

ARTICLE

<https://doi.org/10.1038/s41467-019-13184-4>

OPEN

# Floodplain inundation spectrum across the United States

Durelle T. Scott <sup>1\*</sup>, Jesus D. Gomez-Velez <sup>2</sup>, C. Nathan Jones <sup>3</sup> & Judson W. Harvey <sup>4</sup>

Floodplain inundation poses both risks and benefits to society. In this study, we characterize floodplain inundation across the United States using 5800 stream gages. We find that between 4% and 12.6% of a river's annual flow moves through its floodplains. Flood duration and magnitude is greater in large rivers, whereas the frequency of events is greater in small streams. However, the relative exchange of floodwater between the channel and floodplain is similar across small streams and large rivers, with the exception of the water-limited arid river basins. When summed up across the entire river network, 90% of that exchange occurs in small streams on an annual basis. Our detailed characterization of inundation hydrology provides a unique perspective that the regulatory, management, and research communities can use to help balance both the risks and benefits associated with flooding.

---

<sup>1</sup>Department of Biological Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg 24061 VA, USA. <sup>2</sup>Department of Civil and Environmental Engineering, Vanderbilt University, Nashville 37212 TN, USA. <sup>3</sup>Department of Biological Sciences, University of Alabama, Tuscaloosa 35487 AL, USA. <sup>4</sup>U.S. Geological Survey, Earth System Processes Division, Reston 20192 VA, USA. \*email: [dscott@vt.edu](mailto:dscott@vt.edu)

River flooding has posed risks and benefits for early civilizations to modern society. Historically, river flooding replenished fertile floodplain soils in water scarce societies<sup>1</sup>. Urban areas that preferentially occupied river banks are now a locus for catastrophic flood damages, with a potential for increasing flood frequency associated with upstream land uses and effects of climate change that may exacerbate flooding<sup>2,3</sup>. These risks are being increasingly well defined by regional to continental-scale water balance models and flood risk assessments<sup>4,5</sup>, as well as recent efforts that incorporate hydrology into global climate simulations<sup>6,7</sup>. However, moderate floodplain inundation performs services relevant to society ranging from filtering pollutants<sup>8–11</sup> and flood attenuation<sup>12</sup> to food web support and habitat for healthy fisheries<sup>13</sup>.

Studies of a broad range of floodplain inundation, from flood initiation when water first enters a floodplain through recession, are needed to understand the role of floodplains in storing water, transforming water quality, and supporting ecosystems. A large obstacle in exploring floodplain inundation throughout river basins has been the ability to quantify when stream water exchanges with the adjacent floodplain. Rather than quantifying floodplain inundation directly, thresholds for floodplain inundation are often estimated by a given return period for maximum flows (e.g. 1.5 year)<sup>14</sup> based on the inferred significance of bankfull flow in shaping the channel and the frequency analysis of annual maximum flows<sup>15</sup>. In addition, most previous studies only resolve flooding at a daily timescale, which substantially misrepresents short-lived floods in small streams. As a result, previous regional to continental-scale investigations have not fully characterized floodplain inundation<sup>16</sup>.

This study represents one of the first continental-scale investigations examining all floods from the onset to end of inundation across a long-term record. We analyzed sub-hourly flow data from 5800 USGS gaging stations across the conterminous U.S. (<https://waterdata.usgs.gov/nwis>) and quantified all floodplain inundation events and their characteristic metrics (magnitude, frequency, duration, and seasonal timing) as they vary with river size and hydroclimatic region over a 10-year flow record. First, we estimated a threshold for floodplain inundation and its uncertainty at each gage station using a breakpoint analysis of each gage's stream gaging measurements—our analysis differs in method but is consistent with quantifying bankfull flow, a widely used metric in fluvial geomorphology<sup>17</sup>. Second, we estimated the floodplain fraction of total river flow, the timing, and duration for each event over the 10-year flow record. As explained in the methods and the SI, we propagated the uncertainty from the breakpoint analysis to bound these metrics. Third, we extrapolated results across the U.S. stream network by building predictive relationships between floodplain inundation metrics and drainage area and compiling metrics by major river basin and stream order (1st–7th order).

