

## Article

# Explaining the Flood Behavior for the Bridge Collapse Sites

Fahmidah Ashraf <sup>1,\*</sup>, Hristos Tyralis <sup>2</sup>  and Georgia Papacharalampous <sup>3</sup> <sup>1</sup> Department of Civil Engineering and Construction, Bradley University, Peoria, IL 61625, USA<sup>2</sup> Construction Agency (YPEPA), Hellenic Air Force, 15 561 Cholargos, Greece<sup>3</sup> Department of Water Resources and Environmental Engineering, National Technical University of Athens, 157 80 Zographou, Greece

\* Correspondence: fashraf@fsmail.bradley.edu

**Abstract:** Given the increasing intensity and frequency of flood events, and the casualties and cost associated with bridge collapse events, explaining the flood behavior for the collapse sites would be of great necessity. In this study, annual peak flows of two hundred and five watersheds, associated with two hundred and ninety-seven collapse sites, are analyzed. Generalized Extreme Value distribution together with other statistical analyses are used to derive and analyze the shape parameters of the distributions which represent the extremeness of flood events. Random forest mechanism is employed in order to identify the predictor variables (and the associated importance levels) for the shape parameters. Peak flows are also classified in order to find the extremes and the associated return periods. The results indicate that most of the bridge collapse sites across different physiographic regions, i.e., Appalachian Highland, Central Lowland, Coastal Plain, and Interior Highlands, exhibit common characteristics such as (a) variation of important predictor variables, (b) human interference, (c) extremeness of flood events similar to the regions with hydrologic heterogeneity, and (d) frequent occurrence of extreme flows. These results indicate a commonality in flood behavior, as stems from specific settings, for the collapse sites studied. The findings instigate the revisiting of the bridge design practices and guidelines and provide some basis to assess the risk of future collapse.



**Citation:** Ashraf, F.; Tyralis, H.; Papacharalampous, G. Explaining the Flood Behavior for the Bridge Collapse Sites. *J. Mar. Sci. Eng.* **2022**, *10*, 1241. <https://doi.org/10.3390/jmse10091241>

Academic Editors: Denis Istrati, Ian Buckle and Michael Scott

Received: 21 June 2022

Accepted: 26 August 2022

Published: 3 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** flood behavior; bridge collapse; shape parameter; extreme flows

## 1. Introduction

Floods and other hydraulic events are perceived to be the most common causes of total or partial bridge collapse in the United States (U.S.) [1]. About 62.23% of overwater bridge collapses correspond to hydraulic events [1], with an annual hydraulic collapse frequency of approximately 1/5000 [1,2]. Probable hydraulic effects causing bridge collapses include scour, hydrodynamic horizontal loads applied on the piers, hydrodynamic loads applied on the superstructure (i.e., deck), debris jam, and/or a combination of these effects. The topics of scouring [1,3–5] and hydrodynamic horizontal loads (on piers) [6,7] have been studied extensively, whereas hydrodynamic loads on the superstructure are being studied more recently [8,9]. Although several studies exist focusing on the effects of debris jam [10–16], it is one of the most challenging topics under investigation by the hydraulic engineering community. Given that all of these hydraulic effects can be induced and exaggerated by floods, investigating flood events are of particular concern with the increasing intensity and frequency of recent floods in the U.S. It is noted that scour—particularly scour during floods—alone has been estimated to cause the collapse of 20–100 bridges per year in the U.S. [1,3,4]. More frequent or intense flooding is linked with climate and land use change [17]; the associated high risk is linked with direct and indirect costs, casualties, and user delays [1,3,4,18]. The 2021 American Society of Civil Engineers (ASCE) Report Card noted that a \$22.7 billion annual investment is needed to substantially improve the current bridge conditions [19].

In response to reducing collapse risk due to floods, engineers primarily aim at determining the magnitude and frequency of design floods. Methods commonly used for estimating the return period of a flood event include block maxima approaches, in which a series of annual peak flows is used to define an extreme value distribution. The generalized extreme value distribution (GEV) is a widely used model for extreme events [20]. The major benefit of the GEV model is its ability to fit highly skewed data since it combines three distributions: Gumbel, Frechet, and Weibull [20]. The GEV distribution has three parameters, namely, the location and scale parameters which are mostly related to the magnitude of flow [21], and the shape parameter that determines the behavior (or shape) of the tail of the distribution [21].

Although it is obvious that flood behavior and/or flood hazard depends on the climate, soil, topography, and other watershed conditions, the specific attributes which can be used as predictors for flood behavior are not apparent. Whereas there exists a plethora of studies on site specific flood hazard, there is a knowledge gap in relation to finding the exact relationships among flood behavior and regional attributes. Regional study and/or large-sample catchment hydrology is a common practice in relation to the 2003–2012 initiative of the hydrological community for predictions in ungauged basins [22]. In particular, such approach is called regionalization of parameters using regression algorithms [22]. The large spatial scale is considered an advantage for hydrologic studies in that it includes more diverse basins; therefore, statistical characterizations would be more complete and generalizable in their extrapolation ability. Whereas local study provides more refined, site-specific analysis, their uses are limited because of their limited extrapolation ability. The regional-based approach is also in advantage since the comparison study can be conducted to identify the similarities and differences among different regions within a country or across countries, and to interpret the findings in terms of underlying climate–watershed–soil–topography–human controls [22].

Many hydraulic studies have also drawn attention to the regional scale [23–27], particularly in the presence of human intervention in the water system [23] for assessing the risk of the bridge collapse. There is an emphasis on physiographic settings when assessing bridge–stream intersections since the risk issues/events are intrinsically different for different physiographic regions [27]. Designing bridge scour countermeasures ignoring the regional scale hydrology sometimes results in continuous local maintenance of the structure [23]. For instance, debris jam is often caused by factors acting at a larger scale and addressing such a problem requires consideration of a large spatial scale [23]. In fact, assessing channel instability in relation to both regional and local scales is recommended by the Federal Highway Administration [27].

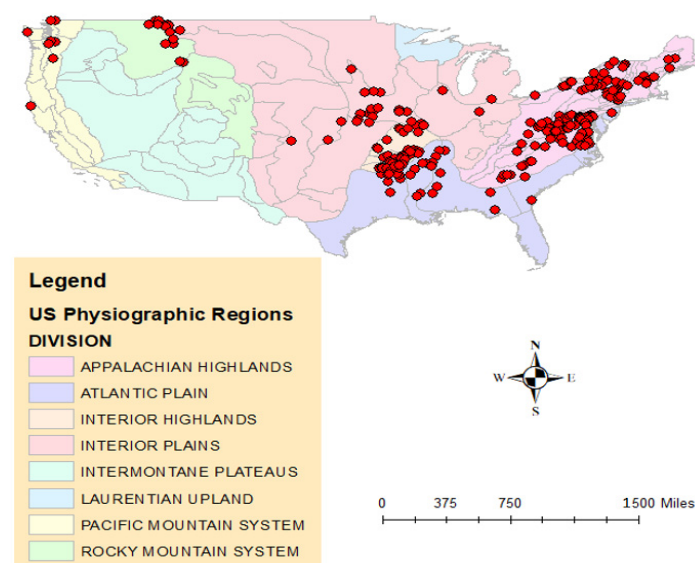
The goal of the study is to explain flood behavior for 297 bridge collapse sites in relation to the shape parameters of the GEV distribution within a regional context. To attain the goal, three objectives are identified: (1) to derive shape parameters of GEV distribution using annual peak flow data, (2) to identify predictor variables for the shape parameter in relation to specific physiographic regions and all regions combined, and (3) to identify extreme flows with associated return periods. The values of shape parameters would reveal the overall extremeness of the flood events. The predictor variables within the regional context would help to identify important hydrological signatures in relation to the extremeness of the flood events. Finally, the identification of extreme flows with return periods would help to identify the uniqueness of the flood events for the collapse sites, if any. The results can reveal any anomaly and/or trend in the behavior of peak floods for the bridge collapse sites and can support the understanding of the mechanism behind the generation of collapse-inducing floods. Since analyzing peak flow distribution parameters is not a common practice in bridge design procedures, such regional analysis can provide preliminary data to assess future collapse risk, as it can be highlighted that the flood event occurrence should be expected within a specific context. Concerning the bridge collapse sites, the findings of the present study are also original in comparison to previous studies in which the shape parameters were examined.

## 2. Methods

### 2.1. Selection of Sites

The New York State Department of Transportation (NYSDOT) bridge failure database is used here to identify 297 bridge collapse sites (Figure 1). NYSDOT is the only U.S.-wide database of bridge collapses. The recorded information includes the NYSDOT failure database ID, identifier in the National Bridge Inventory (NBI), failure cause, the location of the collapsed bridge, the feature (or stream) under the bridge, the year of construction, the date or year of collapse, the bridge material and structure type, the type of collapse (total or partial), the number of casualties related to the collapse, and other comments. The NYSDOT bridge failure database includes information based on available information, that is, searches of journalism databases and surveys of other DOTs. An overrepresentation of collapsed bridges in a certain region/state, therefore, does not imply comparatively more collapse events; rather, it only implies the availability of the specific collapse information in that region [28]. In the current study, the selection of hydraulic collapse sites from the NYSDOT database is guided by the availability of the required information at the recorded collapse sites. The selected bridges were located in Appalachian Highland (140 sites), Coastal Plain (47 sites), Central Lowland (42 sites), Interior Highlands (54 sites), Pacific Mountain (8 sites), and Rocky Mountain (8 sites). The selection criteria are discussed in detail below:

- Existence of a stream gauge listed in the U.S. Geological Survey (USGS) National Water Information System Database at the bridge location, near the bridge location (on the same tributary of the river or on a different tributary), or at a further distance (not the same tributary, but on the same river or within the same watershed).



**Figure 1.** A total of 297 bridge collapse sites across different physiographic regions in the U.S.

- Availability of climate, soil, topography, watershed, and human interference data in relation to the USGS stations in the Geospatial Attributes of Gages for Evaluating Streamflow, version II (GAGES II) database.
- Bridges were apparently collapsed (complete or partial collapse) due to hydraulic effects including scour, debris, and hydraulic pressure/load on the piers and the superstructure. Each hydraulic effect and/or a combination of them could result in the washout of a bridge, which is the case for a couple of bridges. Some collapses are caused by specific flood and hurricane events as reported by NYSDOT. Whereas hydraulic effects are exaggerated during floods and hurricane events, hurricanes are much more complex, multifaceted natural hazards. Hurricanes can cause additional rainfall, extreme waves, and storm surge. There are in-depth studies that focus on

the estimation of hurricane-induced hydrodynamic effects on structures [29–33]. Both for bridges [29,30] and buildings [31–33], it is proven that storm surges and waves, induced by hurricanes, can exert large uplift pressures/forces and smaller horizontal forces leading to potential collapse. With the presence of debris, the effect of any hydraulic event can become more severe. The effects of debris include large, short-duration horizontal impact forces on piers [10,11], both horizontal and uplift forces at different locations of the bridge superstructure [12], and long-duration damming effects both on the pier [13–15] as well as on the bridge superstructure [16]. Debris jams can also significantly increase the collapse risk, particularly in three specific ways: (i) increasing the drag force exerted by the flow [34], (ii) substantially increasing the upstream flood levels [35,36], and (iii) exacerbating the scour [14]. In fact, the accumulation of debris at piers is being acknowledged as a key mechanism of affecting the interaction between bridges, rivers, and floods [14], and contribute to more than one-third of bridge collapses in the United States [13].

Retrieved information of the selected bridge collapse sites (with associated USGS stations) can be found at <https://github.com/fahmidah/Shape-Parameter-Bridge-Collapse-Sites/tree/main/List%20of%20Collapse%20Sites> (accessed on 10 March 2022).

## 2.2. Shape Parameters

Shape parameters are retrieved for the GEV distribution using the *extRemes R* package. Annual peak flow data, retrieved from the United States Geological Survey (USGS), are used to fit with the GEV distribution. The cumulative distribution function of the GEV distribution is given by the following equation [37–41].

$$F(x/\theta) = \exp\left(-(1 + k(x - \mu)/\sigma)^{\pm 1/k}\right), \theta = (\mu, \sigma, k), \sigma > 0$$

Here,  $\mu$  is the location parameter,  $\sigma$  is the scale parameter, and  $k$  is the shape parameter. The shape parameter is of critical significance in that it represents the type of distribution: Gumbel (shape parameter = 0), Frechet or heavy tail (shape parameter > 0), and Weibull (shape parameter < 0) for fitting to block maxima series of data. Therefore, the shape parameter is related to how extreme the floods are. The higher values of  $k$  result in heavier tails.

## 2.3. Predictor Variables

The variables used in the current work are retrieved from the “GAGES II dataset” [42]. The dataset provides geospatial data and classifications for a total of 9322 stream gages maintained by USGS. The GAGES II dataset consists of gages that have had either 20+ complete years (not necessarily continuous) of discharge records since 1950 or are currently active. The geospatial data include watershed characteristics compiled from national data sources, including environmental features (e.g., climate—including precipitation, geology, soils, topography) and anthropogenic influences (e.g., land use, road density, or presence of dams). The dataset also includes comments from local USGS Water Science Centers, based on Annual Data Reports, pertinent to hydrologic modifications and influences. The dataset and associated detail description are available to the public at [https://water.usgs.gov/lookup/getspatial?gagesII\\_Sept2011](https://water.usgs.gov/lookup/getspatial?gagesII_Sept2011) (accessed on 29 August 2022). For the current study, specific variables are retained which are continuous, have physical meaning, and are not related to each other in an obvious way. For instance, ‘DDENS\_2009’ and ‘pre1990\_DDENS’ refer to dam density currently and pre-1990, respectively; both data are continuous and have physical meaning. However, they are related to each other; therefore, only the most updated data, ‘DDENS\_2009’, is included in the analysis. There are a total of 40 distinct variables selected for the current study, as shown below:

- 15 variables related to climate: watershed average of annual number of days of measurable precipitation, watershed average of mean day of the year of last freeze, watershed average of monthly maximum number of days of measurable precipitation, mean an-

nual precipitation (cm) for the watershed, watershed average of mean day of the year of first freeze, watershed average of minimum monthly air temperature (degrees C), average annual air temperature for the watershed (degrees C), mean annual potential evapotranspiration, standard deviation of maximum monthly air temperature, watershed average of maximum monthly air temperature, precipitation seasonality index of how much annual precipitation falls seasonally, watershed average of monthly minimum number of days of measurable precipitation, standard deviation of minimum monthly air temperature (degrees C), watershed average relative humidity (percent), and snow percent of total precipitation estimate.

