/6.2 = 12.0 times larger in zone 73F compared to zone 75W.

Flood potential is generally reduced on the Atlantic coastal plain (P_f: 8.5, 13.4), though flood magnitudes are, on average, 60% higher along the Carolina coast than in Virginia (13.4/8.5 = 1.6). Flood potential is higher on the piedmont than the coastal plain, and then dramatically increases in portions of the Appalachian front (zone 73F; zone 73S: Grandfather Mountain area, northeast of Ashville, NC, $P_f = 69.5$). Further to the west, flood potential decreases substantially in the central portion of the Appalachian Mountains (valley and ridge, zone = 72; $P_f = 11.4$). Large floods in zone 72 are, on average, only 15% of the magnitude of floods in zone 73F (11.4/74.3 = 0.15). Further to the west in the eastern portion of the Appalachian plateaus (zone 71; $P_f = 25.6$), flood potential increases compared to the valley and ridge, with floods in zone 71 being, on average, 2.2 times larger than in zone 72 (25.6/11.4 = 2.23).

Other attributes provided by the FPP of most value for understanding how floods vary in space include the watershed scale ratio, Beard flashiness index (*F*), Richards-Baker flashiness index (*R-B*), bimodality index (B_i), flood variability index (V_f), flood hazard index (H_f), and large flood seasonality (dominant flooding month and the secondary flooding month, of the largest 5% of annual peak discharges). These attributes are described in the user manual (Yochum et al., 2024).

Extreme floods are events beyond the peak discharge magnitude expected given the streamgage records within each zone; these floods have been systematically identified and ranked. Watersheds for these events are illustrated in the FPP, with on-hover details ("View Extreme Flood Watersheds" toggle in the left panel). These extreme floods have been quantified for streamgages with at least 10 years of record, though many other extreme floods have occurred that do not have such records (or any records). These extreme events are relative to the zone that their watersheds fall within, and are identified as having experienced discharges greater than the upper 90% prediction limit of the expected flood potential regressions, which define the maximum likely flood potential discharge ($Q > Q_{mlf}$). Ranking of extreme floods is performed using the flood extreme index (E_f), where higher values (and warmer colors) indicate greater extremity. The E_f value is a multiplier of the expected flood potential discharge (Q_{efp}), the central tendency of large flood magnitudes—a watershed that has experienced an $E_f = 5.0$ has experienced a flood 5 times the expected discharge magnitude. These extreme floods are defined relative to the events quantified as being typical in each zone, and are not reliant on flood frequency (Smith et al., 2018) or hydrologic modeling (Tarouilly et al., 2021) approaches for identifying and ranking unusually large floods.

Eight trend tests in flood magnitudes, frequency, and flashiness are utilized to understand how floods vary over time within each zone, from nonstationarity (Francois et al., 2019; Galloway, 2011; Lins & Cohn, 2011; Milly et al., 2008). The results of these trend tests are presented as mapping attributes (Table 1), including on-hover figures showing significant, possible, or undetected trends. The philosophical approach toward nonstationarity implemented in this project is to *monitor and adjust*, utilizing the streamgaging record to quantify how floods are changing for practitioners to make informed decisions,

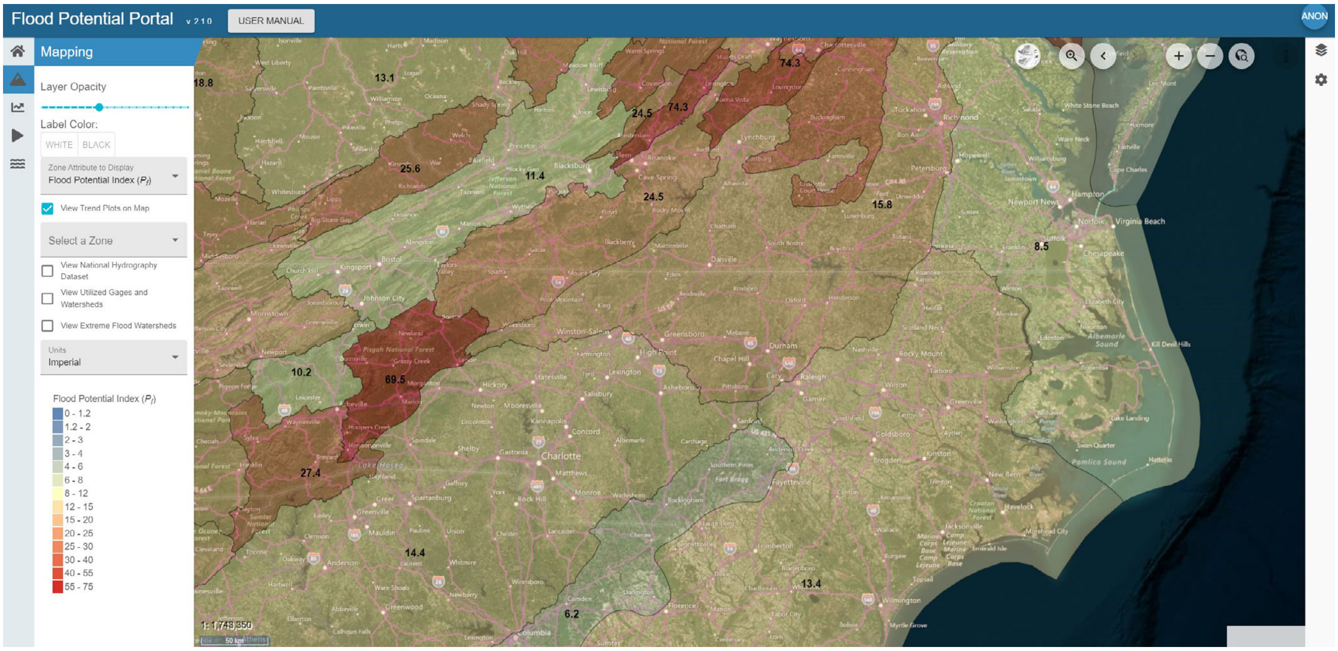

FIGURE 3 Flood Potential Portal mapping page showing a region consisting of portions of Virginia, West Virginia, North Carolina, and South Carolina. Labels indicate flood potential index (P_f) values, with warmer colors indicating higher flood potential and cooler colors indicating lower flood potential. [Color figure can be viewed at wileyonlinelibrary.com]