We found that floodplains convey a substantial proportion (approximately 4–12.6%) of a river's annual flow. The median duration of floodplain inundation increased with river size, from less than a day in the smallest streams to many days in the largest rivers. Floodplain event inundation volumes are greater and floods last longer in large rivers, but flooding occurs much more often in small streams; as a result, large rivers and small streams individually account for similar magnitudes of river-floodplain exchange (except in the arid west where exchange decreases as stream order increases). Furthermore, more of the river's total annual flow interacts with floodplains in the smaller streams compared with larger rivers. A cumulative estimate of river-floodplain exchange suggested that greater than 90% of the river-floodplain exchange occurs in small streams that dominate the cumulative stream length within basins. An improved quantification of inundation

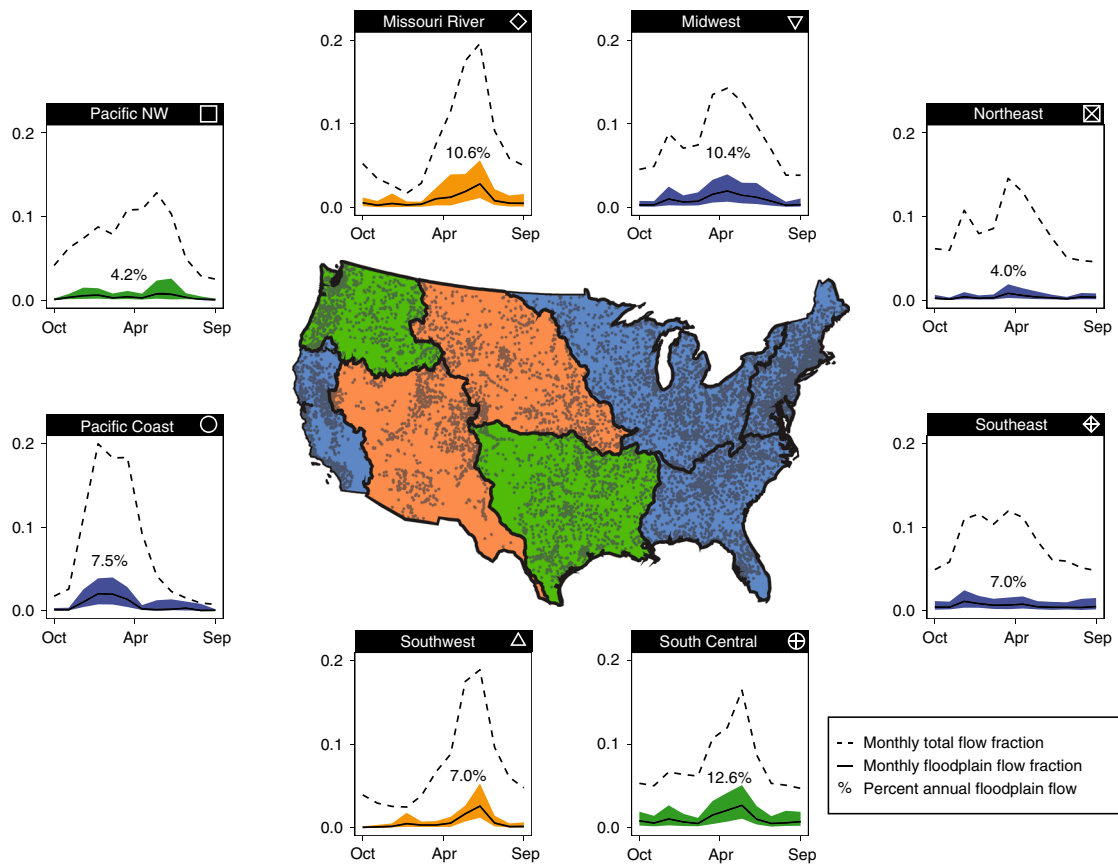
hydrology will be useful to isolate where in the network the water-quality functions of floodplains are most important, in order to inform regulatory strategies and help to prioritize management practices that protect water resources and aquatic health of river corridors.

## Results

**Magnitude of floodplain inundation.** The median annual magnitude of floodplain inundation ranged between 4.0 and 12.6% of the total annual river flow in eight major river basins (Fig. 1). The proportion of total river flow occurring on floodplains is therefore substantial, especially when compared with the median duration of floodplain inundation for U.S. stream gages, indicating that floodplains are inundated on average for only 2.6% of the year (Supplementary Fig. 1). The magnitude of floodplain inundation varied greater in some basins more than others, ranging from less than 2% within the Northeast and Pacific Northwest basins to over 29% in the South Central basin, as expressed by the difference between the 75th and 25th quantiles in each river basin (Fig. 1).

**Seasonal timing of floodplain inundation.** Regional differences in the timing of floodplain inundation are evident across river basins. In the Missouri and Southwest basins (Fig. 1), 56% and 69% of the annual floodplain inundation occurs between April and June, respectively. In contrast, over 71% of the annual floodplain inundation occurs between January and March in the Pacific Coast basin. For the Midwest and South Central basins, a high degree of seasonality in floodplain inundation is shown where 48% and 49% of the annual inundation occurs between March and May. Timing of floodplain inundation is influenced by three important hydroclimatic characteristics: annual snowmelt, balance of snowmelt/rain-on-snow contributions versus rainfall, and proportion of high-intensity rainfall (Fig. 1). These large scale hydroclimatic variables are important drivers that influence the frequency and timing of floodplain inundation<sup>18</sup>. There is a prevalence of snowmelt derived flooding in the spring and early summer in basins such as the Missouri and Southwest basins, compared with the Pacific coast basin where monsoonal driven flooding occurs during the fall and early winter. In the Midwest and South Central basins, a combination of snowmelt, rain on snow, and spring rain events on saturated soils when evaporation is low contribute to the seasonality in the central U.S. Our hydroclimatic interpretations are generally consistent with earlier flood studies that generally separate mechanisms of extreme floods into snowmelt, rain-on snow, and precipitation based events<sup>19,20</sup> and often highlighted strong gradients in these mechanisms at the continental scale<sup>20,21</sup>.