- 7 variables related to watershed: fragmentation index of “undeveloped” land in the watershed, watershed percent “planted/cultivated” (agriculture), watershed percent “forest”, watershed percent “developed” (urban), watershed drainage area (sq km), watershed compactness ratio, density (per square km) of lakes/ponds, and reservoir water bodies.
- 6 variables related to soil: average value for the range of available water capacity for the soil layer, average value of sand content (percentage), average value of silt content (percentage), average value of clay content (percentage), average value of bulk density (grams per cubic centimeter), and average permeability (inches/hour).
- 5 variables related to topography: mean watershed elevation (meters), standard deviation of elevation (meters) across the watershed, aspect “north”, aspect “east, and mean watershed slope (percent).
- 3 variables related to population infrastructure: watershed percent impervious surfaces, population density in the watershed, and road density (km of roads per watershed square km).
- 4 variables related to dams: dam density, dam storage in watershed (normal storage), dam storage in watershed (NID storage), and major dam density.

A brief summary of the predictor variables used in the study is provided in Appendix A. The summary includes variable names, descriptions, types, units, time period, extent of data, data processing method, the scale of source data, and source of data.

For the selected variables, the random forests mechanism [43] is used to derive physiography-based important levels. Pacific Mountain and Rocky Mountain System were excluded from the physiography-based study because of the low number of sites. The methodology is then also implemented considering all the collapse sites together across different physiographic regions in the U.S. Random forests is a machine learning algorithm of increasing interest because of its excellent predictive performance [43]. Because of its ability to find important predictor variables, the mechanism has been used in geoscience [44,45], and in hydrology, particularly, to interpret the characteristics and relationships among different hydrological signatures [46,47]. In the current study, the random forests algorithm is used for regression [48], and the estimation of the importance levels [48,49] of different predictor variables, that is, the contribution of each input variable in predicting the response within a regional context. A negative importance level means that inclusion of the predictor variable results in a decrease in the performance of the algorithm. Positive values indicate a positive contribution to the prediction of the algorithm. The *randomForest* R package [50,51] is used to implement the algorithm. The algorithm requires little tuning, and its performance is very good when using the default parameters, i.e., the number of trees, the maximum number of terminal nodes of the trees, and the minimum size of terminal nodes. R scripts used for the analysis can be retrieved from <https://github.com/fahmidah/Shape-Parameter-Bridge-Collapse-Sites/blob/main/Random%20Forest/Analysis> (accessed on 10 March 2022).

#### 2.4. Classification of Peak Flows

This study employs Jiang’s classification scheme [52] to categorize annual peak flows into heavy tail and light tail flows. Heavy tail flow implies that flows in the heavy tail increase more than the exponential order as compared to the flows in the light tail, and



such high flows are expected to be with higher return periods. With Jiang's classification scheme [52], the minimum heavy tail flow would be identified. To retrieve the return periods of minimum and maximum heavy tail flows, the annual peak flow data are fitted with GEV distribution, and the task is performed by implementing the *extRemes* R package.

### 3. Results

#### 3.1. Shape Parameters

For physiographic regions studied, as shown in Figure 2, shape parameters do present fairly normal distributions ( $p$ -value of the Shapiro-Wilk normality test  $> 0.05$ ). The extremeness of the flood behavior is quite similar across the study regions, as evident by the similar shape parameter values (mean and median). In fact, the median value of the shape parameter (Table 1) within each physiographic region is close to the median value of 0.19 as obtained for the watersheds with a diversity of climate types [43]. The site with the largest shape parameter value (0.74) is located within Central Lowland (USGS station #6108000) (Table 1), which implies a higher likelihood of obtaining extreme flows at this specific location. Here, the heavy tail flows are defined as extreme flows since they increase more than the exponential order as compared to the flows in the light tail. Sites within Coastal Plain experience much broader fluctuations (e.g., three or more standard deviations from the mean) in peak flows; such finding is evident by the value of the kurtosis ( $>3$ ) in Table 1. A distribution is said to be leptokurtic when the kurtosis is greater than 3, implying that there is a greater potential for extreme fluctuations in the observed values. Coastal areas typically experience various kinds of flooding caused by high tides, heavy rain, storm surge, and high waves. Extreme coastal hazards (i.e., hurricanes) do not occur frequently, but when they do occur, they have a dominant effect on the rainfall and the level of flooding. Such unique characteristics of the coastal area might explain the broader fluctuation in the peak flow values. All the shape parameters' values for 205 stations can be retrieved from <https://github.com/fahmidah/Shape-Parameter-Bridge-Collapse-Sites/tree/main/Shape%20Parameters> (accessed on 10 March 2022).

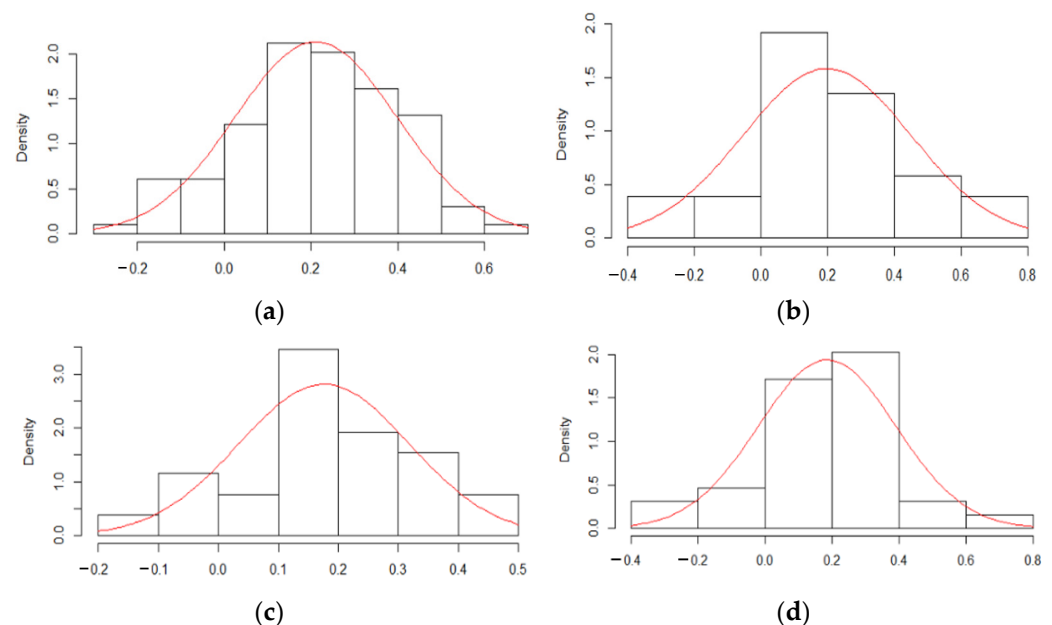
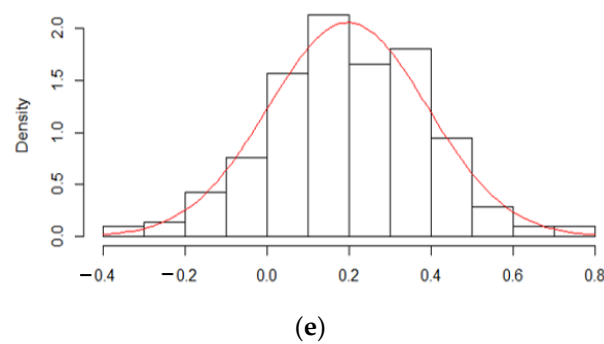


Figure 2. Cont.



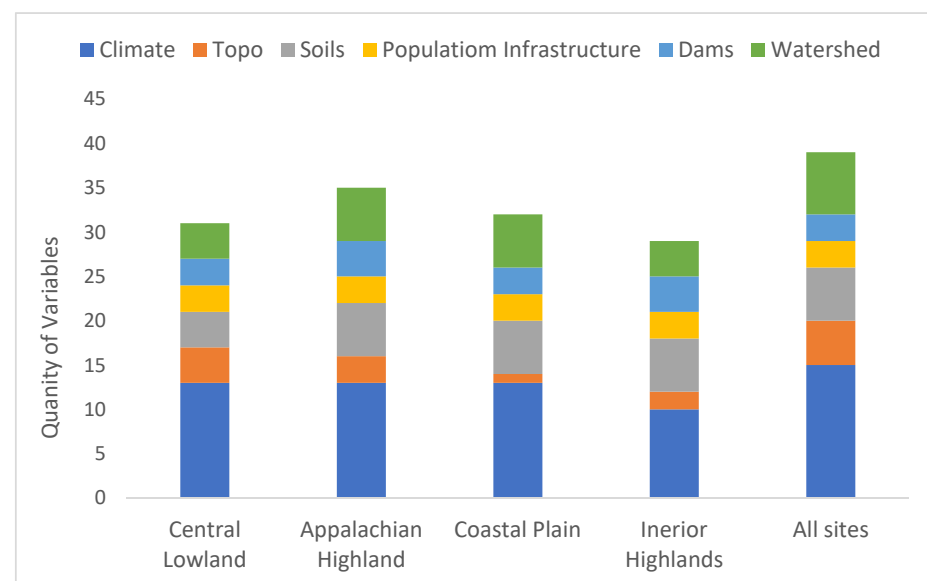
**Figure 2.** Normal distributions of the shape parameters derived from GEV estimates of the annual peak flows for the bridge collapse sites: (a) Appalachian Highland; (b) Central Lowland; (c) Interior Highland; (d) Coastal Plain; (e) All sites.

**Table 1.** Descriptive statistics of the shape parameter values for the bridge collapse sites.

Physiographic Region	Range	Mean	Median	Standard Deviation	Kurtosis	Shapiro-Wilk Normality Test
Appalachian Highland	−0.21–0.61	0.21	0.20	0.19	2.58	0.16
Central Lowland	−0.31–0.74	0.20	0.19	0.25	2.96	0.65
Interior Highland	−0.11–0.43	0.18	0.17	0.14	2.43	0.53
Coastal Plain	−0.32–0.63	0.18	0.2	0.21	3.28	0.64
All Sites	−0.32–0.74	0.20	0.19	0.19	3.02	0.85

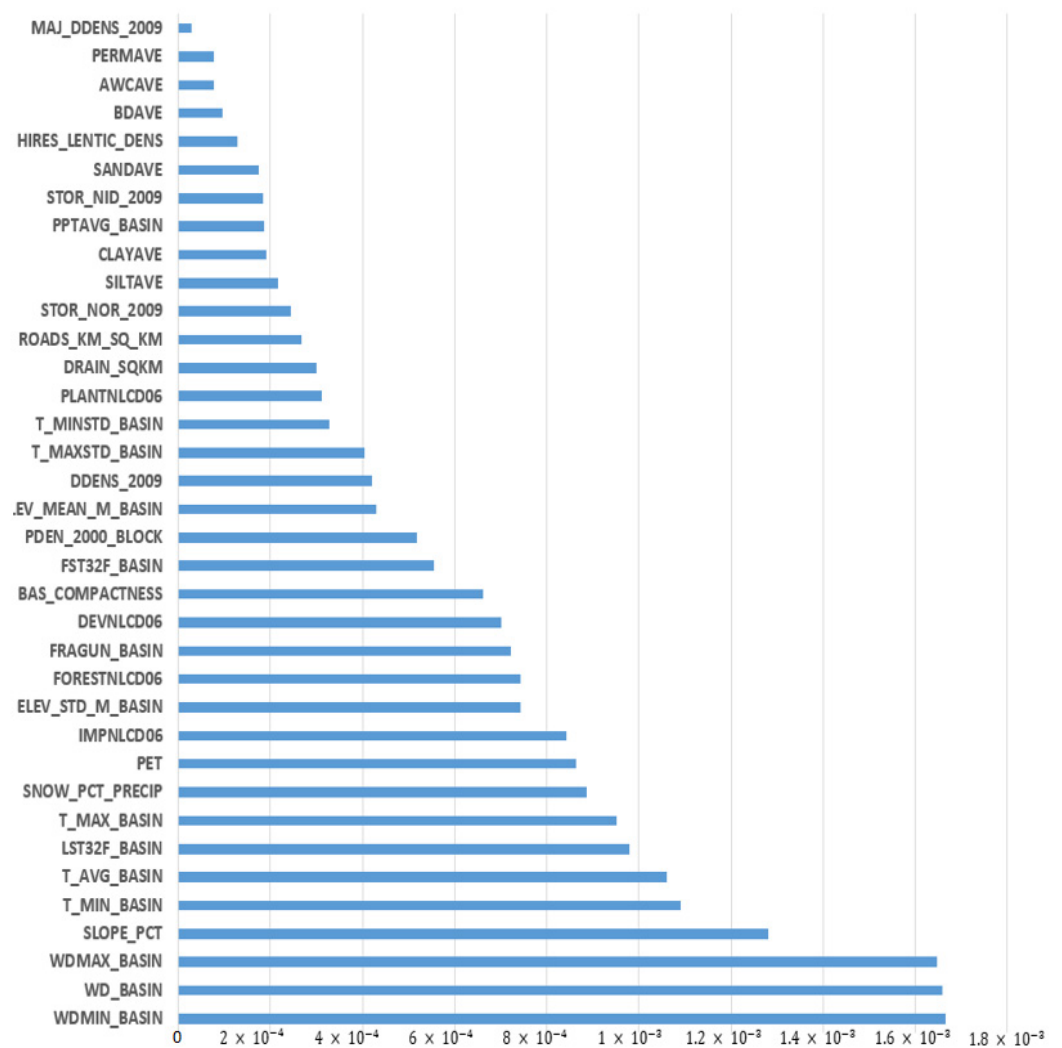
### 3.2. Predictor Variables

Among the six types of predictor variables, climate variables are found in a greater number across all regions (Figure 3). Predictor variables with associated importance levels for each physiographic region and all sites combined are provided in Appendix B.



**Figure 3.** Types of predictor variables across different physiographic regions.

For Appalachian Highland, there are a total of 36 predictor variables (with positive importance levels) identified: climate (13), watershed (7), soils (6), topography (3), dams (4), and population infrastructure (3) (Figure 4, see also Appendix B).



**Figure 4.** Predictor variables of the shape parameters (with associated importance) for Appalachian Highland. Description of predictor variables is provided in Appendix A.