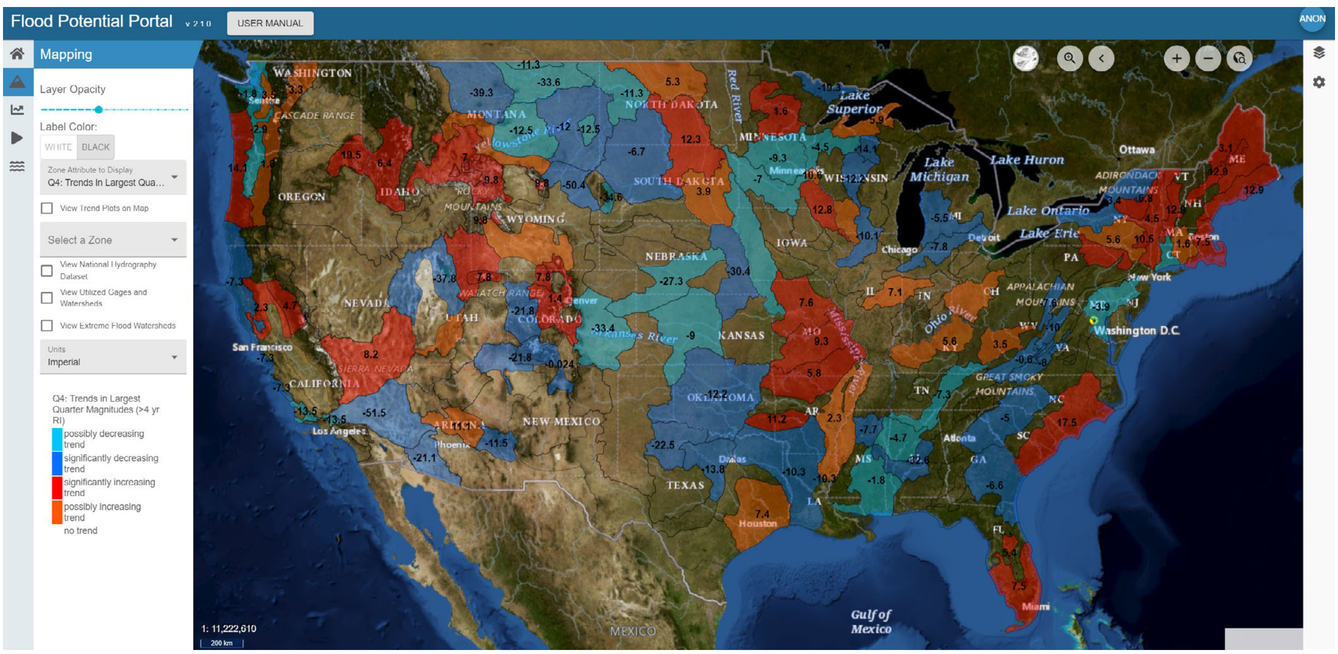

FIGURE 4 Significant and possible trends in flood magnitudes of the largest quarter (Q4) of annual peak discharges for streamgages within each flood potential zone, as viewed from within the Flood Potential Portal. Reds indicate observed increasing trends while blues indicate decreasing trends. Regions with the most consistent increasing trends in large floods include New England, the Interior Highlands, and the central portion of the Rocky Mountains. [Color figure can be viewed at wileyonlinelibrary.com]

to maximize resilience. Analyses are performed using all utilized streamgages within each zone, to maximize statistical power. A trend with a p-value ≤0.05 is considered significant, and ≤0.15 is considered possible. Trend analysis periods vary by zone and test; analysis periods are presented in the zone summary information. Trend tests (and flood potential analyses) are updated periodically, prioritizing zones that have experienced recent extensive flooding. A summary of the methods used for trend detection are provided below, with additional details provided in Yochum et al. (2019) and the user manual (Yochum et al., 2024).

Trends in magnitudes consist of tests of the largest 5% and four quarters of the annual peak discharge record, for the entire data extent in each zone. Results from trend tests of the largest quarter (Q4; >4-year return interval) are provided as an example (Figure 4). Flood magnitudes are normalized using the flood extreme index (E_f). Where trends in flood magnitudes have been detected, the percent change in E_f values are computed by comparison of the averages of the most recent 30 years of record with the entire record. Percent change is labeled and can be used to adjust magnitudes predicted in the Watershed Analysis and Streamgage Analysis modules. Trends in the frequency of floods are evaluated for 1945 to present using both annual and daily data, with analyses that counted zonal floods in each year that exceeded a threshold based on a fraction of the expected flood potential discharge at each streamgage. Variabilities in streamgaging periods and record lengths are addressed by dividing the annual counts by the number of streamgages operating during each year. Trends in slope are provided as labels for relative comparison. Trends in flashiness are computed using the *R-B* flashiness index (Baker et al., 2004).

2.3 | Cross-section analysis

The cross section analysis module allows users to cut regional-scale sections across multiple flood potential zones to inspect and interpret the role of topography in driving orographic forcing of precipitation, which generally enhances flood potential on windward slopes and reduces flood potential on leeward slopes. Cross sections can be cut along transects, as well as along rivers, transportation networks, and so forth. An example is provided (Figure 5) for Northern California through western Nevada. Coloring from the flood potential index (P_f) is shown overlaid on the cross-section.

Atmospheric river events during the winter are the primary driver of large floods in Northern California (Dettinger et al., 2011; Guirguis et al., 2019; Ralph et al., 2016). As lower-level water vapor (<~2.5 km; Dettinger et al., 2011) is pushed upslope after moving onshore from the Pacific Ocean into the California's Coastal Range, orographic uplift occurs and the highest flood potential is experienced ($P_f = 31.2$). As an atmospheric river penetrates the leeward side of the coastal range, and across the Sacramento Valley, substantially lesser flood potential exists ($P_f = 13.4$), with large flood magnitudes $31.2/13.4 = 2.3$ times larger, on average, in zone 25 on windward side than in zone 24 on the leeward side. As an atmospheric river penetrates further inland, into the Sierra Nevada Mountains, uplift again occurs and flood potential increases ($P_f = 24.3$), before large flood magnitudes decrease by 4/5th ($5.0/24.3 = 0.21$) in the high elevation eastern portion of the northern Sierra Nevada in a transition zone ($P_f = 5.0$), and reduced by another half ($2.7/5.0 = 0.54$) in the Great Basin ($P_f = 2.7$).