**Event inundation metrics.** We found that the flood event frequency, duration, fraction of flood that inundates floodplain, and the event inundation volumes varied systematically with river size across the U.S. The smallest streams flood more often (Fig. 2a) with a median frequency of four inundation events per year across the U.S. Event frequency in 1st order streams varied by basins, ranging from less than two events per year in the Northeast to over 10 in the Midwest. In contrast, 7th order rivers had a median event frequency of two events per year, but ranged in frequency from one event per year in the Northeast to six events per year in the Midwest. While frequency of flood events generally decreased as a function of stream size across the U.S., the median duration of flooding increased from 5 h to 4.7 days as stream size increased from 1st order streams to 7th order rivers (Fig. 2b). Our results also show that event fractions of floods that inundate floodplains, which represents the fraction of the event's

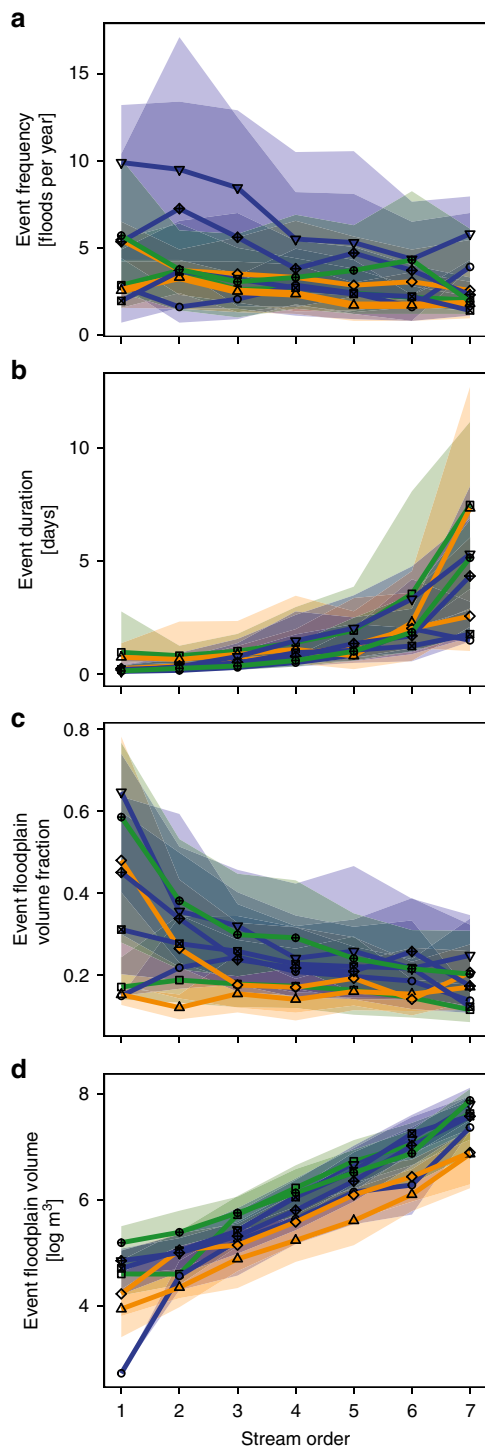


**Fig. 1** Timing of annual stream flow and inundation. (outside) Regional variation across major river basins of the mean monthly proportion of total annual river flow (dotted line) and mean monthly proportion of floodplain flow to total annual river flow. The highlighted number for each basin represents the percent of total annual river flow on floodplains. The shading represents the 25th and 75th quantiles based on analysis of all gage stations within each basin. (inside) Gage stations and basin boundaries (Northeast; Southeast; Midwest; South Central; Missouri; Southwest; Pacific Northwest; Pacific Coast) are displayed, and the background color (blue, orange, and green) indicates the hydroclimatic signature (rain dominated = blue; snow dominated = orange; and mixed = green)

total river flow that inundates the floodplain, generally decreased with river size (Fig. 2c), from a median of 65% in small streams in the Midwest to less than 12% in larger rivers in the Northeast. The exception was in the western U.S. (Pacific Northwest, Pacific Coast, and Southwest), where the event floodplain inundation volumes averaged 17% of total river flow across all stream orders. While the fraction of floods that inundate the floodplain decreased with river size, the event inundation volumes increased three orders of magnitude from 1st order streams to 7th order rivers across all river basins (Fig. 2d).