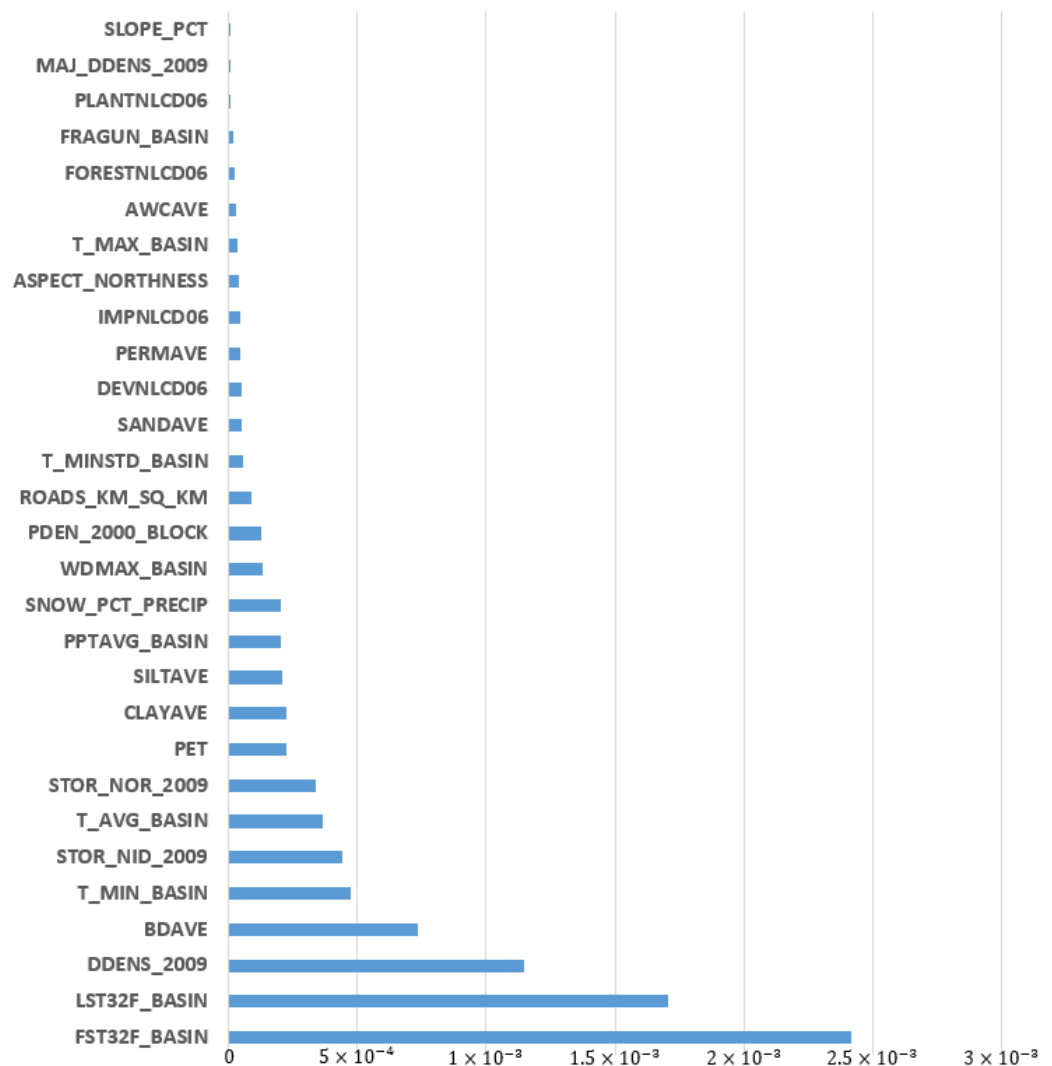
For Interior Highlands, 29 predictor variables (with positive importance levels) are identified: climate (10), soils (6), watershed (4), dams (4), population infrastructure (3), and topography (2) (Figure 5, see also Appendix B).

In total, 32 predictor variables (with positive importance levels) are identified for Central Lowland: climate (13), watershed (5), soils (4), topography (4), dams (3), and population infrastructure (3) (Figure 6, see also Appendix B).

For Coastal Plain, 32 predictor variables (with positive importance levels) are identified: climate (13), watershed (6), soils (6), dams (3), population infrastructure (3), and topography (1) (Figure 7, see also Appendix B).

When considering all sites together, 39 predictor variables (with positive importance levels) are identified: climate (15), topography (5), soils (6), population infrastructure (3), dams (3), and watershed (7) (Figure 8, see also Appendix B).



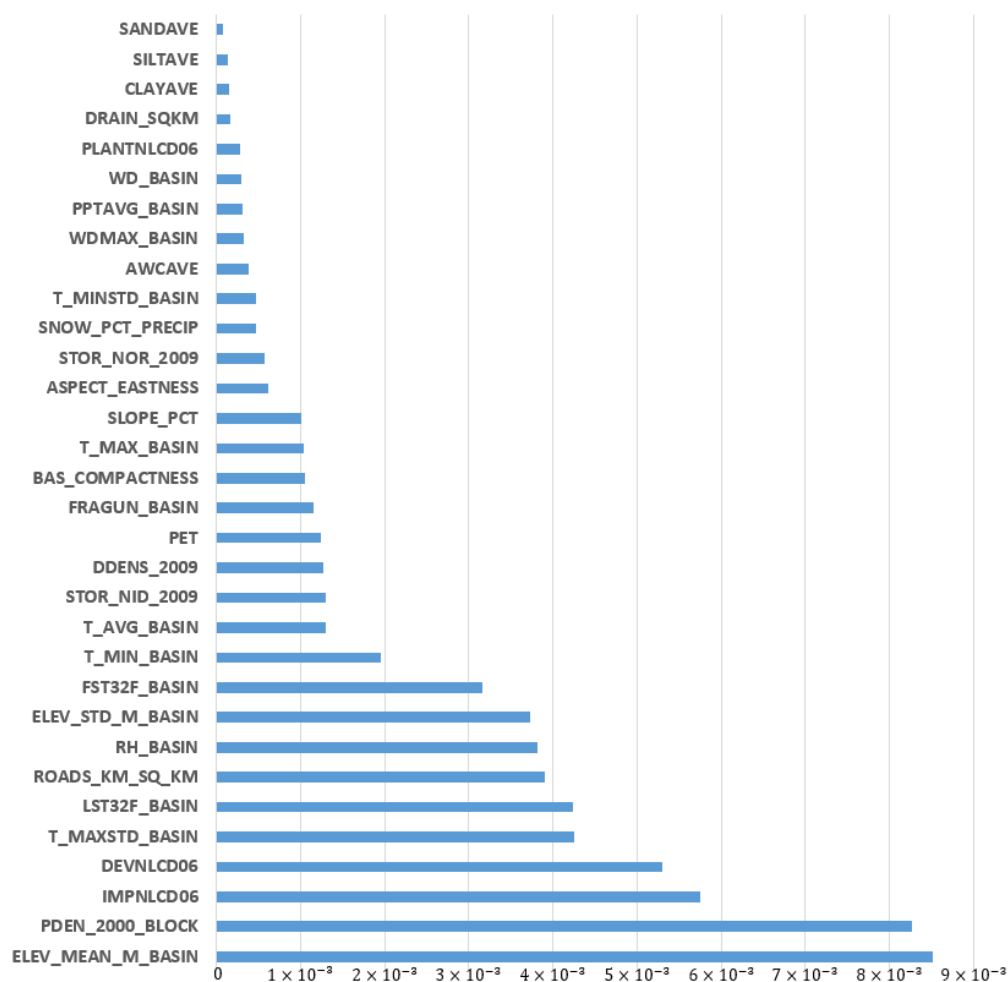


**Figure 5.** Predictor variables of the shape parameters (with associated importance) for Interior Highland. Description of predictor variables is provided in Appendix A.

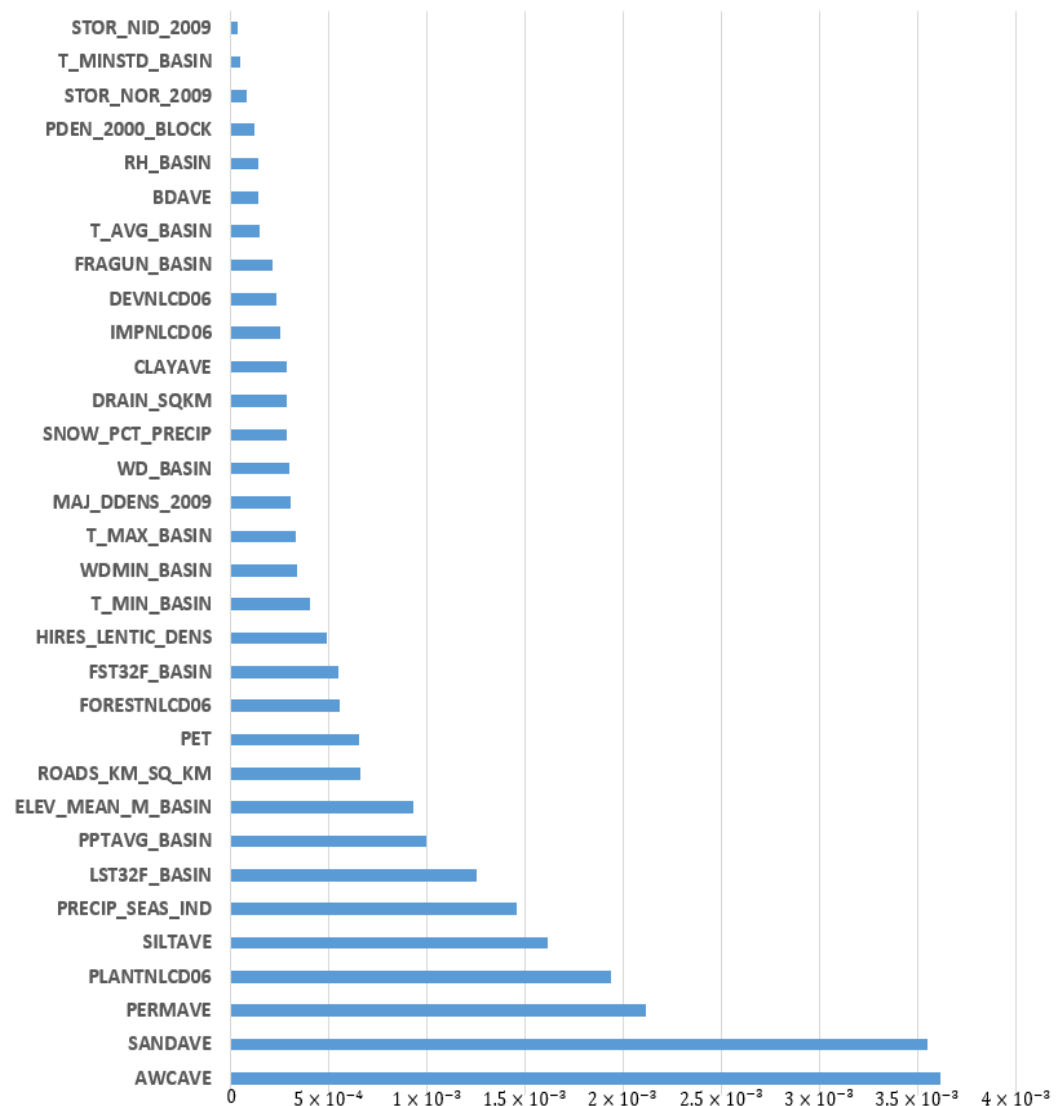
Considering the importance levels, certain climate variables are found to be most important for Appalachian Highland; 9 out of 10 variables (ranking between 1 to 10) are climate variables (Table 2). For Interior Highlands, certain climate, dam, and soil variables are found to be most important (Table 2). Finding dam properties as the important predictor variables (third, sixth, eighth) (Table 2) might be noted here specifically as it suggests the apparent importance of human interference within the watershed. The importance of human interference is also apparent within Central Lowland with the rankings of certain variables: population density (second), percentage of impervious area (third), percentage of developed area in the watershed (fourth), and road density (seventh) (Table 2). Certain topography, population infrastructure, and climate variables are also found to be most important for Central Lowland (Table 2). For Coastal Plain, certain soil and climate variables are found to be the most important (Table 2). The findings, considering all sites together, are quite similar to the Appalachian Highland; 9 out of 10 important variables are found to be climate type (Table 2).

**Table 2.** Types of most important (ranking 1 to 10) predictor variables across different physiographic regions. Description of predictor variables is provided in Appendix A.