2.4 | Watershed analysis

Practitioners need to quantify peak discharge at ungaged locations to design road-stream crossings (bridges and culverts) and other infrastructure. A standard practice is to use the US Geological Survey (USGS) StreamStats application (https://streamstats.usgs.gov/ss/) for this purpose. This tool is powerful for providing a wide variety of streamflow characteristics, however it depends upon only one method for the selection of design flood discharges: regional regression equations. A better practice is to utilize an ensemble for the selection. The FPP Watershed Analysis module includes three approaches for prediction at user-selected points: (a) the flood potential method, including trend reporting of flood magnitudes (with percent change), flood frequency, and flashiness; (b) an index flood-frequency method, using flood potential zones; and (c) USGS regional regression flood-frequency equations, as presented in StreamStats (Figure 6). The expected flood potential (Q_{efp}), 100-year index ($Q_{100,index}$, 1% chance of exceedance), and the 100-year regional regression ($Q_{100,regional\ regression}$) discharges are directly comparable for the selection of design flood discharges. Percent change from observed increasing trends in flood magnitudes can be used to adjust these values, at the discretion of the practitioner.

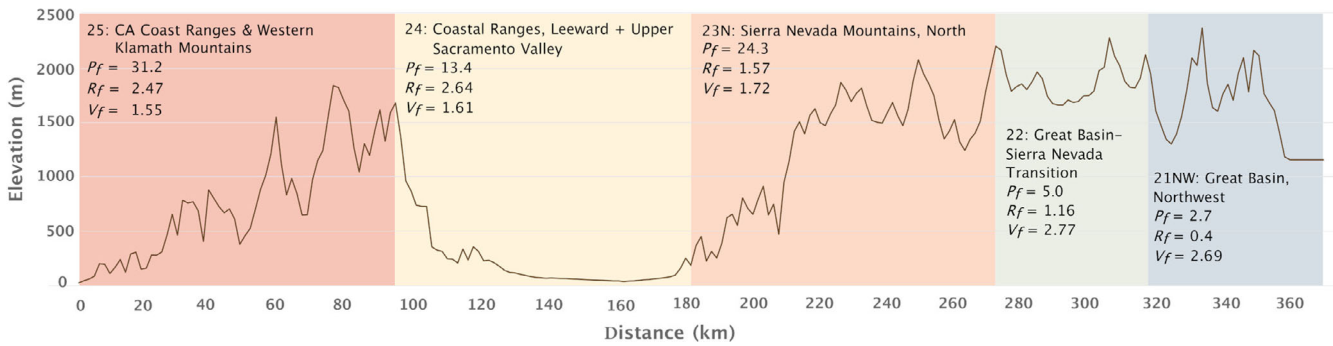

FIGURE 5 Exported regional-scale cross section showing ground surface and flood potential variability across Northern California into Western Nevada., from the City of Fort Bragg on the Pacific Coast, through Chico, CA in the Sacramento Valley, to Pyramid Lake, NV in the Great Basin (50 km north of Reno). Warmer colors and higher flood potential index (P_f) values indicate higher flood potential. R_f = watershed scale ratio; V_f = flood variability index. Generated using 2 km elevation interval. [Color figure can be viewed at wileyonlinelibrary.com]

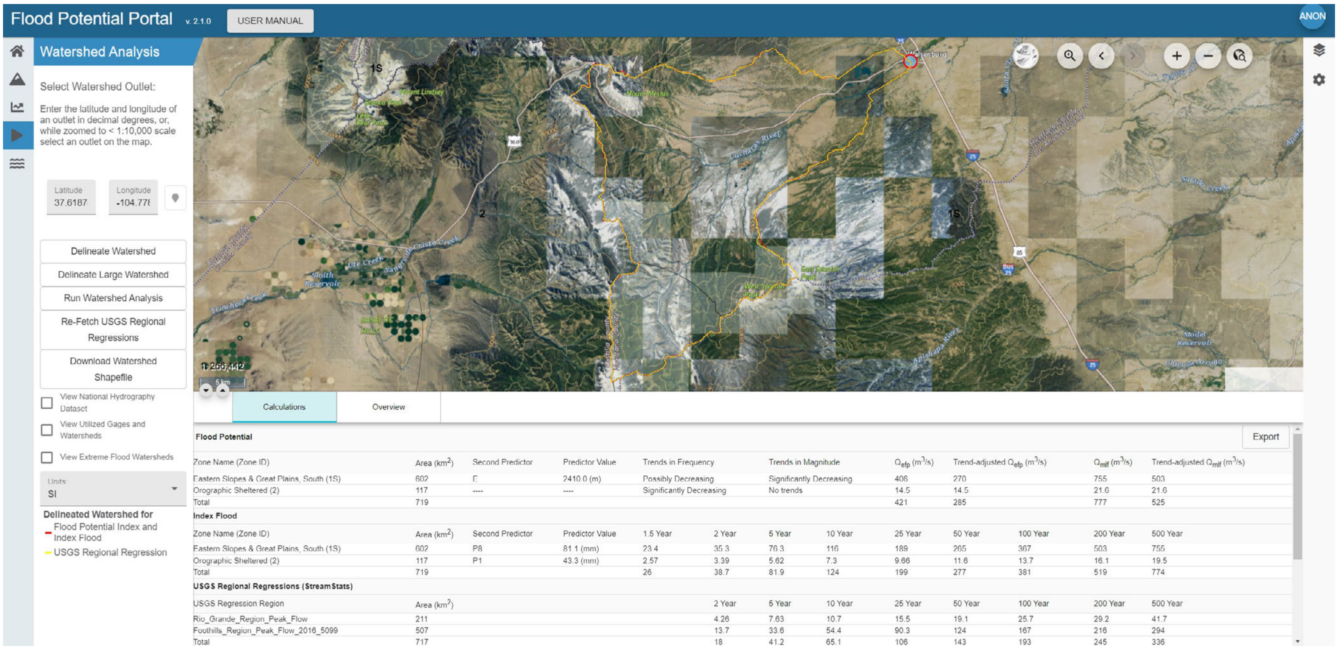

FIGURE 6 Watershed analyses module results (Calculations tab) within the Flood Potential Portal for the Cucharas River at Walsenburg, Colorado. Results: $Q_{efp} = 421$ cms (with no trends or decreasing trends in large flood magnitudes and frequency); $Q_{100\ index} = 381$ cms; $Q_{100\ regional\ regression} = 193$ cms. Selected design flood discharge (median) = 381 cms (13,400 cfs). [Color figure can be viewed at wileyonlinelibrary.com]