**Reach-scale and cumulative river-floodplain exchange.** Using the estimates from the gage observational network, we fitted power-law relationships between drainage area and the key descriptors of inundation (e.g., inundation volume and duration at the event and annual time scales), obtaining a parsimonious tool to map these variables into the NHD Plus river network and to get a better perspective of their spatial variability. Figure 3a illustrates this map for the case of floodplain inundation volume during typical events. From this mapped variable, we then calculated event river-floodplain exchange (Fig. 3b) at each NHD reach and also estimated the cumulative river-floodplain exchange throughout all river reaches of each major river basin (Fig. 3c). Event river-floodplain exchange is the difference in floodplain inundation volume between the upstream and downstream reach (Fig. 3a).

Event river-floodplain exchange distinguished two groups, i.e., basins where exchange was constant with stream size and basins where exchange decreased with stream size. Within the snowmelt basins of the arid west (Southwest, Missouri) and the South Central basins, river-floodplain exchange decreased as a function of stream size (Fig. 3b) from a median of  $0.9 \text{ m}^3 \text{ m}^{-1} \text{ event}^{-1}$  in 1st order streams to a median of  $0.1 \text{ m}^3 \text{ m}^{-1} \text{ event}^{-1}$  in 7th order streams. In contrast, event river-floodplain exchange within the remaining basins (Northeast, Southeast, Midwest, Pacific Coastal, and Pacific Northwest) was higher and generally less variable, ranging from  $4.1 \text{ m}^3 \text{ m}^{-1} \text{ event}^{-1}$  in 1st order streams to  $1.9 \text{ m}^3 \text{ m}^{-1} \text{ event}^{-1}$  in 7th order streams. Across all basins, cumulative annual river-floodplain exchange decreased over 5-orders of magnitude throughout the river network (Fig. 3c), as controlled by the dendritic nature of river networks<sup>22</sup> and its associated exponential decrease in total river length in each stream order. Although our simple scaling with drainage area to estimate cumulative effects does not explicitly account for detailed spatial factors such as reach-to-reach geomorphic heterogeneity along the river corridor<sup>23</sup>, including valley morphology<sup>24</sup>, channelization<sup>25</sup>, reservoirs for water storage, and flood control structures such as levees<sup>26,27</sup>, it does demonstrate the general tendency for flood magnitude, frequency, duration, and river-floodplain exchange to scale with drainage area, which provides a useful first-order assessment of the cumulative effects that can undoubtedly be improved.



**Fig. 2** Flood characteristics as a function of stream size. Median (a) event frequency [floods year<sup>-1</sup>], (b) event duration [days], (c) fraction of flood that inundates floodplain, and (d) event floodplain inundation volume [log<sub>10</sub> m<sup>3</sup>]. The shading represents the 25th and 75th quartiles of the stations within each major river basin of a given stream size (small—1st order to large—8th order)

**Implications and future outlook.** Continental-scale flood studies typically focus on large, catastrophic flood events. Our study focused on a broad range of flood inundation, from small to large floods and from the initiation of flood inundation through recession. We found that the duration of flooding is generally less than 2.6% of the year across the conterminous U.S., and varied