Ranking	Appalachian Highland	Interior Highlands	Coastal Plain	Central Lowland	All Sites
1	WDMIN_BASIN	FST32F_BASIN	AWCAVE	ELEV_MEAN_M_BASIN	WD_BASIN
2	WD_BASIN	LST32F_BASIN	SANDAVE	PDEN_2000_BLOCK	LST32F_BASIN
3	WDMAX_BASIN	DDENS_2009	PERMAVE	IMPNLCD06	WDMAX_BASIN
4	SLOPE_PCT	BDAVE	PLANTNLCD06	DEVNLCD06	PPTAVG_BASIN
5	T_MIN_BASIN	T_MIN_BASIN	SILTAVE	T_MAXSTD_BASIN	FST32F_BASIN
6	T_AVG_BASIN	STOR_NID_2009	PRECIP_SEAS_IND	LST32F_BASIN	ELEV_MEAN_M_BASIN
7	LST32F_BASIN	T_AVG_BASIN	LST32F_BASIN	ROADS_KM_SQ_KM	RH_BASIN
8	T_MAX_BASIN	STOR_NOR_2009	PPTAVG_BASIN	RH_BASIN	T_MIN_BASIN
9	SNOW_PCT_PRECIP	PET	ELEV_MEAN_M_BASIN	ELEV_STD_M_BASIN	T_AVG_BASIN
10	PET	CLAYAVE	ROADS_KM_SQ_KM	FST32F_BASIN	PET
		Climate Soils Population Infrastructure		Watershed Dams Topography	



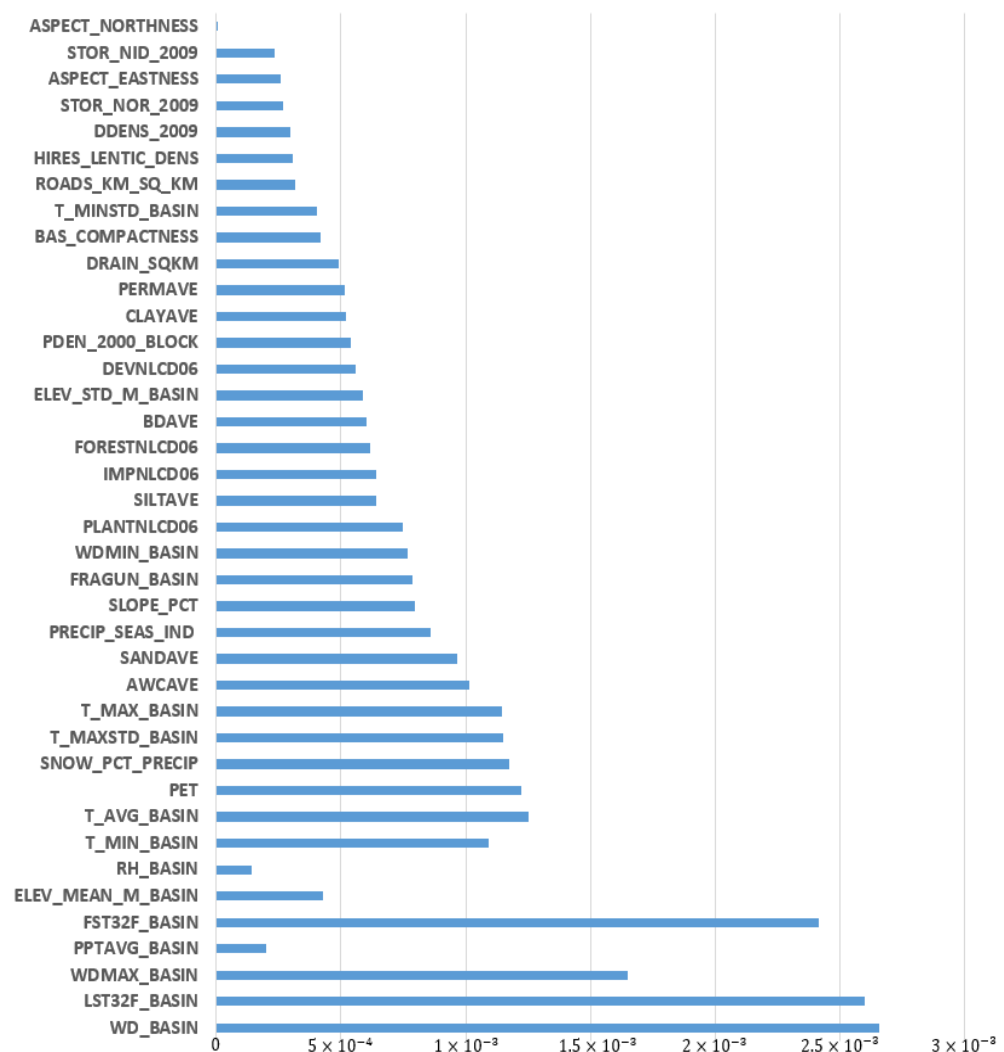
**Figure 6.** Predictor variables of the shape parameters (with associated importance) for Central Lowland. Description of predictor variables is provided in Appendix A.



**Figure 7.** Predictor variables of the shape parameters (with associated importance) for Coastal Plain. Description of predictor variables is provided in Appendix A.

Among 40 distinct variables, 10 predictor variables are found common across all regions (Table 3); climate (6), soil (2), population infrastructure (1), and watershed (1). Each of the common variables is found to be important for one or more than one physiographic region (Table 3), except for the variables ‘T\_MIN\_STD\_BASIN’, and ‘FRAGUN\_BASIN’. Only one common variable, ‘LST32F\_BASIN’, is found to be important (ranking between 1 to 10) for all regions (Table 3).

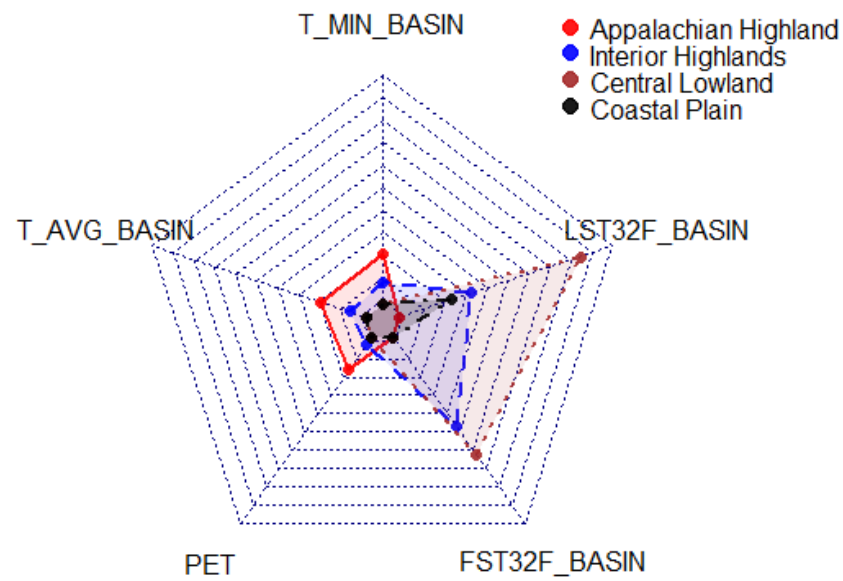
When comparing important (ranking between 1 to 10) variables common to more than one region, comparatively higher importance levels are obtained for certain regions (Figures 9–13). For instance, for common soil variables (ranking 1 to 10), relatively higher important levels are retrieved for Coastal Plain except for ‘BDAVE’ (Figure 10). For common topography variables (ranking 1 to 10), comparatively higher importance levels are retrieved for Central Lowland (Figure 11). On the other hand, for common topography and watershed variables (ranking 1 to 10), Interior Highlands have the least importance levels (Figures 11 and 12). Higher importance levels imply higher predictability of the flood behavior for Central Lowland and Coastal Plain while considering certain variables as predictors. Since the system has higher predictability, accurate forecasts based on current observations (predictor variables with higher importance levels) would be comparatively easier.



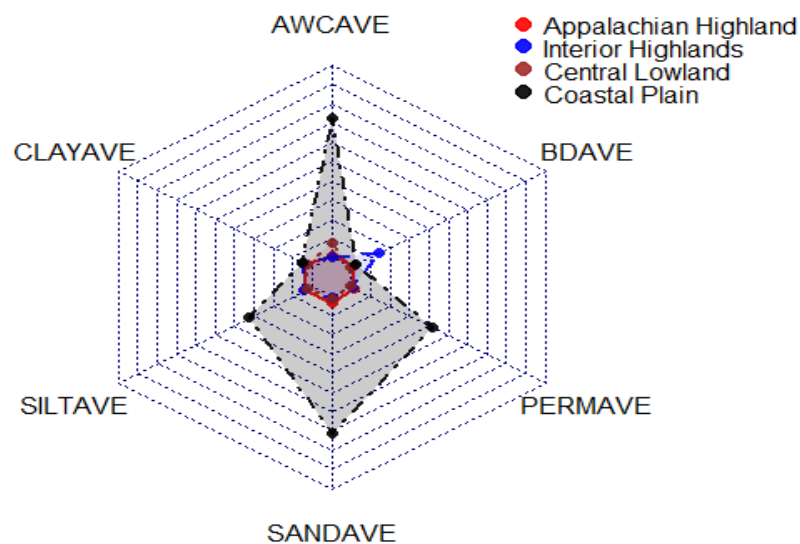
**Figure 8.** Predictor variables of the shape parameters (with associated importance) for all sites across different physiographic regions. Description of predictor variables is provided in Appendix A.

**Table 3.** Common predictor variables across all study regions. Description of predictor variables is provided in Appendix A.

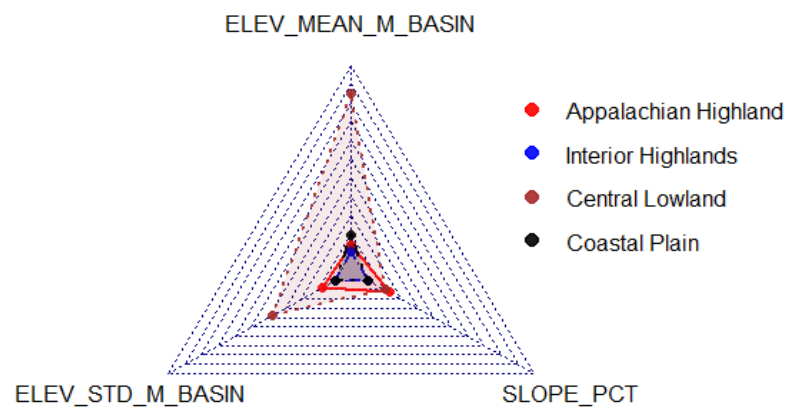
Variables	Variable Type	Important (Ranking between 1 to 10) Regions
LST32F_BASIN	Climate	All regions
T_MIN_BASIN	Climate	Appalachian Highland, Interior Highlands
T_AVG_BASIN	Climate	Appalachian Highland, Interior Highlands
SNOW_PCT_PRECIP	Climate	Appalachian Highland
FST32F_BASIN	Climate	Interior Highlands
T_MINSTD_BASIN	Climate	NA
CLAYAVE	Soils	Interior Highlands
SANDAVE	Soils	Coastal Plain
ROADS_KM_SQ_KM	Population Infrastructure	Central Lowland, Coastal Plain
FRAGUN_BASIN	Watershed	NA



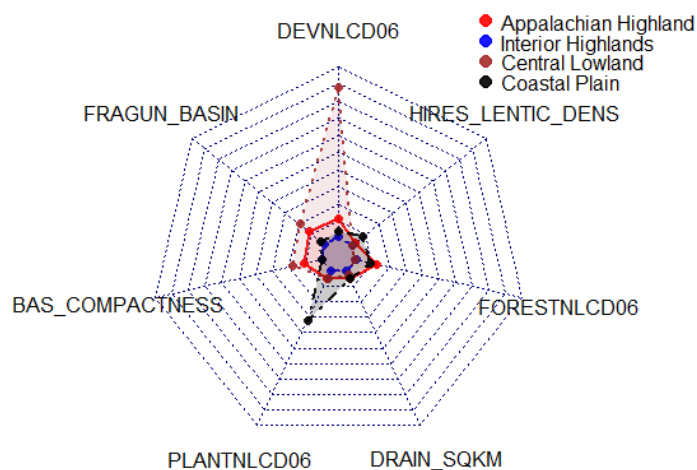
**Figure 9.** Comparison of importance levels for climate variables (ranking 1 to 10) common to more than one region.



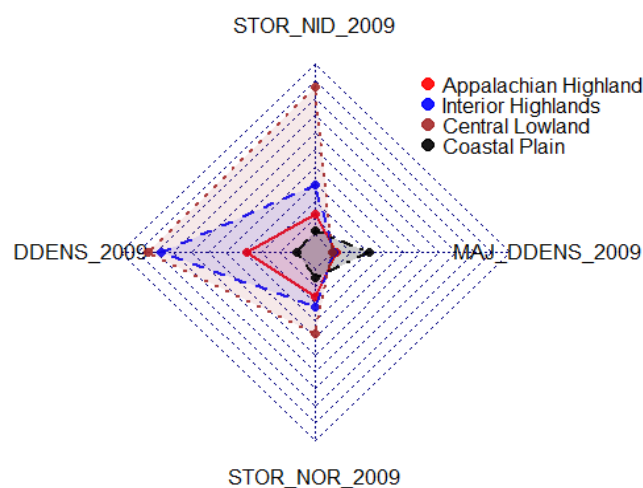
**Figure 10.** Comparison of importance levels for soil variables (ranking 1 to 10) common to more than one region.



**Figure 11.** Comparison of importance levels for topography variables (ranking 1 to 10) common to more than one region.



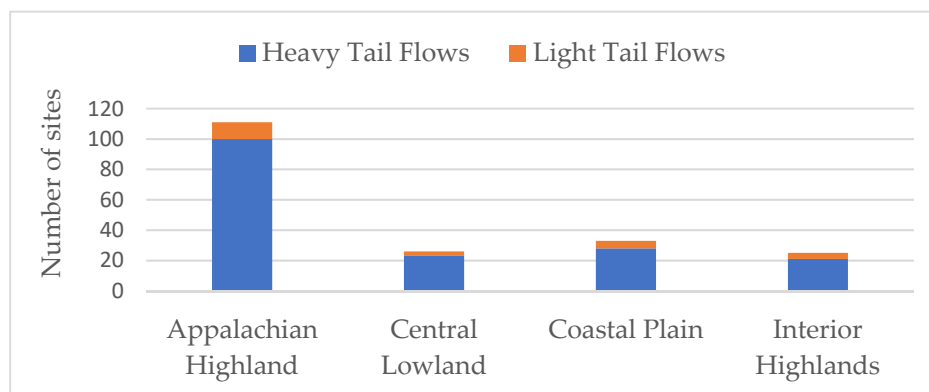
**Figure 12.** Comparison of importance levels for watershed variables (ranking 1 to 10) common to more than one region.



**Figure 13.** Comparison of importance levels for dam variables (ranking 1 to 10) common to more than one region.

### 3.3. Heavy Tail Flows

Most of the USGS stations (172 out of 195) associated with the collapse sites are found to be exhibiting a ‘heavy tail’ distribution (Figure 14, Table 4) for peak flows, that is, the shape parameter values of the GEV distributions are greater than zero.



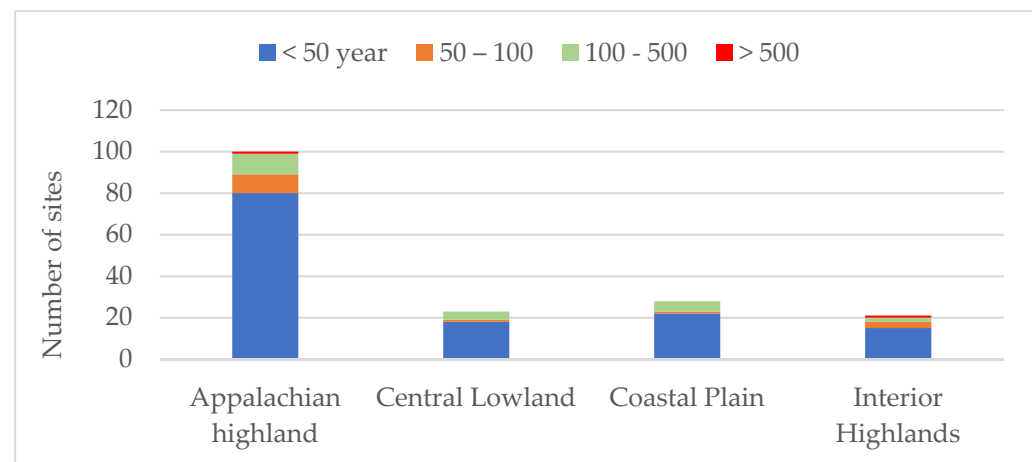
**Figure 14.** Number of sites with heavy and light tail flows of the bridge collapse sites across different regions.