The flood potential method provides expected flood potential and maximum likely flood potential (Q_{mlf}) discharges (Yochum et al., 2019; Yochum & Levinson, 2023). Q_{efp} is computed from a regression of the record peak discharges for the streamgaged watersheds within each zone, with the Q_{mlf} being the upper 90% prediction limit, floods beyond this being extreme, and departure indicating the degree of extremity. Peak discharge predictions utilizing Q_{efp} predict flood magnitudes based on the central tendency of large floods experienced within zones of similar flood response. Considering the streamgaging records, and the meteorological and hydrological processes that generate large floods (within the range of variability for each zone), it is argued that a reasonable design flood discharge is a value that has an even chance of both falling short and exceeding a flood magnitude; this central tendency of large floods is Q_{efp}. Overall, there is an insignificant difference between the Q_{efp} and the 100-year discharge (Q_{100}; 1% chance of occurrence; Yochum et al., 2019), though variability between Q_{efp} and Q_{100} may likely exist by zone and watershed size, with systematic differences in some areas. This approach does not utilize a frequency distribution, avoiding bias arising from data bimodality (or multimodality), also referred to as heavy tails of flood peak distributions (Merz et al., 2022), from mixed populations of flood-producing mechanisms (England et al., 2018).

The index flood frequency method, initially introduced by the USGS (Dalrymple, 1960), is a widely used flood frequency analysis method for prediction at ungaged locations as well as refining at-a-station estimates through the utilization of available flood data for a given region. Based on Hosking and Wallis (1993), an index-flood analysis based on L-moments is utilized in the FPP by assuming that all sites within a homogeneous region have the same distributions except for the scale or index-flood parameter. L-moments are estimated by linear combinations of order statistics (Hosking, 1990). A dimensionless frequency curve is developed to represent the ratio between flooding of any given frequency and an index flood. The mean annual flood is used as the scale or index-flood parameter (Hosking & Wallis, 1997). Relationships between the geomorphologic characteristics of watersheds and the mean annual flood are then developed enabling estimation of the mean annual flood at any point. A regional frequency curve is formed by combining the mean annual flood with the dimensionless frequency curve.

Within each flood potential zone, and with annual flood data available at N sites with a duration of n_i year at site i, a flood magnitude with a recurrence interval T at site i (Q_T^i) is expressed as:

$$Q_T^i = q_T u_i \ ; i = 1, 2, \ldots\ldots, N, \tag{1}$$

where u_i denotes the mean annual flood at site i, and q_T is the dimensionless regional frequency distribution, which remains the same throughout a zone for each recurrence interval. The ratios between annual peak flows and mean annual flood at site i, $q_{ij} = \frac{Q_{ij}}{u_i}$, where $j = 1, 2, \ldots., n_i$ and $i = 1, 2, \ldots\ldots, N$, are the basis for estimation of q_T (Hosking & Wallis, 1993). To estimate q_T, annual peak discharge data from streamgages in each flood potential zone are used to estimate at-site generalized extreme value distribution (GEV) parameters based on the L-moment approach. The at-site estimates were combined to give regional estimates:

$$\theta_k^{(R)} = \frac{\sum_{i=1}^{N} n_i \theta_k^{(i)}}{\sum_{i=1}^{N} n_i}, \tag{2}$$

where $\theta_k^{(i)}$ is the site i estimate of θ_k and $\theta_k^{(R)}$ is a weighted average, with the site i parameter estimate is given weight proportional to n_i. In this fashion GEV ξ (location), α (scale), and k (shape) regional parameters and q_T are estimated in each flood potential zone for recurrence intervals of $T = 1.5$, 2, 5, 10, 25, 50, 100, 200, and 500. Only streamgages used in the flood potential analysis were utilized in the index flood analysis. Multiple linear regression analysis are then carried out to link the natural logarithm of the mean annual flood (i.e., index-flood) to the watershed area, as well as a second watershed characteristic (where significant). The u_i is defined as $u_i = a A_i^b B_i^c$, where A_i denotes site i watershed area and B_i represents the second watershed characteristic. A computation example is provided in the user manual (Yochum et al., 2024).

The USGS provides a service of developing and publishing regional regression equations for predicting flood-frequency estimates at ungaged locations on a state-by-state basis across the United States (e.g., Capesius & Stephens, 2009; Kenney et al., 2007; Kohn et al., 2016; Miller, 2003; Waltemeyer, 2008). Generally, the development process has been for hydrologists to fit logPearson distributions (England et al., 2018; IACWD, 1982) to streamgage data across regions on a state-by-state basis, and develop regression equations for each standard recurrence interval (percent chance of exceedance) from these fitted distributions. These regression equations are then embedded within the USGS StreamStats application (Ries et al., 2017), for providing flood magnitude predictions alongside of predictions of a variety of other flow characteristics. Due to the dependance of the initial fit to logPearson distributions, this method is susceptible to bias. The FPP has been integrated with web services that support the StreamStats website, including watershed delineation and watershed characteristics calculated by the StreamStats Web Services and the National Stream Statistics Web Services for regional regression locations, metrics, and results.

2.5 | Streamgage analysis

As a fourth method for quantifying flood magnitudes for the selection of design flood discharges, the Streamgage Analyses module has been incorporated into the FPP for determining flood frequency relationships at streamgaged locations (England et al., 2018; IACWD, 1982). The FPP provides an additional software platform for performing these computations, as an alternative to existing tools (US Army Corps of Engineers HEC-SSP, USGS PeakFQ). The Streamgage Analysis module features map-based selection of a streamgage to be analyzed, automated data download, a plot of the annual peak discharges and Q_{efp}, data inspection/exclusion for outliers, and the quantification of generalized skew, peak discharge intervals, and perception thresholds. Flood frequency analysis results are presented using both station and generalized skews and can be exported for presentation with other predictions.