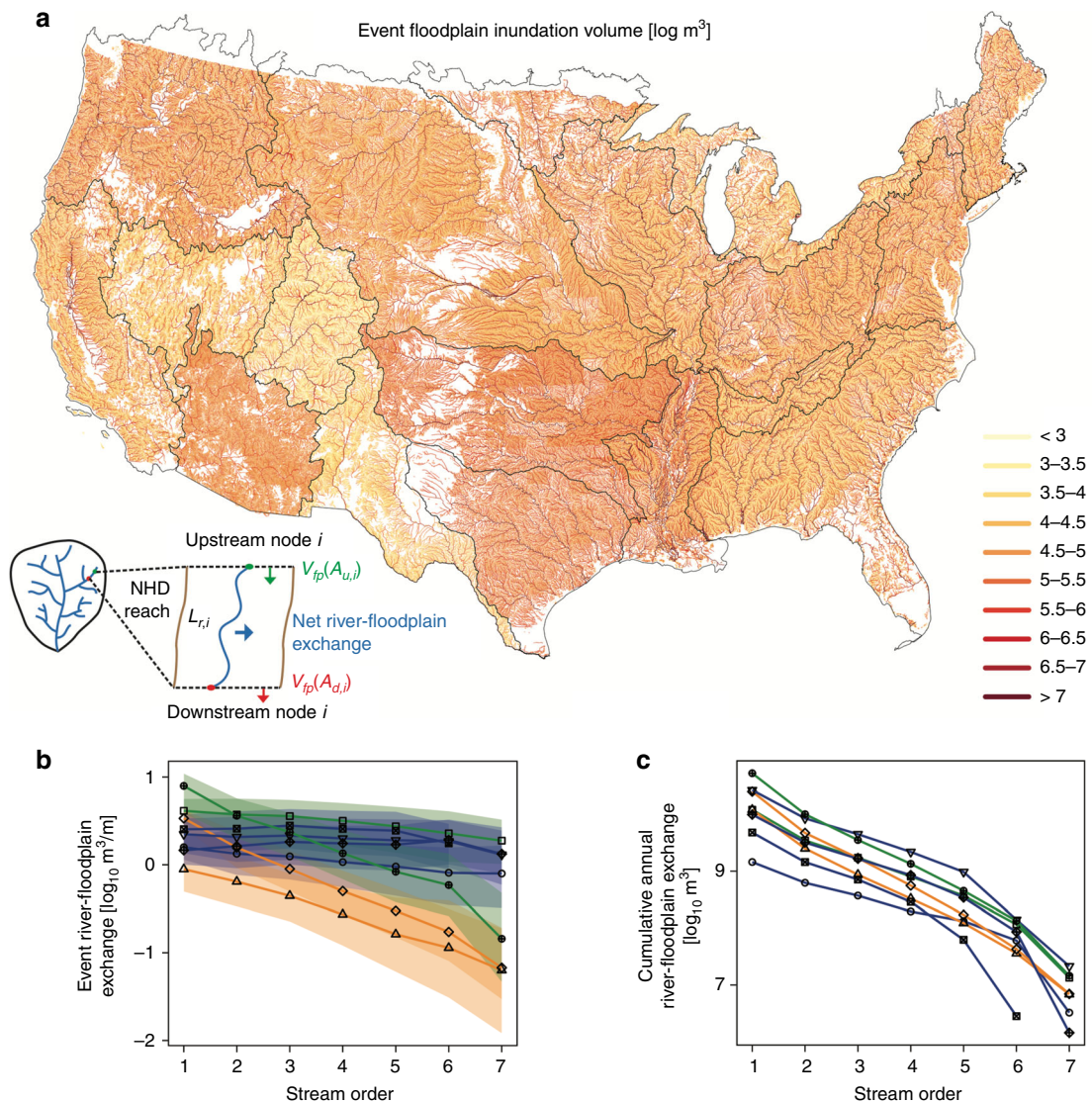
from <1% in 1st order streams to over 10% in the larger rivers (Supplementary Fig. 1). Furthermore, the proportion of total river flow reaching floodplains was determined to be a substantial proportion (3–12% median, as high as 29% in the 75th upper quartile, Fig. 1) of the annual river water discharge across the U.S., providing important functions and services at regional to continental scales.

The seasonality of inundation has implications for food production, water storage, urban infrastructure, water quality, sediment storage and supply, and ecological functioning. For example, spring flooding within the South Central and Midwest can delay planting or destroy young crops<sup>28</sup>, but may also recharge groundwater to offset the effects of extreme droughts that decrease crop yields<sup>29</sup>. Flood events during the spring within larger rivers draining into the Gulf of Mexico could partially offset riverine nutrient loads from upstream, given that spring-time nutrient loads are linked to the extent of the hypoxic zone within the Gulf of Mexico<sup>30</sup>. The majority of the other major river basins experience strong seasonal patterns resulting in periods of both high and low floodplain inundation (Fig. 1). For example, in coastal areas of the Pacific Northwest and Pacific Coast river basins, flooding triggers ecologically and culturally significant fall/winter chinook salmon migrations. Not only do the high flow events provide passage to headwater streams for spawning, nutrient-rich floodplain ecosystems provide key habitat for juvenile development and biomass production<sup>31,32</sup>.