**Table 4.** Return periods and associated information for heavy tail flows of the bridge collapse sites across different regions.

Physiographic Region	# of Sites with Heavy Tail Flows (Total Sites)	# of Sites with Different Return Periods (Years)				Minimum/Maximum Return Periods (Years)	Mean/Median	Kurtosis
		<50 Years	50–100 Years	100–500 Years	>500 Years			
Appalachian Highland	100 (111)	80	9	10	1	3/897	45.58/13.57	47.0
Central Lowland	23 (26)	18	1	4		2/465	55.37/10.12	10.41
Coastal Plain	28 (33)	22	1	5		2/213	45.38/14.81	2.05
Interior Highlands	21 (25)	15	3	2	1	1/627	75/18.41	9.30

Heavy tail flows are expected to be with higher return periods. However, for most of the sites, heavy tail flows exhibit lower return periods (Figure 15, Table 4). The minimum return periods range from 1 to 3 years considering all regions (Table 4). In fact, the median values for the return periods are found to be less than 50 years for each of the physiographic regions studied (Table 4). Within these regions, heavy tail flows occur more frequently than expected. There are also a few collapse sites within each region, for which no heavy tail flow occurred yet. Such results imply that at these specific sites, bridges collapsed at flows with lower return periods, particularly, less than 50 years. On the other hand, the maximum return periods for the heavy tail flows range from 213 to about 897 years (Table 4). Comparing all regions, the likelihood of having extremes with higher return periods (outliers) is greater for Appalachian Highland as implied by the larger kurtosis value of 47 (Table 4). The calculated return periods for the heavy tail flows (for 205 USGS stations) can be retrieved from <https://github.com/fahmidah/Shape-Parameter-Bridge-Collapse-Sites/tree/main/Return%20Periods> (accessed on 10 March 2022).

**Figure 15.** Return periods for heavy tail flows of the bridge collapse sites across different regions.

#### 4. Discussion

The key findings are discussed in detail here.

1. *The normal distribution of shape parameters does not imply climate variables as the most important predictor variables.* Some studies suggest that the shape parameter depends on climate indexes, while other attributes of the catchments are less important [46,53–55]. With such attributes, the GEV shape parameter is most preferably modeled by a normal distribution with a common mean across all sites [46,53,56–58]. However, it is also claimed that this may be a result of an insufficient summary of the catchment attributes by the implemented indexes [59]. For the current study, the retrieved values of shape parameters do agree with recent studies [59–62] in that the values follow a fairly normal

distribution within each physiographic region. However, for the study sites, the distinct association with climate variables is not evident except for Appalachian Highland.

2. *There is a variation in the types of important predictor variables across all study regions.* Although there is an over-representation of climate variables, important predictors (ranking 1 to 10) include more than one type across all regions: climate and topography for Appalachian Highland; climate, dam, and soil for Interior Highlands; topography, population, and climate for Central Lowland; and soil and climate for Coastal Plain. In fact, the median value of the shape parameters obtained for each region is very close to the median value of 0.19, as obtained for the sites with hydrological heterogeneity [46]. Such a result implies that the flood behavior (or extremeness) within each specific region is similar to the flood behavior (or extremeness) within the region of hydrological heterogeneity.
3. *Predictor variables associated with human interference are important.* For all regions, ‘population infrastructure’ and ‘dams’ variables are obtained as the predictor variables. For Interior Highlands, Coastal Plain, and Central Lowland, the importance of human interference is apparent with higher importance levels (ranking between 1 to 10) as compared to other predictor variables. It is well-established that human interference, particularly the presence of dams, can cause and/or exacerbate flood events. Dam break-generated floods can be much more catastrophic than typical riverine floods since dam break generates significant impulsive loads, which can exceed the drag of typical riverine floods by up to 270% [63]. Dam break-generated floods also generate different patterns of vertical forces and moments, which would lead to different probabilities of collapse and different failure modes as compared to riverine floods [63]. Emergency releases of water down spillways can also cause far greater and more sudden floods downstream as compared to the natural riverine floods [64]. The loss of natural sponges for floodwaters, such as deforestation, removal of scrub, and/or drained-out wetlands, caused by building dams, can result in a completely different hydrological regime, and can also increase the risk of extreme floods [64]. In the current study, the peak flows of 29 sites are affected by possible dam breaks as noted by USGS. The reconnaissance studies throughout the mid-west also revealed that the widespread human disturbance together with easily erodible soils has produced thousands of miles of highly unstable streams [27,65]. Therefore, it can be argued that human interference, for instance, through dams, is an important parameter at a watershed and/or regional scale, given the unique flood generation mechanism within such a context.
4. *Return periods of heavy tail flows vary from very low to very high.* For the collapse sites studied, there exists a wide variation of return periods for the heavy tail flows. However, there is a persistence of lower return periods, as low as 1 to 3 years, which implies that the heavy tail flows occur more frequently than expected at the bridge collapse sites. Such a finding is in accordance with a recent study of the bridge collapse sites in the U.S. where the prevalence of collapse-inducing floods with lower return periods has become evident [28]. In fact, certain physiographic settings do possess plausible conditions for which flows with lower return periods can induce collapse [66,67]. For instance, within physiographic regions, namely, Central Lowland, Intermontane, Great Plains, Appalachian Highland, and Coastal Plain, unique conditions such as irregular channel shifting, high erosion at bed and bank, debris jam, backwater from bays, and/or intensive human interventions are exhibited [27,65]. Regions with such intrinsic risk conditions would not have the resiliency to flow with lower return periods as would have been expected. Extreme flows with higher return periods would have catastrophic effects within such settings. The study results do also indicate the occurrence of extreme flows with return periods as high as 213 to 897 years at a few sites within each region studied. Such findings instigate revisiting the guidelines specifying the use of a “100-year flood” (flood with an annual probability of exceedance of 1%) for modern interstate bridges receiving federal funding [68]. Typically, for design

purposes, i.e., calculation of expected scour and hydraulic loads, engineers would consider the magnitude of a single design flood. One single, characteristic flood, such as a 100-year flood, is also frequently used in collapse risk studies [1]. Given that the collapse can be induced by heavy tail flows with both low and high return periods, in relation to certain physiographic settings, bridge design should be modeled over a range of possible collapse-inducing flows. Risk studies should also consider all plausible collapse-inducing floods in relation to certain regional conditions.

### *Summary*

Whereas climate, soil, topography, and watershed attributes are common in flood hazard indexing, recent studies suggest that the shape parameter of the flood distribution mostly depends on climate indexes, while other attributes of the catchments are less important [46,53–55]. The association of the shape parameter mostly with climate variables implies that certain climate variables are the most important predictors of flood behavior. The current study results are novel in that they suggest that other variables, including soil, topography, watershed, and human interference, are more important than climate variables in certain physiographic settings, at least for the collapse sites considered. Widespread human interference at a regional scale is also evident in the bridge collapse sites, particularly within Interior Highlands, Coastal Plain, and Central Lowland. Another key finding is that the collapse sites within each physiographic region exhibit a similar extremeness of flood as within the region of hydrologic heterogeneity. The extremeness of floods is also implied by the frequent occurrence of heavy tail flows within each region. All of these key findings, within a regional context, should be considered in failure modes and effect analysis for assessing collapse risk at a local scale, which could be an area for further research. Here, it should be noted that failure modes and effects analysis for a large dataset of bridge collapse sites are not feasible because of the unavailability of the required information. The NYDOT failure database does not include any information about the failed component or the failure mode for the collapse events. Such information might be available at the state-based DOT (Department of Transportation); however, they are not available to the public. Nonetheless, if an association can be established among regional variables, bridge characteristics, and local conditions in relation to specific failure modes, it can lead to a robust risk assessment framework for the bridge structures. In addition, hydrologic variables, such as the 40 variables selected in the current study, are typically analyzed at a regional and/or watershed scale, whereas hydraulic parameters, such as flood level, bridge type, etc., are considered at a more refined local scale. Some parameters, such as debris jams, are investigated both at the regional and local scales. Integration of these two spatial scales requires extensive research and, therefore, could be a topic of future studies.

### **5. Conclusions**

In the study, analysis of shape parameters, retrieved from the GEV distribution of annual peak flows, are performed for 205 USGS watersheds (or stations) associated with 297 bridge collapse sites across different physiographic regions. Annual peak flows are also analyzed in relation to the heavy tail flows and their return periods. Certain unique findings become apparent for all of the collapse sites in relation to each physiographic region (Appalachian Highland, Central Lowland, Interior Highlands, and Coastal Plain): (a) variation of the most important predictor variables, (b) apparent importance of human interference, (c) flood extremeness similar to the region with hydrologic heterogeneity, and (d) frequent occurrence of extreme peak flows. Such findings necessitate revisiting bridge design and bridge collapse risk study in relation to not only the magnitude and frequency of floods but also the behavior of the floods as stemming from specific physiographic settings.

**Author Contributions:** Conceptualization, F.A., H.T. and G.P.; methodology, F.A., H.T. and G.P.; formal analysis, F.A., H.T. and G.P.; data curation, F.A.; writing—original draft preparation, F.A.; writing—review and editing, H.T. and G.P.; funding acquisition, F.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** The material contained in this document is based upon work supported by a National Aeronautics and Space Administration (NASA) grant or cooperative agreement. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of NASA. This work was supported through a NASA grant awarded to the Illinois/NASA Space Grant Consortium (Grant Number: 2786531F).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Dataset including analysis used to retrieve the predictor variables can be found at <https://github.com/fahmidah/Shape-Parameter-Bridge-Collapse-Sites> (accessed on 10 February 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Description of predictor variables. All information is retrieved from [https://water.usgs.gov/lookup/getspatial?gagesII\\_Sept2011](https://water.usgs.gov/lookup/getspatial?gagesII_Sept2011) (accessed on 10 February 2022).

Variable Name	Description	Units	Time Period	Extent/Processing Method	Additional Note on Processing or Dataset	Scale of Source Data	Source of Data for Details [Citation]
<i>Climate Variables</i>							
WD_BASIN	Watershed average of annual number of days (days) of measurable precipitation, derived from 30 years of record (1961–1990), 2 km PRISM.	days	1961–1990	WATERSHED/ Grid	N/A	2 km grid	PRISM [69,70]
LST32F_BASIN	Watershed average of mean day of the year of last freeze, derived from 30 years of record (1961–1990), 2 km PRISM. For example, value of 100 is the 100th day of the year (10 April).	day of year	1961–1990	WATERSHED/ Grid	Values of 0 here indicate there is no freeze.	2 km grid	PRISM [69,70]
WDMAX_BASIN	Watershed average of monthly maximum number of days (days) of measurable precipitation, derived from 30 years of record (1961–1990), 2 km PRISM.	days	1961–1990	WATERSHED/ Grid	N/A	2 km grid	PRISM [69,70]
PPTAVG_BASIN	Mean annual precip (cm) for the watershed, from 800 m PRISM data. Period of record of 30 years, 1971–2000.	cm	1971–2000	WATERSHED/ Grid	N/A	800 m grid	PRISM [69,70]
FST32F_BASIN	Watershed average of mean day of the year of first freeze, derived from 30 years of record (1961–1990), 2 km PRISM. For example, value of 300 is the 300th day of the year (27 October).	day of year	1961–1990	WATERSHED/ Grid	Values > 365 mean that the first freeze occurs after 31 December. It is conducted this way so that (first freeze–last freeze) should always give a positive number; high numbers for an area with a long frost-free period (e.g., Florida), low numbers for an area with a short frost-free period.	2 km grid	PRISM [69,70]

Table A1. Cont.

Variable Name	Description	Units	Time Period	Extent/Processing Method	Additional Note on Processing or Dataset	Scale of Source Data	Source of Data for Details [Citation]
T_MIN_BASIN	Watershed average of minimum monthly air temperature (degrees C) from 800 m PRISM, derived from 30 years of record (1971–2000).	degrees C	1971–2000	WATERSHED/ Grid	N/A	800 m grid	PRISM [69,70]
T_AVG_BASIN	Average annual air temperature for the watershed, degrees C, from 2 km PRISM data. Period of record of 30 years, 1971–2000.	degrees C	1971–2000 Grid	WATERSHED/ Grid	N/A	2 km grid	PRISM [69,70]
PET	Mean annual potential evapotranspiration (PET), estimated using the Hamon (1961) equation.	mm/year	1961–1990	WATERSHED/ Grid	The Hamon equation computes PET from mean monthly air temperature and latitude. The mean monthly air temperature values were computed from 1961–1990 “normals” (30-year average values) of mean monthly minimum and maximum air temperature generated from a PRISM spatial/statistical model.	1 km grid	PRISM [69,70]
T_MAXSTD_BASIN	Standard deviation of maximum monthly air temperature (degrees C) from 800 m PRISM, derived from 30 years of record (1971–2000).	degrees C	1971–2000	WATERSHED/ Grid	N/A	800 m grid	PRISM [69,70]
T_MAX_BASIN	Watershed average of maximum monthly air temperature (degrees C) from 800 m PRISM, derived from 30 years of record (1971–2000).	degrees C	1971–2000	WATERSHED/ Grid	N/A	800 m grid	PRISM [69,70]
PRECIP_SEAS_IND	Precipitation seasonality index (Markham, 1970; Dingman, 2002). Index of how much annual precipitation falls seasonally (high values) or spread out over the year (low values). Based on monthly precip values from 30 year (1971–2000) PRISM. Range is 0 (precip spread out exactly evenly in each month) to 1 (all precip falls in a single month).	unitless	1971–2000	WATERSHED/ Grid	This is calculated as described in Dingman (2002), pages 140–145. Each month is treated as a vector, and the precip in that month as the “length” of that vector.	800 m grid	PRISM [69,70]
WDMIN_BASIN	Watershed average of monthly minimum number of days (days) of measurable precipitation, derived from 30 years of record (1961–1990), 2 km PRISM.	days	1961–1990	WATERSHED/ Grid	N/A	2 km grid	PRISM [69,70]
T_MINSTD_BASIN	Standard deviation of minimum monthly air temperature (degrees C) from 800 m PRISM, derived from 30 years of record (1971–2000).	degrees C	1971–2000	WATERSHED/ Grid	N/A	800 m grid	PRISM [69,70]

Table A1. Cont.