It is recommended to also utilize the Watershed Analysis module at streamgage analysis locations, to add additional results for ensemble decision making and to account for zone trends in the magnitudes,

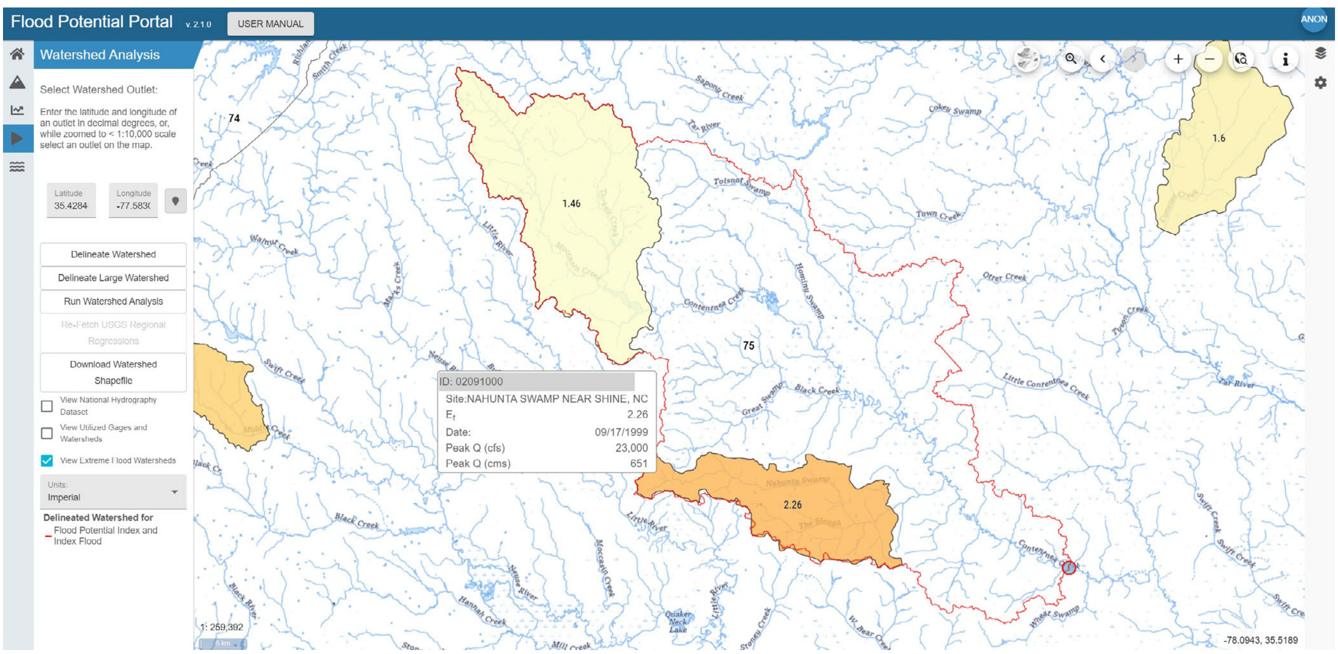

FIGURE 7 Watershed boundary for Contentnea Creek at NC-123, at Hookerton, NC. Two extreme floods have been recorded within this watershed ($E_f = 2.26$ and 1.46) in 1999 (Hurricane Floyd), with on-hover information provided for Nahunta Swamp. [Color figure can be viewed at wileyonlinelibrary.com]

Flood Potential Watershed Analysis: Results Overview

Date Generated: 07/10/2024 Latitude: 35.42846667
Version: 2.1.0 Longitude: -77.58305635

Indices -- Watershed Flooding Characteristics	Value	Percentile [Range]
Flood Potential Index (P_f)	13.4	61.2 [0 - 74.3]
Watershed Scale Ratio (R_f)	0.95	59.4 [0 - 5.79]
Flood Variability Index (V_f)	1.43	22.6 [0 - 2.86]
Beard Flash Flood Index (F)	0.79	59 [0 - 2.82]
Richards-Baker Flashiness Index (R-B)	0.306	50.2 [0 - 1.621]
Bimodality Index (Bi)	9	71.3 [0 - 182.8173]

Trends

(Southern) Atlantic Coastal Plain (75)

Zone 75 is experiencing a possible increasing trend in the magnitudes of largest (5%) of annual peak discharges (percent change = 8.4).
Zone 75 is experiencing a significant increasing trend in the magnitudes of largest quarter (Q4, >4 yr RI) of annual peak discharges (percent change = 17.5).
Zone 75 is experiencing no trend in the magnitudes of moderate quarter (Q3, 4 to 2 yr RI) of annual peak discharges.
Zone 75 is experiencing no trend in the magnitudes of ~bankfull quarter (Q2, 2 to 1.33 yr RI) of annual peak discharges.
Zone 75 is experiencing no trend in the magnitudes of < bankfull quarter (Q1, <1.33 yr RI) of annual peak discharges.

Zone 75 is experiencing a possible increasing trend in the annual frequency of floods.

Zone 75 is experiencing no trend in the event frequency of large floods.

Zone 75 is experiencing a significant increasing trend in flood flashiness.

Flood Discharge Estimates (cfs)	Unadjusted	Trend-Adjusted
Expected Flood Potential (Q_{efp})	47,400	55,700
Maximum Likely Flood Potential (Q_{mlf})	67,900	79,800

Return Interval (yr)	Percent Chance of Occurance	Index Flood	USGS Equations (StreamStats)
500	0.2	68,300	36,200
200	0.5	45,300	30,400
100	1	33,100	26,100
50	2	24,000	22,100
25	4	17,300	17,900
10	10	11,000	13,200
5	20	7,520	9,750
2	50	4,050	5,520
1.5	66.7	3,060	----

: directly comparable

FIGURE 8 Flood Potential Portal watershed analysis module results for Contentnea Creek at NC-123, at Hookerton, NC, including general flooding characteristics, observed trends in the magnitude, frequency, and flashiness of floods, and peak flood discharge estimates for ensemble decision making.

frequency, and flashiness of floods. Considering known issues with log-Pearson distributions biasing results in some situations, through data bimodality (and multi-modality) inducing heavy tails from mixed flood-producing mechanisms (England et al., 2018; Merz et al., 2022), it should not be assumed that streamgage analysis results are most appropriate for determining a design flood discharge. Comparison of results from the Streamgage Analysis and Watershed Analysis modules indicate that longer period streamgages may be more susceptible to such bias. Higher streamgage and zone-average bimodality index values indicate sites and areas where this problem may be more pervasive.

3 | EXAMPLE APPLICATION

To illustrate the value of the FPP for the selection of design flood discharges, an example is provided for Contentnea Creek at NC-123 at Hookerton, NC (at USGS streamgage 02091500) for a hypothetical bridge replacement. This stream is a tributary of the Neuse River on the Atlantic Coastal Plain, with a 1900 km^2 watershed east of Raleigh (Figure 7). For this example, recommendations are initially made assuming this is an ungaged location, and then consequently made using the streamgage flood-frequency results, to illustrate both the most common situation (ungaged) as well as how to incorporate streamgage analysis results where they are available.