The regional variation observed in our flood inundation metrics is attributable to a combination of hydroclimate, geomorphology, and hydrologic alteration across the major river basins. In the arid west (Missouri and Southwest), much of the stream flow is captured within reservoirs for agriculture and human water demands; previous estimates within the Colorado and Rio Grande basins are 100% capture<sup>33</sup>. The influence of water storage and capture is best observed for these two basins and the South Central (which contains the Rio Grande), where river-floodplain exchange decreased as a function of stream size. The Southwest, which has the lowest exchange across all basins also corresponds to the basin that has the lowest precipitation and is the most water stressed<sup>34</sup>. While ultimately much of the water is captured, these basins experience predictable, annual flooding from snowmelt resulting in prolonged flood periods during the late spring. The remaining basins all had relatively uniform river-floodplain exchange and similar event floodplain magnitudes; however, flood events were shorter and less frequent in the Northeast and Pacific Coast. These two basins were characterized by higher floodplain inundation thresholds, which is generally consistent with previously estimated bankfull discharge thresholds from regional analyses of hydraulic geometry<sup>35</sup>; additionally these two regions also have a high degree of hydrologic alteration which could raise or lower floodplain inundation thresholds depending on whether channel incision occurred<sup>36</sup>. Row-crop agriculture can also increase runoff from rainfall<sup>37</sup>, which will impact flood characteristics in agriculturally dominated regions (e.g., Midwest, Southeast). We also found that the Midwest, Southeast, and the Pacific Northwest had higher cumulative river-floodplain exchange throughout the river network, which has implications for water resources management. Positive implications of higher exchange include enhanced habitat, peak-flow reduction, and water quality improvements, but holistic management needs to consider that higher exchange has the potential to increase the delivery of non-point source runoff to freshwater systems.

The major societal costs of catastrophic floods, and the focus of many studies on extreme floods, may leave the impression that river flooding is always detrimental. Recently there has been an emphasis on understanding how moderate flooding, along with





**Fig. 3** Floodplain significance through river basins applied to the National Hydrography Database Plus stream network. **a** Event floodplain inundation volume [ $\log \text{m}^3$ ], with inset highlighting approach to estimate river-floodplain exchange per reach. **b** River-floodplain exchange during events [ $\log_{10} \text{m}^3 \text{m}^{-1}$ ] within all streams of order  $n$  across major river basin. The uncertainty bounds represent 95th confidence derived from the uncertainty analysis. **c** Cumulative annual river-floodplain exchange [ $\log_{10} \text{m}^3$ ] within each stream order across major river basin

other processes such as hyporheic exchange, may influence water-quality and ecological functions of large river networks<sup>11,38–40</sup>. Our results support the need to quantify floodplain inundation throughout the nation’s river corridors to isolate dominant processes, e.g., where and when floodplains dominate the removal of river nutrients, compared with streambed processing during low flow, and its downstream effects on nutrient loading to estuaries. Specifically, the magnitude, duration, and timing of floods affects the contact time of waters on floodplains with reactive floodplain soils<sup>41,42</sup>, which was hypothesized to be a key rate limiting step for biogeochemical reactions and removal efficiencies of riverborne contaminants<sup>11,43</sup>. Field studies have revealed important processes. For example, as flooding begins, floodplain soils and vegetation tend to release nutrients into floodwater which may be a source of nutrients to the river<sup>44</sup>. Following the initial flush, floodplains generally retain nutrients through physical (settling) and biogeochemical (e.g., denitrification) processes that are facilitated by longer contact times with sediment. Thus, flood durations of less than 1 day have generally been found to increase downstream nutrient export as material is

flushed from the floodplain<sup>45,46</sup>, whereas the longer duration floods have sufficient times for settling of suspended particles and contact times with floodplain sediments that promote reactions that remove nutrients, e.g., denitrification<sup>8,47</sup>. Consequently, flooding in small (low-order) streams is likely to increase nutrient loading to rivers, especially where flooding is of short duration (Fig. 2b) and where there is high cumulative water storage on floodplains (Fig. 3b). In fact, small streams (orders 1–3) with flood durations less than 1 day dominate water storage on floodplains, cumulatively representing 92% of the total water storage throughout major river basins. Larger rivers, especially in the Midwest, South Central, and Southeast major river basins are potential hotspots for nutrient storage and biogeochemical processing due to their longer durations of river-floodplain exchange that provide ample residence time for nutrient removal (Fig. 2b). The co-occurrence of high nutrient loading within the agricultural Midwest landscape and higher river-floodplain exchange, for example, present an opportunity for targeted enhancement of floodplain restoration in this region. However, cumulative water storage in floodplains adjacent to larger rivers

was less than 8% across all major river basins, which limits their overall role in nutrient removal from rivers, especially where constructed levees limit water exchange and material processing<sup>48,49</sup>. More research is needed to better understand river-floodplain exchange, and its impact on downstream fate and transport of materials across scales ranging from individual river reaches to larger river networks. While modeling<sup>16,24,50</sup> and experimental<sup>51,52</sup> investigations are beginning to characterize hydrologic exchange flows between rivers and adjacent floodplains, empirical observations are needed to ground those characterizations in reality.

Sustainable river basin management requires balancing competing water demands (e.g., human, ecosystem) combined with both local (reach-scale) and cumulative effects for downstream water storage, flood risk, and water quality functions and river health. Both scientific forefronts and applications by water resource managers can be aided by our broad characterization of floodplain inundation from small streams to large rivers and across hydroclimatic regions of the U.S. Future efforts to maintain, enhance, or reconnect floodplains need to account for the regional differences in the seasonal timing and magnitude of floodplain connectivity, in addition to the effects of short versus longer duration flood events that alter whether the floodplain acts as a nutrient source or sink and cumulative flood attenuation.