Variable Name	Description	Units	Time Period	Extent/Processing Method	Additional Note on Processing or Dataset	Scale of Source Data	Source of Data for Details [Citation]
RH_BASIN	Watershed average relative humidity (percent), from 2 km PRISM, derived from 30 years of record (1961–1990).	percent	1961–1990	WATERSHED/ Grid	N/A	2 km grid	PRISM [69,70]
SNOW_PCT_PRECIP	Snow percent of total precipitation estimate, mean for period 1901–2000. From McCabe and Wolock (submitted, 2008), 1 km Grid.	percent	1901–2000	WATERSHED/ Grid	N/A	1 km grid	David Wolock [71]
<i>Topography Variables</i>							
ELEV_MEAN_M_BASIN	Mean watershed elevation (meters) from 100 m National Elevation Dataset	meters	N/A	WATERSHED/ Grid	Original source of elevation data were 30 m tiles of USGS NED data assembled by Roland Viger and Curtis Price of the USGS in the early 2000 period.	100 m grid (resampled from 30 m)	USGS [72]
ELEV_STD_M_BASIN	Standard deviation of elevation (meters) across the watershed from 100 m National Elevation Dataset	meters	N/A	WATERSHED/ Grid	Original source of elevation data were 30 m tiles of USGS NED data assembled by Roland Viger and Curtis Price of the USGS in the early 2000 period.	100 m grid (resampled from 30 m)	USGS [72]
ASPECT_NORTHNESS	Aspect “northness”. Ranges from –1 to 1. Value of 1 means watershed is facing/draining due north, value of –1 means watershed is facing/draining due south.	unitless	N/A	WATERSHED/ Grid	Original source of elevation data were 30 m tiles of USGS NED data assembled by Roland Viger and Curtis Price of the USGS in the early 2000 period.	100 m grid (resampled from 30 m)	USGS [72]
ASPECT_EASTNESS	Aspect “eastness”. Ranges from –1 to 1. Value of 1 means watershed is facing/draining due east, value of –1 means watershed is facing/draining due West.	unitless	N/A	WATERSHED/ Grid	Original source of elevation data were 30 m tiles of USGS NED data assembled by Roland Viger and Curtis Price of the USGS in the early 2000 period.	100 m grid (resampled from 30 m)	USGS [72]
SLOPE_PCT	Mean watershed slope, percent. Derived from 100 m resolution National Elevation Dataset, so slope values may differ from those calculated from data of other resolutions.	percent	N/A	WATERSHED/ Grid	Original source of elevation data were 30 m tiles of USGS NED data assembled by Roland Viger and Curtis Price of the USGS in the early 2000 period.	100 m grid (resampled from 30 m)	USGS [72]
<i>Soil Variables</i>							
AWCAVE	Average value for the range of available water capacity for the soil layer or horizon (inches of water per inches of soil depth).	unit less (fraction inches of water per inches of soil depth)	N/A	WATERSHED/ Grid	Each STATSGO mapping unit contains multiple soil components, and each component can have multiple soil layers. The average value for each mapping unit is a weighted average based on soil layer thickness and component area.	The base data layer is built on the STATSGO mapping units, the average size of which is about 750 sq km	USGS NAWQA. Data are based on STATSGO, but as aggregated by Wolock, 1997 [73]



Table A1. Cont.

Variable Name	Description	Units	Time Period	Extent/Processing Method	Additional Note on Processing or Dataset	Scale of Source Data	Source of Data for Details [Citation]
SANDAVE	Average value of sand content (percentage)	percent	N/A	WATERSHED/ Grid	Each STATSGO mapping unit contains multiple soil components, and each component can have multiple soil layers. The average value for each mapping unit is a weighted average based on soil layer thickness and component area.	The base data layer is built on the STATSGO mapping units, the average size of which is about 750 sq km	USGS NAWQA. Data are based on STATSGO, but as aggregated by Wolock, 1997 [73]
SILTAVE	Average value of silt content (percentage)	percent	N/A	WATERSHED/ Grid	Each STATSGO mapping unit contains multiple soil components, and each component can have multiple soil layers. The average value for each mapping unit is a weighted average based on soil layer thickness and component area.	The base data layer is built on the STATSGO mapping units, the average size of which is about 750 sq km	USGS NAWQA. Data are based on STATSGO, but as aggregated by Wolock, 1997 [73]
CLAYAVE	Average value of clay content (percentage)	percent	N/A	WATERSHED/ Grid	Each STATSGO mapping unit contains multiple soil components, and each component can have multiple soil layers. The average value for each mapping unit is a weighted average based on soil layer thickness and component area.	The base data layer is built on the STATSGO mapping units, the average size of which is about 750 sq km	USGS NAWQA. Data are based on STATSGO, but as aggregated by Wolock, 1997 [73]
BDAVE	Average value of bulk density (grams per cubic centimeter)	grams per cubic centimeter	N/A	WATERSHED/ Grid	Each STATSGO mapping unit contains multiple soil components, and each component can have multiple soil layers. The average value for each mapping unit is a weighted average based on soil layer thickness and component area.	The base data layer is built on the STATSGO mapping units, the average size of which is about 750 sq km	USGS NAWQA. Data are based on STATSGO, but as aggregated by Wolock, 1997 [73]
PERMAVE	Average permeability (inches/hour)	inches/hour	N/A	WATERSHED/ Grid	Each STATSGO mapping unit contains multiple soil components, and each component can have multiple soil layers. The average value for each mapping unit is a weighted average based on soil layer thickness and component area.	The base data layer is built on the STATSGO mapping units, the average size of which is about 750 sq km	USGS NAWQA. Data are based on STATSGO, but as aggregated by Wolock, 1997 [73]
Watershed variables							
FRAGUN_BASIN	Fragmentation Index of “undeveloped” land in the watershed. High numbers = more disturbance by development and fragmentation; a very pristine basin with a lot of contiguous undeveloped land cover would have a low number.	unitless	2000–2002	WATERSHED/ Grid	Based on Ritters and others (2000) method, using $3 \times 3$ processing window. Number given here is: (100-percent “interior” pixels of undeveloped land (Riitters class 1)). Definition of “undeveloped” land = all land which is not urban nor agriculture, from NLCD01 data.	30 m grid	USGS NLCD01 [74]

Table A1. Cont.

Variable Name	Description	Units	Time Period	Extent/Processing Method	Additional Note on Processing or Dataset	Scale of Source Data	Source of Data for Details [Citation]
PLANTNLCD06	Watershed percent “planted/cultivated” (agriculture), 2006 era (2001 for AK-HI-PR). Sum of classes 81 and 82.	percent	2006 (2001 for AK-HI-PR)	WATERSHED/ Grid	These are sums of Level II classes. Data are from NLCD01 for Alaska-Hawaii-PR.	30 m grid	USGS NLCD06 (NLCD01 for AK-HI-PR) [74]
FORESTNLCD06	Watershed percent “forest”, 2006 era (2001 for AK-HI-PR). Sum of classes 41, 42, and 43.	percent	2006 (2001 for AK-HI-PR)	WATERSHED/ Grid	These are sums of Level II classes. Data are from NLCD01 for Alaska-Hawaii-PR.	30 m grid	USGS NLCD06 (NLCD01 for AK-HI-PR) [74]
DEVNLCD06	Watershed percent “developed” (urban), 2006 era (2001 for AK-HI-PR). Sum of classes 21, 22, 23, and 24.	percent	2006 (2001 for AK-HI-PR)	WATERSHED/ Grid	These are sums of Level II classes. Data are from NLCD01 for Alaska-Hawaii-PR.	30 m grid	USGS NLCD06 (NLCD01 for AK-HI-PR) [74]
DRAIN_SQKM	Watershed drainage area, sq km, as delineated in our basin boundary.	square km	N/A	WATERSHED/Vector polygon area	See report for more info on basin delineation process.	1:24,000–1:100,000	USGS NAWQA. Most are based on Mike Wicczorek basin boundaries, derived from 30-m NHDPlus data [75,76]
BAS _COMPACTNESS	Watershed compactness ratio, $=\text{area}/\text{perimeter}^2 \times 100$ ; higher number = more compact shape.	unitless	N/A	WATERSHED/Vector polygon area	Basin grids were reconverted to polygons for calculation, because original perimeters could be of inconsistent detail, depending on method used to delineate basins (some were from NAWQA).	1:24,000–1:100,000	USGS NWIS and NAWQA [75,76]
HIRES_LENTIC _DENS	Density (#/sq km) of Lakes/Ponds + Reservoir water bodies from NHD Hi-Resolution (1:24k) data.	number/sq km	N/A	WATERSHED/Vector polygon area	NHD Hi-Res data assembled September 2011.	1:24,000	National Hydrography Dataset
Population Infrastructure							
IMPNLCD06	Watershed percent impervious surfaces from 30-m resolution NLCD06 data.	percent	2006	WATERSHED/ Grid	N/A	30 m grid	USGS NLCD06 [74]
PDEN_2000 _BLOCK	Population density in the watershed, persons per sq km, from 2000 Census block data regridded to 100 m. This variable is maintained to support models built from original GAGES dataset.	persons/sq km	2000	WATERSHED/ Grid	Block data is the most detailed Census scale. Original SILVIS data in polygon form (all U.S. Census blocks), subsequently gridded at 100 m.	Census block	SILVIS Lab [77]
ROADS_KM _SQ_KM	Road density, km of roads per watershed sq km, from Census 2000 TIGER roads.	km/sq km	2000	WATERSHED/Vector line	N/A	1:100,000	Census, repackaged by GeoLytics [78]

Table A1. Cont.

Variable Name	Description	Units	Time Period	Extent/Processing Method	Additional Note on Processing or Dataset	Scale of Source Data	Source of Data for Details [Citation]
<i>Dams</i>							
DDENS_2009	Dam density; number per 100 km sq.	number of dams/100 sq km	All dams' data are based on an enhanced version of 2009 National Inventory of Dams (NID)	WATERSHED/Vector point	We took the 2009 NID (about 83,000) and merged that with a subset of about 59,000 dams whose locations had been verified and, in some cases, adjusted by the USGS National Map/NHD program, which had located those dams on the NHDPlus network. The result is that the locations of the 83,000 NID dams are in many cases improved/corrected from their original location. We also eliminated several hundred which were duplicates or near-duplicates.	1:100,000	NID 2009 (83,000 records) merged with verified/corrected locations for 59,000 from USGS National Map/NHD work [79]
STOR_NOR_2009	Dam storage in watershed ("NOR-MAL_STORAGE"); megaliters total storage per sq km (1 megaliters = 1,000,000 L = 1000 cubic meters).	megaliters/sq km	All dams' data based on an enhanced version of 2009 National Inventory of Dams (NID)	WATERSHED/Vector point	We took the 2009 NID (about 83,000) and merged that with a subset of about 59,000 dams whose locations had been verified and, in some cases, adjusted by the USGS National Map/NHD program, which had located those dams on the NHDPlus network. The result is that the locations of the 83,000 NID dams are in many cases improved/corrected from their original location. We also eliminated several hundred which were duplicates or near-duplicates.	1:100,000	NID 2009 (83,000 records) merged with verified/corrected locations for 59,000 from USGS National Map/NHD work [79]
STOR_NID_2009	Dam storage in watershed ("NID_STORAGE"); megaliters total storage per sq km (1 megaliters = 1,000,000 L = 1000 cubic meters). See note to the right as well.	megaliters/sq km	All dams data based on an enhanced version of 2009 National Inventory of Dams (NID)	WATERSHED/Vector point	NID storage in the National Inventory of Dams is generally (but not always, because of apparent errors in the NID) equal to the maximum storage of the reservoir. This was the default we used because of inconsistencies in the Normal storage (STOR_NOR_2009); i.e., dams which obviously had storage—huge lakes—but had Normal storage of 0.	1:100,000	NID 2009 (83,000 records) merged with verified/corrected locations for 59,000 from USGS National Map/NHD work [79]
MAJ_DDENS_2009	Major dam density; number per 100 km sq.	number of major dams/100 sq km	All dams' data based on an enhanced version of 2009 National Inventory of Dams (NID)	WATERSHED/Vector point	Of the 83,262 dams in our version of the 2009 NID dataset, 8576 met this definition, i.e., roughly 10%.	1:100,000	NID 2009 (83,000 records) merged with verified/corrected locations for 59,000 from USGS National Map/NHD work [79]

## Appendix B

**Table A2.** Predictor variables retrieved for Central Lowland.

Variable Type	Variable	Importance Level
Climate	T_MAXSTD_BASIN	$4.25 \times 10^{-3}$
Climate	LST32F_BASIN	$4.24 \times 10^{-3}$
Climate	FST32F_BASIN	$3.16 \times 10^{-3}$
Climate	T_MIN_BASIN	$1.96 \times 10^{-3}$
Climate	T_AVG_BASIN	$1.30 \times 10^{-3}$
Climate	SNOW_PCT_PRECIP	$4.70 \times 10^{-4}$
Climate	T_MINSTD_BASIN	$4.62 \times 10^{-4}$
Climate	WDMAX_BASIN	$3.19 \times 10^{-4}$
Climate	PPTAVG_BASIN	$3.06 \times 10^{-4}$
Climate	WD_BASIN	$2.93 \times 10^{-4}$
Climate	RH_BASIN	$3.82 \times 10^{-3}$
Climate	PET	$1.24 \times 10^{-3}$
Climate	T_MAX_BASIN	$1.03 \times 10^{-3}$
Topography	ELEV_MEAN_M_BASIN	$8.52 \times 10^{-3}$
Topography	ELEV_STD_M_BASIN	$3.72 \times 10^{-3}$
Topography	SLOPE_PCT	$1.00 \times 10^{-3}$
Topography	ASPECT_EASTNESS	$6.12 \times 10^{-4}$
Soils	CLAYAVE	$1.48 \times 10^{-4}$
Soils	SILTAVE	$1.33 \times 10^{-4}$
Soils	SANDAVE	$7.12 \times 10^{-5}$
Soils	AWCAVE	$3.86 \times 10^{-4}$
Population Infrastructure	PDEN_2000_BLOCK	$8.26 \times 10^{-3}$
Population Infrastructure	IMPNLCD06	$5.75 \times 10^{-3}$
Population Infrastructure	ROADS_KM_SQ_KM	$3.91 \times 10^{-3}$
Watershed	DEVNLCD06	$5.29 \times 10^{-3}$
Watershed	FRAGUN_BASIN	$1.15 \times 10^{-3}$
Watershed	BAS_COMPACTNESS	$1.05 \times 10^{-3}$
Watershed	PLANTNLCD06	$2.84 \times 10^{-4}$
Watershed	DRAIN_SQKM	$1.69 \times 10^{-4}$
Dams	STOR_NID_2009	$1.30 \times 10^{-3}$
Dams	DDENS_2009	$1.27 \times 10^{-3}$
Dams	STOR_NOR_2009	$5.68 \times 10^{-4}$