Good practice for the use of the FPP in the selection of a design flood discharge is to initially explore flooding characteristics within the Mapping module. The Contentnea Creek watershed is in zone 75, in the (Southern) Atlantic Coastal Plain. Clicking on this zone within the Mapping module provides zone summary, data, and regional comparison information. Zone 75 experiences moderate-scale floods, when compared to zones across the United States: $P_f = 13.4$ (58th percentile), $R_f = 0.95$ (57th), $R\text{-}B = 0.31$ (47th), $B_i = 9.0$ (69th), and $V_f = 1.43$ (17th). Flood potential is less in the adjacent Carolina Sandhills zone to the immediate west (zone 75W, $P_f = 6.2$), with watersheds in zone 75 experiencing floods, on average, $13.4/6.2 = 2.2$ times larger. Flood potential is slightly higher (floods are larger) in the Piedmont (zone 74), to the west ($P_f = 15.8$). Zone 75 has experienced large floods most frequently in September (47.4% of total) and October (32.6%), but has also experienced large floods from February through August. Two extreme floods have been experienced within the Contentnea Creek watershed, from Hurricane Floyd in 9/1999, with flood extreme index (E_f) values of 2.26 and 1.46. Extreme floods also occurred within this zone during Hurricane Florence (9/2018), as well as an event in 10/2019. Considering both the dominant seasonality and these extreme events, it is apparent that tropical cyclones are a primary producer of large floods within this zone. Possibly increasing trends exist for the largest 5% and quarter (Q4) flood magnitudes (+8.4% and + 17.5%, respectively), with possibly increasing trends in annual flood frequency and significantly increasing trends in flashiness. Flood potential regressions utilize only watershed area, with the index flood regression also utilizing average annual precipitation. Explained variances are high ($R^2 = 0.97$ and 0.94, respectively).

The watershed analysis consists of the selection of the analysis point, delineation of the watershed using the automated tool (Figure 7), and computation of the flood values using flood potential, index flood, and regional regression methodologies. Results are exported in Excel format (Figure 8). For selection of the design flood discharge, grey-highlighted values are directly comparable. These comparable values are $Q_{efp} = 1340$ cms (47,400 cfs), $Q_{100,index} = 937$ cms (33,100 cfs), and $Q_{100,regional} = 739$ cms (26,100 cfs), with a trend-adjusted $Q_{efp} = 1580$ cms (55,700 cfs). The median value is often best for use as the design flood discharge (937 cms). The results of the USGS regional regression equations (from StreamStats) should not be used at this site since this method provides the lowest values and is least conservative. Additionally, it would be most appropriate to account for the observed increasing trend in large floods (to maximize resiliency), especially considering the observed increasing trends in annual flood frequency and flashiness, the occurrence of multiple extreme floods in this watershed, and increasing tropical cyclone rainfall in the southeast (Knight & Davis, 2009). Hence, use of the Watershed Analysis module indicates a recommended design flood discharge of $Q_{design} = 937*1.175 = 1100$ cms (38,900 cfs).

The streamgage analysis for this site was performed using 95 years of data, with systematic data collected from 1929 to 2023 (Figure 9). Annual peak discharge data indicate bimodal data, with two high-magnitude events (1999 and 2017: 903 and 782 cms, respectively) and the remaining 93 values averaging 136 cms. Results of the streamgage analysis, computed using the station skew and

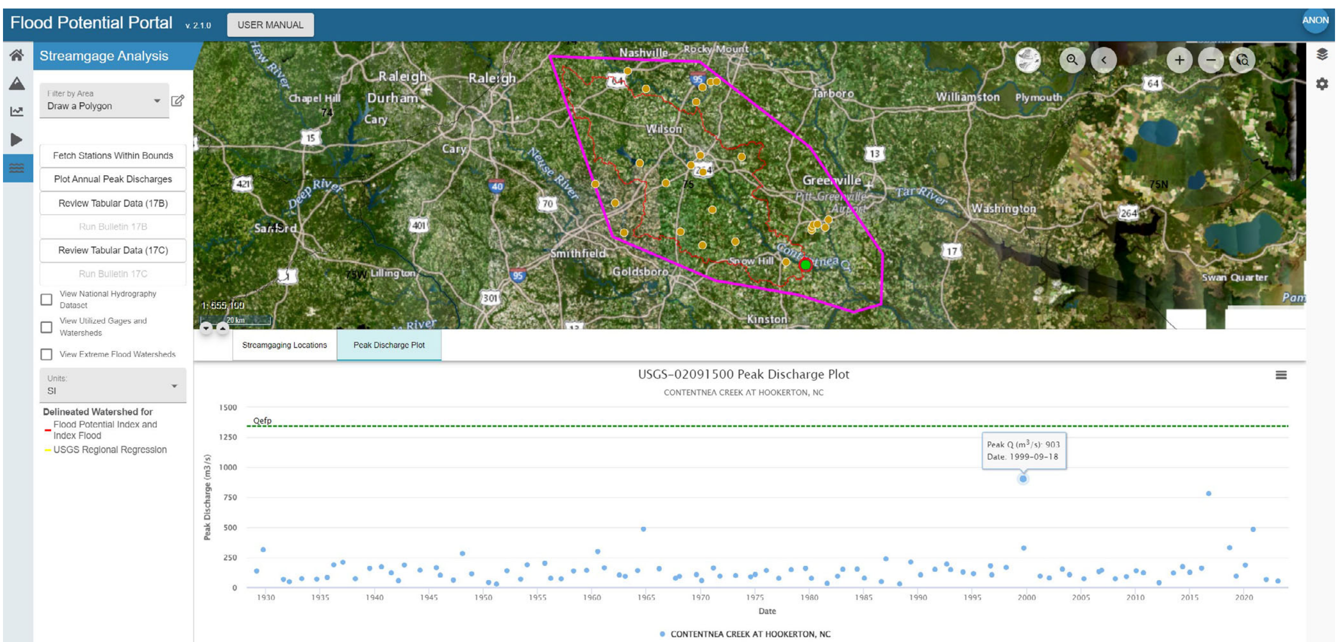

FIGURE 9 Annual peak discharge data for a streamgage on Contentnea Creek, NC (ID: 02091500). An on-hover box details the maximum recorded discharge of 903 cms, on 9/18/1999 (Hurricane Floyd), similar to the unadjusted (for the Q4 trend) $Q_{100\ index} = 933$ cms. [Color figure can be viewed at wileyonlinelibrary.com]