**Methods**

**General approach.** Our inundation analysis consists of four main steps. First, we identified a flooding threshold using a breakpoint analysis and estimate its uncertainty. To this end, we used sub-hourly flow data, stage-discharge ( $s - Q$ ) field measurements and rating curves ( $Q = f_{rc}(s)$ ) from 5800 USGS gaging stations (USGS NWIS, <http://waterdata.usgs.gov>) in the GAGESII database<sup>53</sup> (see Supplementary Note 1). Second, we identified individual inundation events above the threshold and applied the modified divided channel method (DCM) to separate flood flow into channel and floodplain components. Third, we estimated the following metrics for each flood event across the record: channel volume ( $V_c$ ), floodplain volume ( $V_{fp}^e$ ), event duration ( $d_f$ ), and time between individual events ( $b_f$ ) (a full list of variables and their definition is available in Supplementary Table 1). The uncertainty in the flooding threshold is propagated through the metric estimation process. Lastly, we generalized our inundation metrics to sub-divisions of Major River Basins and stream order ( $\omega$ ) across the stream network using the National Hydrography Database [NHD Plus V2, <http://nhd.usgs.gov>] (see Supplementary Note 2) by applying regional power-law relationships based on drainage area. This approach allowed us to estimate event reach-scale and cumulative annual floodplain exchange throughout the river network.

**Breakpoint analysis for inundation.** For each gage, we established the threshold for floodplain inundation through a segmented regression analysis<sup>54</sup> of the stage ( $s$ ) and discharge ( $Q$ ) measurements used to define the gage’s rating curve ( $Q = f_{rc}(s)$ ). This approach, allowed us to identify breakpoints in the  $s$ - $Q$  relationship as well as their statistical significance and confidence intervals within a rigorous statistical inference framework<sup>55</sup>. The breakpoint between regression segments often has physical meaning<sup>56</sup>, and in this application, we assume it represents the point of incipient flooding, or the minimum flow where river-floodplain inundation occurs ( $Q_{bkf}$ ). This is consistent with hydrogeomorphic theory suggesting that there is an abrupt change in cross sectional hydraulic geometry when stream flow transitions from channel flow to overbank flooding<sup>17,57</sup>. Uncertainty associated with this estimate of  $Q_{bkf}$  is largely associated with the hysteretic nature of flood flows and river stage<sup>41,58</sup>, flow measurement error<sup>59</sup>, and non-stationarity in stream/floodplain bathymetry<sup>60</sup>. While possibly unsuitable for site level engineering applications (e.g., sizing a channel for balanced sediment transport), we suggest our estimates of the threshold for floodplain inundation stage ( $\hat{s}_{bkf}$ ) and discharge ( $\hat{Q}_{bkf}$ ) and their confidence intervals (CIs) provide a useful characterization of floodplain inundation to reveal patterns at the continental scale.

For each gaging site, a linear segmented model of the form

$$Q(s) = \beta_0 + \beta_1 s + \beta_2 (s - \psi_b) I(s > \psi_b) \tag{1}$$

was iteratively fitted to find one or multiple breakpoints  $\psi_b$  and their statistical significance. In Eq. (1),  $\psi_b$  is the breakpoint and  $I(\cdot)$  is the indicator function equal to one when the argument is true. To this end, we used the package segmented implemented in the R statistical software<sup>61</sup>—details of the theoretical and iterative approaches used by segmented can be found in the package’s documentation and the reference<sup>62</sup>.

For each gage station, we tested the models with zero, one, two, and three breakpoints and used the Bayesian information criterion (BIC) to select the most parsimonious model<sup>62</sup>. Based on the BIC, all of the stations were characterized by one or two breakpoints, and in the cases where two breakpoints were identified, the breakpoint with the highest stage value was physically reasonable and had a better agreement with field observations.

The majority of the gage stations resulted in breakpoint-models with high adjusted  $R^2$  fits (Supplementary Fig. 2). The 25th percentile adjusted  $R^2$  was 0.75, the 50th was 0.89, and the 75th was 0.96. Less than 5% of our gages had adjusted  $R^2$  values of less than 0.40. Examples at each of the 25th, 50th, and 75th are provided in Supplementary Fig. 3. To validate the breakpoint analysis, we compared best-fit estimates of  $\hat{Q}_{bkf}$  to measured bankfull at 537 gages from across the US. Measured bankfull values were obtained from 36 different regional curve studies across 32 state<sup>63</sup>. Results suggest strong agreement between estimated and measured bankfull (RMSE = 0.5, see Supplementary Fig. 4).