**Table A3.** Predictor variables retrieved for Coastal Plain.

Variable Type	Variable	Importance Level
Climate	PRECIP_SEAS_IND	$1.45 \times 10^{-3}$
Climate	LST32F_BASIN	$1.25 \times 10^{-3}$
Climate	PPTAVG_BASIN	$9.99 \times 10^{-4}$
Climate	PET	$6.54 \times 10^{-4}$
Climate	FST32F_BASIN	$5.50 \times 10^{-4}$
Climate	T_MIN_BASIN	$4.04 \times 10^{-4}$
Climate	WDMIN_BASIN	$3.41 \times 10^{-4}$
Climate	T_MAX_BASIN	$3.30 \times 10^{-4}$
Climate	WD_BASIN	$3.01 \times 10^{-4}$
Climate	SNOW_PCT_PRECIP	$2.86 \times 10^{-4}$
Climate	T_AVG_BASIN	$1.48 \times 10^{-4}$
Climate	RH_BASIN	$1.40 \times 10^{-4}$
Climate	T_MINSTD_BASIN	$5.00 \times 10^{-5}$
Topography	ELEV_MEAN_M_BASIN	$9.33 \times 10^{-4}$
Soils	AWCAVE	$3.61 \times 10^{-3}$
Soils	SANDAVE	$3.55 \times 10^{-3}$
Soils	PERMAVE	$2.11 \times 10^{-3}$
Soils	SILTAVE	$1.62 \times 10^{-3}$
Soils	CLAYAVE	$2.84 \times 10^{-4}$
Soils	BDAVE	$1.42 \times 10^{-4}$
Watershed	PLANTNLCD06	$1.94 \times 10^{-3}$
Watershed	FORESTNLCD06	$5.55 \times 10^{-4}$
Watershed	HIRES_LENTIC_DENS	$4.87 \times 10^{-4}$
Watershed	FRAGUN_BASIN	$2.12 \times 10^{-4}$
Watershed	DRAIN_SQKM	$2.85 \times 10^{-4}$
Watershed	DEVNLCD06	$2.29 \times 10^{-4}$
Population Infrastructure	IMPNLCD06	$2.49 \times 10^{-4}$
Population Infrastructure	ROADS_KM_SQ_KM	$6.57 \times 10^{-4}$
Population Infrastructure	PDEN_2000_BLOCK	$1.20 \times 10^{-4}$
Dams	MAJ_DDENS_2009	$3.08 \times 10^{-4}$
Dams	STOR_NOR_2009	$7.78 \times 10^{-5}$
Dams	STOR_NID_2009	$3.70 \times 10^{-5}$

**Table A4.** Predictor variables retrieved for Interior Highlands.

Variable Type	Variable	Importance Level
Climate	FST32F_BASIN	$2.41 \times 10^{-3}$
Climate	LST32F_BASIN	$1.71 \times 10^{-3}$
Climate	T_MIN_BASIN	$4.73 \times 10^{-4}$
Climate	T_AVG_BASIN	$3.67 \times 10^{-4}$
Climate	PET	$2.22 \times 10^{-4}$
Climate	PPTAVG_BASIN	$2.02 \times 10^{-4}$
Climate	SNOW_PCT_PRECIP	$2.02 \times 10^{-4}$
Climate	WDMAX_BASIN	$1.30 \times 10^{-4}$
Climate	T_MINSTD_BASIN	$5.67 \times 10^{-5}$
Climate	T_MAX_BASIN	$3.52 \times 10^{-5}$
Topography	ASPECT_NORTHNESS	$4.16 \times 10^{-5}$
Topography	SLOPE_PCT	$3.50 \times 10^{-6}$
Soils	CLAYAVE	$2.21 \times 10^{-4}$
Soils	SILTAVE	$2.10 \times 10^{-4}$
Soils	SANDAVE	$5.28 \times 10^{-5}$
Soils	PERMAVE	$4.45 \times 10^{-5}$
Soils	AWCAVE	$2.83 \times 10^{-5}$
Soils	BDAVE	$7.33 \times 10^{-4}$
Watershed	DEVNLCD06	$5.24 \times 10^{-5}$
Watershed	FORESTNLCD06	$2.56 \times 10^{-5}$
Watershed	FRAGUN_BASIN	$1.96 \times 10^{-5}$
Watershed	PLANTNLCD06	$6.84 \times 10^{-6}$
Population Infrastructure	PDEN_2000_BLOCK	$1.25 \times 10^{-4}$
Population Infrastructure	ROADS_KM_SQ_KM	$8.74 \times 10^{-5}$
Population Infrastructure	IMPNLCD06	$4.33 \times 10^{-5}$
Dams	DDENS_2009	$1.15 \times 10^{-3}$
Dams	STOR_NID_2009	$4.42 \times 10^{-4}$
Dams	STOR_NOR_2009	$3.37 \times 10^{-4}$
Dams	MAJ_DDENS_2009	$4.79 \times 10^{-6}$



**Table A5.** Predictor variables retrieved for Appalachian Highland.

Variable Type	Variable	Importance Level
Climate	WDMIN_BASIN	$1.67 \times 10^{-3}$
Climate	WD_BASIN	$1.66 \times 10^{-3}$
Climate	WDMAX_BASIN	$1.65 \times 10^{-3}$
Climate	T_MIN_BASIN	$1.09 \times 10^{-3}$
Climate	T_AVG_BASIN	$1.06 \times 10^{-3}$
Climate	LST32F_BASIN	$9.79 \times 10^{-4}$
Climate	T_MAX_BASIN	$9.52 \times 10^{-4}$
Climate	SNOW_PCT_PRECIP	$8.87 \times 10^{-4}$
Climate	PET	$8.64 \times 10^{-4}$
Climate	T_MAXSTD_BASIN	$4.05 \times 10^{-4}$
Climate	T_MINSTD_BASIN	$3.28 \times 10^{-4}$
Climate	FST32F_BASIN	$5.55 \times 10^{-4}$
Climate	PPTAVG_BASIN	$1.85 \times 10^{-4}$
Topography	SLOPE_PCT	$1.28 \times 10^{-3}$
Topography	ELEV_STD_M_BASIN	$7.44 \times 10^{-4}$
Topography	ELEV_MEAN_M_BASIN	$4.30 \times 10^{-4}$
Soils	SILTAVE	$2.17 \times 10^{-4}$
Soils	CLAYAVE	$1.90 \times 10^{-4}$
Soils	BDAVE	$9.56 \times 10^{-5}$
Soils	AWCAVE	$7.78 \times 10^{-5}$
Soils	SANDAVE	$1.74 \times 10^{-4}$
Soils	PERMAVE	$7.63 \times 10^{-5}$
Watershed	FORESTNLCD06	$7.43 \times 10^{-4}$
Watershed	FRAGUN_BASIN	$7.23 \times 10^{-4}$
Watershed	DEVNLCD06	$7.01 \times 10^{-4}$
Watershed	BAS_COMPACTNESS	$6.63 \times 10^{-4}$
Watershed	PLANTNLCD06	$3.10 \times 10^{-4}$
Watershed	DRAIN_SQKM	$2.99 \times 10^{-4}$
Watershed	HIRES_LENTIC_DENS	$1.29 \times 10^{-4}$
Population Infrastructure	IMPNLCD06	$8.42 \times 10^{-4}$
Population Infrastructure	PDEN_2000_BLOCK	$5.19 \times 10^{-4}$
Population Infrastructure	ROADS_KM_SQ_KM	$2.67 \times 10^{-4}$
Dams	DDENS_2009	$4.22 \times 10^{-4}$
Dams	STOR_NOR_2009	$2.44 \times 10^{-4}$
Dams	STOR_NID_2009	$1.85 \times 10^{-4}$
Dams	MAJ_DDENS_2009	$2.93 \times 10^{-5}$

**Table A6.** Predictor variables retrieved for all sites combined.

Variable Type	Variable	Importance Level
Climate	WD_BASIN	$2.66 \times 10^{-3}$
Climate	LST32F_BASIN	$2.60 \times 10^{-3}$
Climate	WDMAX_BASIN	$1.65 \times 10^{-3}$
Climate	PPTAVG_BASIN	$2.02 \times 10^{-4}$
Climate	FST32F_BASIN	$2.41 \times 10^{-3}$
Climate	RH_BASIN	$1.40 \times 10^{-4}$
Climate	T_MIN_BASIN	$1.09 \times 10^{-3}$
Climate	T_AVG_BASIN	$1.25 \times 10^{-3}$
Climate	PET	$1.22 \times 10^{-3}$
Climate	SNOW_PCT_PRECIP	$1.18 \times 10^{-3}$
Climate	T_MAXSTD_BASIN	$1.15 \times 10^{-3}$
Climate	T_MAX_BASIN	$1.14 \times 10^{-3}$
Climate	WDMIN_BASIN	$7.69 \times 10^{-4}$
Climate	PRECIP_SEAS_IND	$8.60 \times 10^{-4}$
Climate	T_MINSTD_BASIN	$4.04 \times 10^{-4}$
Topography	ASPECT_EASTNESS	$2.57 \times 10^{-4}$
Topography	ELEV_MEAN_M_BASIN	$4.30 \times 10^{-4}$
Topography	SLOPE_PCT	$7.97 \times 10^{-4}$
Topography	ELEV_STD_M_BASIN	$5.87 \times 10^{-4}$
Topography	ASPECT_NORTHNESS	$8.12 \times 10^{-6}$
Soils	AWCAVE	$1.02 \times 10^{-3}$
Soils	SANDAVE	$9.67 \times 10^{-4}$
Soils	SILTAVE	$6.42 \times 10^{-4}$
Soils	BDAVE	$6.00 \times 10^{-4}$
Soils	CLAYAVE	$5.20 \times 10^{-4}$
Soils	PERMAVE	$5.14 \times 10^{-4}$
Watershed	FRAGUN_BASIN	$7.84 \times 10^{-4}$
Watershed	PLANTNLCD06	$7.48 \times 10^{-4}$
Watershed	FORESTNLCD06	$6.18 \times 10^{-4}$
Watershed	DEVNLCD06	$5.60 \times 10^{-4}$
Watershed	DRAIN_SQKM	$4.89 \times 10^{-4}$
Watershed	BAS_COMPACTNESS	$4.17 \times 10^{-4}$
Watershed	HIRES_LENTIC_DENS	$3.05 \times 10^{-4}$
Population Infrastructure	IMPNLCD06	$6.42 \times 10^{-4}$
Population Infrastructure	PDEN_2000_BLOCK	$5.39 \times 10^{-4}$
Population Infrastructure	ROADS_KM_SQ_KM	$3.15 \times 10^{-4}$
Dams	DDENS_2009	$2.98 \times 10^{-4}$
Dams	STOR_NOR_2009	$2.66 \times 10^{-4}$
Dams	STOR_NID_2009	$2.36 \times 10^{-4}$