Flood Potential Portal Streamgage Analysis: Flood Frequency Results

Date Generated: 07/10/2024
Version: 2.1.0

USGS 02091500

Return Interval (yr)	Percent Chance of Occurence	Bulletin 17B Flood Method Estimates (cms)					
		Computed with Station Skew (0.4691)			Computed with Weighted Generalized Skew (0.2065) *		
		Magnitude	Confidence Interval Lower	Confidence Interval Upper	Magnitude	Confidence Interval Lower	Confidence Interval Upper
500	0.2	**1,030**	802	1,410	**844**	671	1,120
200	0.5	**785**	628	1,040	**675**	548	875
100	1	**633**	517	813	**564**	465	714
50	2	**504**	420	630	**464**	390	574
25	4	**394**	336	479	**375**	321	453
10	10	**275**	241	322	**272**	238	318
5	20	**201**	179	228	**203**	181	231
2	50	**116**	104	129	**119**	107	132
1.5	66.7	**90**	80	100	**92**	82	102
1.25	80	**72**	63	80	**72**	63	81
1.05	95.2	**47**	40	55	**45**	38	52

Return Interval (yr)	Percent Chance of Occurence	Bulletin 17C Flood Method Estimates (cms)					
		Computed with Station Skew (0.4616)			Computed with Weighted Generalized Skew (0.207) ^		
		Magnitude	Confidence Interval Lower	Confidence Interval Upper	Magnitude	Confidence Interval Lower	Confidence Interval Upper
500	0.2	**1,030**	699	2,420	**845**	658	1,170
200	0.5	**785**	568	1,550	**676**	539	901
100	1	**633**	478	1,100	**564**	458	732
50	2	**504**	397	775	**464**	385	586
25	4	**394**	324	545	**375**	317	461
10	10	**275**	236	339	**272**	236	322
5	20	**201**	176	234	**203**	180	234
2	50	**116**	103	130	**119**	107	132
1.5	66.7	**90**	80	100	**92**	82	102
1.25	80	**72**	64	80	**72**	64	80
1.05	95.2	**48**	41	54	**46**	39	52

Footnote
* Generalized skew coefficient sourced from: State Skew Map
^ Generalized skew coefficient sourced from: State Skew Map

FIGURE 10 Streamgage flood frequency results for Contentnea Creek, NC (ID: 02091500) using both Bulletins 17B (IACWD, 1982) and 17C (England et al., 2018) methodologies. The $Q_{100} = 633$ cms result, using the station skew, are identical for 17B and 17C at this streamgage and are substantially less than floods experienced at this site in 1999 (903 cms) and 2017 (782 cms). These two floods were not extreme as defined using the flood potential method but do contribute to the observed increasing trend (+17.5%) in the largest quarter of annual peak discharges in zone 75. The trend-adjusted watershed analysis result of $Q = 1100$ cms (38,900 cfs) is most appropriate as the design flood discharge for this hypothetical bridge replacement.

the weighted generalized skew (as populated using the "Use State Map" automated selection button), and using both the Bulletins 17B (IACWD, 1982) and 17C (England et al., 2018) methods, provides four sets of flood frequency results (Figure 10). The predictions are very similar between the 17B and 17C methods. Use of the station skew results in the largest peak flow estimates, with $Q_{100} = 633$ cms (22,300 cfs) and $Q_{500} = 1030$ cms (36,400 cfs). Application of the observed increasing trend from the Q4 flood potential analysis (+17.5%) increases the Q_{100} to 743 cms (26,300 cfs), 32% less than the Watershed Analysis module results and insufficiently conservative. Despite (or as a result of) a long record length that encourages confidence in the results of streamgage analyses, the logPearson distribution is biasing the Q_{100} low, likely due to the bimodal dataset (streamgage $B_i = 7.32$). The Watershed Analysis module result, utilizing the observed trend adjustment ($Q_{design} = 1100$ cms $= 38{,}900$ cfs), should be used as the design flood discharge for this hypothetical bridge replacement.

4 | SUMMARY AND CONCLUSIONS

The Flood Potential Portal (https://floodpotential.erams.com/) is introduced for use by practitioners, for the design of road-stream crossings and other stream valley infrastructure, as well as for floodplain and riparian ecosystem management. The observational network (streamgages) are implemented in these analyses. This decision support system was developed to enhance understanding of how floods vary in magnitude from place to place, how flood severity is observed to be changing due to climate change and other sources of nonstationarity, and to provide tools for peak discharge prediction using multiple methods at both ungaged and streamgaged locations. Providing results from multiple methods allows for ensemble decision making for the selection of design flood and base flood discharge, to increase resilience. Additional research is needed to project trends in floods into the future and enhance streamgage flood frequency analyses to address data bimodality/multimodality, to extend the capabilities of the FPP. Additionally, research is needed to understand the mechanisms causing flood variability across the United States, in both space and time.

ACKNOWLEDGEMENTS

Reviews of an earlier version of this manuscript by Dan Scott and an anonymous reviewer have been greatly appreciated; their suggestions have improved the article and enhanced the utility of the Flood Potential Portal. David Levinson, Heidi Klingel, Joel Murray, and Julian Scott, of the 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, staff at One Water Solutions Institute at Colorado State University, namely Dave Patterson, Lucas Yaege, and Holm Kipka, are appreciated for their work with the Flood Potential Portal development. Appreciation is also expressed to leadership support from the Forest Service Biological and Physical Resources Washington Office directorate and the National Forest System deputy chief's office (Challenge Cost Share agreement #20-CS-11132422-274), and 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 results presented within the Flood Potential Portal.

FUNDING INFORMATION

Funding for development of the Flood Potential Portal was provided through the National Stream and Aquatic Ecology Center by a US Forest Service challenge cost share agreement (#20-CS-11132422-274).

DATA AVAILABILITY STATEMENT

Supporting data for the flood potential method and the Flood Potential Portal are available for download from the project page (https://www.fs.usda.gov/biology/nsaec/projects-floodpotential.html). Algorithms are provided within this project page or as zone summary information within the Mapping module.

ORCID

Steven E. Yochum https://orcid.org/0000-0002-7973-1134

REFERENCES

Baker, D. B., Richards, R. P., Loftus, T. T., & Kramer, J. W. (2004). A new flashiness index: Characteristics and applications to midwestern rivers and streams. *Journal of the American Water Resources Association*, *40*, 503–522.

Capesius, J. P., & Stephens, V. C. (2009). Regional regression equations for estimation of natural streamflow statistics in Colorado. U.S. Geological Survey Scientific Investigations Report 2009–5136.