## References

1. Cook, W.; Barr, P.J.; Halling, M.W. Bridge failure rate. *J. Perform. Constr. Facil.* **2015**, *29*, 04014080. [\[CrossRef\]](#)
2. Nowak, A.S.; Collins, K.R. *Reliability of Structures*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2012.
3. Briaud, J.L.; Brandimarte, L.; Wang, J.; D'Odorico, P. Probability of scour depth exceedance owing to hydrologic uncertainty. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2007**, *1*, 77–88. [\[CrossRef\]](#)
4. Stein, S.M.; Sedmera, K. *Risk-Based Management Guidelines for Scour at Bridges with Unknown Foundations*; NCHRP Report No. 107; Transportation Research Board of the National Academies: Washington, DC, USA, 2006.
5. Arneson, L.A.; Zevenbergen, L.W.; Lagasse, P.F.; Clopper, P.E. *Evaluating Scour at Bridges*; FHWA-HIF-12-003 HEC-18; National Highway Institute: Arlington, VA, USA, 2012.
6. Apelt, C.J.; Isaacs, L.T. Bridge piers—Hydrodynamic force coefficients. *J. Hydraul. Div.* **1968**, *94*, 17–30. [\[CrossRef\]](#)
7. Wang, Y.; Zou, Y.; Xu, L.; Luo, Z. Analysis of water flow pressure on bridge piers considering the impact effect. *Math. Probl. Eng.* **2015**, *2015*, 687535. [\[CrossRef\]](#)
8. Kerenyi, K.; Sofu, T.; Guo, J. *Hydrodynamic Forces on Inundated Bridge Decks*; FHWA-HRT-09-028; Federal Highway Administration: McLean, VA, USA, 2009.
9. Istrati, D.; Buckle, I.; Lomonaco, P.; Yim, S. Deciphering the tsunami wave impact and associated connection forces in open-girder coastal bridges. *J. Mar. Sci. Eng.* **2018**, *6*, 148. [\[CrossRef\]](#)
10. Haehnel, R.B.; Daly, S.F. *Daly. Maximum Impact Force of Woody Debris on Floodplain Structures*; Engineer Research and Development Center Hanover NH Cold Regions Research and Engineering Lab: Hanover, NH, USA, 2002.
11. Hasanpour, A.; Istrati, D.; Buckle, I. Coupled SPH–FEM Modeling of Tsunami-Borne Large Debris Flow and Impact on Coastal Structures. *J. Mar. Sci. Eng.* **2021**, *9*, 1068. [\[CrossRef\]](#)
12. Hasanpour, A.; Istrati, D.; Buckle, I.G. Multi-Physics Modeling of Tsunami Debris Impact on Bridge Decks. In Proceedings of the 3rd International Conference on Natural Hazards & Infrastructure, Athens, Greece, 5–7 July 2022.
13. Diehl, T.H. *Potential Drift Accumulation at Bridges*; Federal Highway Administration, U.S. Department of Transportation: Washington, DC, USA, 1997.
14. Lagasse, P.F.; Zevenbergen, L.W.; Clopper, P.E. *Effects of Debris on Bridge Pier Scour*; National Cooperative Highway Research Program, Transportation Research Board, USA: Washington, DC, USA, 2010.
15. Lyn, D.A.; Cooper, T.J.; Yi, Y.; Sinha, R.N.; Rao, A.R. *Factors in Debris Accumulation at Bridge Piers*; Federal Highway Administration, U.S. Department of Transportation, USA: Washington, DC, USA, 2003.
16. Istrati, D.; Hasanpour, A.; Buckle, I. Numerical investigation of tsunami-borne debris damming loads on a coastal bridge. In Proceedings of the 17th World Conference Earthquake Engineering, Sendai, Japan, 27 September–2 October 2021.
17. Meyer, M.D.; Flood, M.; Keller, J.; Lennon, J.; McVoy, G.; Dorney, C.; Leonard, K.; Hyman, R.; Smith, J. *Climate Change, Extreme Weather Events and the Highway System: A Practitioner's Guide*; NCHRP: Washington, DC, USA, 2013.
18. Neumann, J.E.; Price, J.; Chinowsky, P.; Wright, L.; Ludwig, L.; Streeter, R.; Jones, R.; Smith, J.B.; Perkins, W.; Jantarasami, L.; et al. Climate change risks to US infrastructure: Impacts on roads, bridges, coastal development, and urban drainage. *Clim. Chang.* **2014**, *131*, 97–109.
19. ASCE. *A Comprehensive Assessment of America's Infrastructures*; ASCE: Reston, VA, USA, 2021.
20. Eljabri, S.S.M. New Statistical Models for Extreme Values. Ph.D. Thesis, The University of Manchester, Manchester, UK, 2013.
21. Northrop, P.J. Likelihood-based approaches to flood frequency estimation. *J. Hydrol.* **2004**, *292*, 96–113. [\[CrossRef\]](#)
22. Hrachowitz, M.; Savenije, H.H.G.; Blöschl, G.; McDonnell, J.J.; Sivapalan, M.; Pomeroy, J.W.; Arheimer, B.; Blume, T.; Clark, M.P.; Ehret, U.; et al. A decade of Predictions in Ungauged Basins (PUB)—A review. *Hydrol. Sci. J.* **2013**, *58*, 1198–1255. [\[CrossRef\]](#)
23. Liu, X.; Ashraf, F.U.; Strom, K.B.; Wang, K.H.; Briaud, J.L.; Sharif, H.; Shafique, S.B. *Assessment of the Effects of Regional Channel Stability and Sediment Transport on Roadway Hydraulic Structures: Final Report*; Texas Department of Transportation: Austin, TX, USA, 2014.
24. Shen, H.W.; Schumm, S.A.; Nelson, J.D.; Doehring, D.O.; Skinner, M.M. *Methods for Assessment of Stream Related Hazards to Highways and Bridges*; Federal Highway Administration: Washington, DC, USA, 1980.
25. Brice, J.C. *Stability of Relocated Stream Channels*; U.S. Geological Survey: Washington, DC, USA, 1981.
26. Lagasse, P.F.; Clopper, P.E.; Pagan-Ortiz, J.E.; Zevenbergen, L.W.; Arneson, L.A.; Schall, J.D.; Girard, L.G. *Bridge Scour and Stream Instability Countermeasures: Experience, Selection and Design Guidance*; Federal Highway Administration: Washington, DC, USA, 2009.
27. Johnson, P.A. Physiographic Characteristics of Bridge-Stream Intersections. *River Res. Appl.* **2006**, *22*, 617–630. [\[CrossRef\]](#)
28. Flint, M.M.; Fringer, O.; Billington, S.L.; Freyberg, D. Historical analysis of hydraulic bridge collapses in the continental United States. *J. Infrastruct. Syst.* **2017**, *23*, 04017005. [\[CrossRef\]](#)
29. Bozorgnia, M. *Computational Fluid Dynamic Analysis of Highway Bridge Superstructures Exposed to Hurricane Waves*; University of Southern California: California, LA, USA, 2012.
30. Xiang, T.; Istrati, D. Assessment of Extreme Wave Impact on Coastal Decks with Different Geometries via the Arbitrary Lagrangian-Eulerian Method. *J. Mar. Sci. Eng.* **2021**, *9*, 1342. [\[CrossRef\]](#)
31. Robertson, I.N.; Riggs, H.R.; Yim, S.C.; Young, Y.L. Lessons from Hurricane Katrina storm surge on bridges and buildings. *J. Waterw. Port Coast. Ocean Eng.* **2007**, *133*, 463–483. [\[CrossRef\]](#)

32. Hasanpour, A.; Istrati, D. Reducing extreme flooding loads on essential facilities via elevated structures. In Proceedings of the ASCE Lifelines Conference 2021–2022, University of California Los Angeles, Los Angeles, CA, USA, 1–3 February 2022.
33. Tomiczek, T.; Wyman, A.; Park, H.; Cox, D.T. Modified Goda Equations to Predict Pressure Distribution and Horizontal Forces for Design of Elevated Coastal Structures. *J. Waterw. Port Coast. Ocean Eng.* **2019**, *145*, 04019023. [\[CrossRef\]](#)
34. Panici, D.; Almeida, G.A.M. Formation, growth, and failure of debris jams at bridge piers. *Water Resour. Res.* **2018**, *54*, 6226–6241. [\[CrossRef\]](#)
35. Gippel, C.; O'Neill, I.; Finlayson, B.; Schnatz, I. Hydraulics guidelines for the reintroduction and management of large woody debris in lowland rivers. *Regul. Rivers Res. Manag.* **1996**, *12*, 223–236. [\[CrossRef\]](#)
36. Parola, A.C.; Apelt, C.J.; Jempson, M.A. *Debris Forces on Highway Bridges*; National Cooperative Highway Research Program Transportation Research Board, USA: Washington, DC, USA, 2000.
37. Coles, G.S. *An Introduction to Statistical Modeling of Extreme Values*; Springer: New York, NY, USA, 2001.
38. Dey, D.K.; Roy, D.; Yan, J. Univariate Extreme Value Analysis. In *Extreme Value Modeling and Risk Analysis, Methods and Applications*; Dey, D.K., Yan, J., Eds.; CRC Press: Boca Raton, FL, USA, 2016; pp. 1–2.
39. Stedinger, J.R.; Vogel, R.M.; Foufoula-Georgiou, E. Frequency Analysis of Extreme Events. In *Handbook of Hydrology*, 1st ed.; Maidment, D.R., Ed.; McGraw Hill Education: New York, NY, USA, 1993; pp. 18.1–18.66.
40. Hosking, J.R.M.; Wallis, J.R. *Regional Frequency Analysis*; Cambridge University Press: New York, NY, USA, 1997.
41. Koutsoyiannis, D. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrol. Sci. J.* **2004**, *49*, 575–590. [\[CrossRef\]](#)
42. Falcone, J. *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow*; U.S. Geological Survey: Reston, VA, USA, 2011.
43. Tyralis, H.; Papacharalampous, G.A.; Tantane, S. How to explain and predict the shape parameter of the generalized extreme value distribution of streamflow extremes using a big dataset. *J. Hydrol.* **2019**, *574*, 628–645. [\[CrossRef\]](#)
44. Papacharalampous, G.A.; Tyralis, H. Evaluation of random forests and Prophet for daily streamflow forecasting. *Adv. Geosci.* **2018**, *45*, 201–208. [\[CrossRef\]](#)
45. Papacharalampous, G.A.; Tyralis, H.; Koutsoyiannis, D. One-step ahead forecasting of geophysical processes within a purely statistical framework. *Geosci. Lett.* **2018**, *5*, 12. [\[CrossRef\]](#)
46. Tyralis, H.; Dimitriadis, P.; Koutsoyiannis, D.; O'Connell, P.E.; Tzouka, K.; Iliopoulou, T. On the long-range dependence properties of annual precipitation using a global network of instrumental measurements. *Adv. Water Resour.* **2018**, *111*, 301–318. [\[CrossRef\]](#)
47. Addor, N.; Nearing, G.; Prieto, C.; Newman, A.J.; Le Vine, N.; Clark, M.P. A ranking of hydrological signatures based on their predictability in space. *Water Resour. Res.* **2018**, *54*, 8792–8812. [\[CrossRef\]](#)
48. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2009.
49. Grömping, U. Variable importance in regression models. *WIREs Comput. Stat.* **2015**, *7*, 137–152. [\[CrossRef\]](#)
50. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News.* **2002**, *2*, 18–22.
51. Breiman, L.; Cutler, A.; Liaw, A.; Wiener, M. *randomForest: Breiman and Cutler's Random Forests for Classification and Regression*; R Package Version 4.6-14; 2018; Available online: <https://cran.r-project.org/web/packages/randomForest/index.html> (accessed on 1 September 2022).
52. Jiang, B. A New Classification Scheme for Data with a Heavy-Tailed Distribution. *Prof. Geogr.* **2012**, *65*, 482–494. [\[CrossRef\]](#)
53. Lima, C.H.R.; Lall, U.; Troy, T.; Devineni, N. A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates. *J. Hydrol.* **2016**, *541*, 816–823. [\[CrossRef\]](#)
54. Wallis, J.R.; Schaefer, M.G.; Barker, B.L.; Taylor, G.H. Regional precipitation frequency analysis and spatial mapping for 24-hour and 2-hour durations for Washington State. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 415–442. [\[CrossRef\]](#)
55. He, J.; Anderson, A.; Valeo, C. Bias compensation in flood frequency analysis. *Hydrol. Sci. J.* **2015**, *60*, 381–401. [\[CrossRef\]](#)
56. Morrison, J.E.; Smith, J.A. Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resour. Res.* **2002**, *38*, 41–1–41–12. [\[CrossRef\]](#)
57. Gupta, V.K.; Waymire, E. Multiscaling properties of spatial rainfall and river flow distributions. *J. Geophys. Res.* **1990**, *95*, 1999–2009. [\[CrossRef\]](#)
58. Burlando, P.; Rosso, R. Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation. *J. Hydrol.* **1996**, *187*, 45–64. [\[CrossRef\]](#)
59. Villarini, G.; Smith, J.A. Flood peak distributions for the eastern United States. *Water Resour. Res.* **2010**, *46*. [\[CrossRef\]](#)
60. Villarini, G.; Smith, J.A.; Serinaldi, F.; Ntelekos, A.A. Analyses of seasonal and annual maximum daily discharge records for central Europe. *J. Hydrol.* **2011**, *399*, 299–312. [\[CrossRef\]](#)
61. Villarini, G.; Smith, J.A.; Baack, M.L.; Krajewski, W.F. Examining Flood Frequency Distributions in the Midwest, U.S. *J. Am. Water Resour. Assoc.* **2011**, *47*, 447–463. [\[CrossRef\]](#)
62. Gvoždíková, B.; Müller, M. Evaluation of extensive floods in western/central Europe. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3715–3725. [\[CrossRef\]](#)
63. Istrati, D.; Hasanpour, A. Numerical Investigation of Dam Break-Induced Extreme Flooding of Bridge Superstructures. In Proceedings of the 3rd Intl Conference on Natural Hazards & Infrastructure, Athens, Greece, 5–7 July 2022.
64. Lempérière, F. Dams and floods. *Engineering* **2017**, *3*, 144–149. [\[CrossRef\]](#)
65. Simon, A.; Rinaldi, M. Channel instability in the loess area of the mid-western United States. *J. Am. Water Resour. Assoc.* **2000**, *36*, 133–150. [\[CrossRef\]](#)

66. Ashraf, F.; Flint, M.M. Retrospective analysis of U.S. hydraulic bridge collapse sites to assess HYRISK performance. In Proceedings of the ASCE World Environmental and Water Resources Congress 2021, Online, 7–11 June 2021.
67. Ashraf, F.; Flint, M.M. A Novel paradigm of risk study for bridge infrastructure. In Proceedings of the ASCE World Environmental and Water Resources Congress 2019, Pittsburgh, PA, USA, 19–23 May 2019.
68. US Code of Federal Regulations. *Bridges, Structures and Hydraulics*; US Code of Federal Regulations: Washington, DC, USA, 2009.
69. Diluzio, M.; Johnson, G.L.; Daly, C.; Eischeid, J.K.; Arnold, J.G. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 475–497. [[CrossRef](#)]
70. Daly, C.; Halbleib, M.; Smith, J.I.; Gibson, W.P.; Doggett, M.K.; Taylor, G.H.; Curtis, J.; Pasteris, P.A. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *Int. J. Climatol.* **2008**, *28*, 2031–2064. [[CrossRef](#)]
71. McCabe, G.J.; Wolock, D.M. Recent declines in western U.S. snowpack in the context of twentieth-century climate variability. *Earth Interact.* **2009**, *13*, 1–15. [[CrossRef](#)]
72. USGS. *U.S. Geological Survey 2008h*; USGS: Reston, VA, USA, 2008. Available online: <https://www.usgs.gov/> (accessed on 10 March 2022).
73. Wolock, D.M. *STATSGO Soil Characteristics for the Conterminous United States*; U.S. Geological Survey: Reston, VA, USA, 1997. Available online: <https://water.usgs.gov/GIS/metadata/usgswrd/XML/muid.xml> (accessed on 10 March 2022).
74. USGS. *U.S. Geological Survey 2008d*; USGS: Reston, VA, USA, 2008. Available online: <https://www.mrlc.gov/> (accessed on 10 March 2022).
75. USGS. *U.S. Geological Survey 2008a*; USGS: Reston, VA, USA, 2008. Available online: <http://water.usgs.gov/nawqa/> (accessed on 10 March 2022).
76. USGS. *U.S. Geological Survey 2008b*; USGS: Reston, VA, USA, 2008. Available online: <http://waterdata.usgs.gov/nwis/> (accessed on 10 March 2022).
77. SILVIS Lab. *The Wildland-Urban Interface, GIS Data Library*; Forest and Wildlife Ecology, University of Wisconsin-Madison: Madison, WI, USA, 2008. Available online: [http://silvis.forest.wisc.edu/projects/WUI\\_Main.asp](http://silvis.forest.wisc.edu/projects/WUI_Main.asp) (accessed on 10 March 2022).
78. GeoLytics. *Census 2000 and Street 2000*; GeoLytics Inc.: Somerville, NJ, USA, 2001. Available online: <http://www.geolytics.com/> (accessed on 10 March 2022).
79. USGS. *U.S. Geological Survey 2010b*; USGS: Reston, VA, USA, 2010. Available online: <https://www.usgs.gov/national-hydrography/national-hydrography-dataset> (accessed on 10 March 2022).