Dalrymple, T. (1960). Flood-Frequency Analyses. Manual of Hydrology Part 3. Flood-flow techniques. Geological Survey Water-Supply Paper, 1543-A.

David, O., Lloyd, W., Rojas, K., Arabi, M., Geter, F., Ascough, J., Green, T., Leavesley, G., & Carlson, J. (2014). Model-as-a-service (MaaS) using the cloud services innovation platform (CSIP), international environmental modelling and software society (iEMSs). In D. P. Ames, N. W. T. Quinn, & A. E. Rizzoli (Eds.), *Proceedings of 7th int. congress on env. Modelling and software*.

Dettinger, M. D., Ralph, F. M., Das, T., Neiman, P. J., & Cayan, D. R. (2011). Atmospheric rivers, floods, and water resources of California. *Water*, *3*, 445–478. https://doi.org/10.3390/w3020445

England, J. F., Cohn, T. C., Faber, B. A., Stedinger, J. R., Thomas, W. O., Veilleux, A. G., Kiang, J. E., & Mason, R. R. (2018). Guidelines for determining flood flow frequency—bulletin 17C. In *U.S. Geological Survey Techniques and Methods, book 4*, Chap. B5 (p. 148). U.S. Geological Survey.

Francois, B., Schlef, K. E., Wi, S., & Brown, C. M. (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*(3), 563–570. https://doi.org/10.1111/j.1752-1688.2011.00550.x

Guirguis, K., Gershunov, A., Shulgina, T., Clemesha, R. E. S., & Ralph, F. M. (2019). Atmospheric rivers impacting northern California and their modulation by a variable climate. *Climate Dynamics*, *52*, 6569–6583. https://doi.org/10.1007/s00382-018-4532-5

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 (Methodological)*, *52*(1), 105–124.

Hosking, J. R. M., & Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research*, *29*(92), 271–281.

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: An approach based on L-moments*. Cambridge University Press. https://doi.org/10.1017/CBO9780511529443

Interagency Advisory Committee on Water Data. (1982). *Guidelines for determining flood flow frequency: Bulletin #17B for the hydrology subcommittee*. US Department of the Interior, Geological Survey, Office of Water Data Coordination.

Kenney, T. A., Wilkowske, C. D., & Wright, S. J. (2007). Methods for estimating magnitude and frequency of peak flows for natural streams in Utah. US Geological Survey Scientific Investigations Report 2007–5158.

Knight, D. B., & Davis, R. E. (2009). Contribution of tropical cyclones to extreme rainfall events in the southeastern United States. *Journal of Geophysical Research*, *114*, D23102. https://doi.org/10.1029/2009JD012511

Kohn, M. S., Stevens, M. R., Harden, T. M., Godaire, J. E., Klinger, R. E., & Mommandi, A. (2016). Paleoflood investigations to improve peak-streamflow regional-regression equations for natural streamflow in eastern Colorado, 2015. US Geological Survey Scientific Investigations Report 2016–5099.

Lins, H. F., & Cohn, T. A. (2011). Stationarity: Wanted dead or alive? *Journal of the American Water Resources Association*, *47*(3), 475–480. https://doi.org/10.1111/j.1752-1688.2011.00542.x

Lloyd, W., Pallickara, S., David, O., Arabi, M., Wible, T., & Ditty, J. (2015). Demystifying the clouds: Harnessing resource utilization models for cost effective infrastructure alternatives. *IEEE Transactions on Cloud Computing*, *5*(4), 667–680.

Lloyd, W., Pallickara, S., David, O., Lyon, J., Arabi, M., & Rojas, K. (2013). Performance implications of multi-tier application deployments on infrastructure-as-a-service clouds: Towards performance modeling. *Future Generation Computer Systems*, *29*(5), 1254–1264. https://doi.org/10.1016/j.future.2012.12.007

Merz, B., Basso, S., Fischer, S., Lun, D., Bloschi, G., Merz, R., Guse, B., Viglione, A., Vorogushyn, S., Macdonald, E., Wietzke, L., & Schumann, A. (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. US Geological Survey Water-Resources Investigations Report 03–4107.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, *319*, 573–574. https://doi.org/10.1126/Science.1151915

Ralph, F. M., Cordeira, J. M., Neiman, P. J., & Highes, M. (2016). Landfalling atmospheric rivers, the sierra barrier jet, and extreme daily precipitation in northern California's upper Sacramento River watershed. *Journal of Hydrometeorology*, *17*, 1905–1914. https://doi.org/10.1175/JHM-D-15-0167.1

Ries, K. G., Newson, J. K., Smith, M. J., Guthrie, J. D., Steeves, P. A., Haluska, T., Kolb, K., Thompson, R. F., Santoro, R. D., & Vraga, H. W. (2017). Streamstats, version 4. US Geological Survey Fact Sheet. 2017–3046 https://doi.org/10.3133/fs20173046

Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L., & Bates, P. (2018). Strange floods: The upper tail of flood peaks in the United States. *Water Resources Research*, *54*, 6510–6542. https://doi.org/10.1029/2018WR022539

Tarouilly, E., Dongyue, L., & Lettenmaier, D. P. (2021). Western US Superfloods in the recent instrumental record. *Water Resources Reseearch*, *57*, 1–15. https://doi.org/10.1029/2020WR029287

Waltemeyer, S. D. (2008). Analysis of the magnitude and frequency of peak discharge and maximum observed peak discharge in New Mexico and surrounding areas. *US Geological Survey Scientific Investigations Report*, 2008–5119.

Yochum, S. E. (2019). Flood potential in the southern Rocky Mountains region and beyond. In *SEDHYD-2019 Conference Proceedings*, June 14 to 28, Reno, NV.

Yochum, S. E., & Levinson, D. L. (2023). Flood variability in the Western United States: Overview and examples. In *SEDHYD-2023 Conference Proceedings*, May 8 to 12, St. Louis, MO.

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

Yochum, S. E., Wible, T., Ghanbari, M., & Millonig, S. (2024). Flood Potential Portal: User Manual. One Water Solutions Institute and the National Stream and Aquatic Ecology Center, Colorado State University and the US Forest Service, Version 2.1. https://floodpotential.erams.com/floodpotential/download_user_manual/

How to cite this article: Yochum, S. E., Wible, T., Korsa, M., Ghanbari, M., & Arabi, M. (2024). Flood Potential Portal: A web tool for understanding flood variability and predicting peak discharges. *River Research and Applications*, 1–13. https://doi.org/10.1002/rra.4354