Once a statistically significant breakpoint ( $\hat{s}_{bkf}$ ) and its 95% confidence interval [ $s_{bkf}^-, s_{bkf}^+$ ] were identified, we used the published USGS stage-discharge rating curve to obtain an estimator for the bankfull discharge  $\hat{Q}_{bkf} = f_{rc}(\hat{s}_{bkf})$  and its corresponding 95% confidence interval [ $Q_{bkf}^-, Q_{bkf}^+$ ] = [ $f_{rc}(s_{bkf}^-), f_{rc}(s_{bkf}^+)$ ]. We then addressed the uncertainties in the estimation of  $s_{bkf}$  and  $Q_{bkf}$  by using a Monte Carlo approach<sup>64,65</sup>, where we assumed that the probability distribution function (pdf) for the bankfull stage  $s_{bkf}$  is normally distributed with mean  $\hat{s}_{bkf}$  and standard deviation  $\sigma_{s_{bkf}} = (s_{bkf}^+ - \hat{s}_{bkf}) / (\sqrt{2} \operatorname{erf}^{-1}(2 \cdot 0.975 - 1))$ . That is,  $s_{bkf} \sim N(\hat{s}_{bkf}, \sigma_{s_{bkf}}^2)$ . Once the pdf for  $s_{bkf}$  was defined, we generated random realizations of stage (1000 in our analysis) and mapped them through the USGS stage-discharge rating curve  $f_{rc}(s)$  to obtain an empirical pdf for bankfull discharge ( $Q_{bkf}$ )<sup>65</sup>. The one thousand realizations of  $Q_{bkf}$  allowed us to account for the uncertainty of the flooding metrics, repeating the analysis described below for each realization. In other words, this methodology allowed us to estimate empirical probability distribution functions for bankfull stage and discharge and all the metrics characterizing the temporal evolution of river-floodplain inundation at each station.

Our next step was to apply a Gaussian smoothing filter to remove spurious fluctuations of stage and discharge within the time series (Supplementary Fig. 5). In this case, a filtered time series is obtained by using a convolution of the form<sup>66</sup>

$$Y_f(t) = \int_{-\infty}^{\infty} g(t - u) Y(t) du \tag{2}$$

where  $Y(t)$  is the original time series and  $g(t)$  is a Gaussian kernel given by

$$g(t) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{3}$$

with a smoothing parameter  $\sigma$ .

**Event analysis across flow record.** For each realization of  $Q_{bkf}$ , we identified individual inundation events that exceed the randomly generated bankfull discharge across the gage flow record (10 years) and calculated their associated metrics. We calculated the total volume during each event ( $V_t$ ) by integrating under the hydrograph with respect to time over the duration of each flood event (Supplementary Fig. 6). The total volume can be conceptualized as the sum of the channel ( $V_c$ ) and floodplain inundation ( $V_{fp}^e$ ) volumes during the event, that is  $V_t = V_{fp}^e + V_c$  (Supplementary Fig. 7). A modification of the Direct Channel Method (DCM)<sup>67</sup> was used to estimate  $V_c$  and  $V_{fp}^e$ . To find  $V_c$ , it is necessary to further subdivide the channel compartment into the bankfull channel and bankfull excess compartments (Supplementary Fig. 7).  $V_c$  is the sum of total volume of flow in the bankfull channel ( $V_{bkf}$ ) and a bankfull excess compartment that remains associated with the channel ( $V_{bfe}$ ):

$$V_c = V_{bkf} + V_{bfe} \tag{4}$$

Here,  $V_{bkf}$  is calculated by integrating the realization of the bankfull flow ( $Q_{bkf}$ ) over the course of the event using

$$V_{bkf} = \int_{t_{start}}^{t_{end}} Q_{bkf} dt = Q_{bkf} (t_{end} - t_{start}). \tag{5}$$

Similarly,  $V_{bfe}$  is calculated integrating  $U_{bkf}$  across the bankfull excess cross sectional area over the course of the individual storm:

$$V_{bfe} = \int_{t_{start}}^{t_{end}} U_{bkf} w_{bkf} (s - s_{bkf}) dt. \tag{6}$$

$U_{bkf}$  was assumed to represent average velocity in the bankfull excess compartment.  $U_{bkf}$  is estimated using the empirical relationships developed by Bjerklie<sup>68</sup> with predictors derived from the NHD Plus V2 dataset (see Supplementary Note 3 and Supplementary Fig. 8). Bankfull channel width ( $w_{bkf}$ ) and depth ( $d_{bkf}$ ) were estimated using the downstream hydraulic geometry relationships developed by Wilkerson et al.<sup>69</sup>. Finally, we calculated the floodplain inundation as  $V_{fp}^e = V_t - V_c$ .

The information from the analysis applied across the flow record was then used to quantify variation in the frequency, duration, and magnitude of river-floodplain

connectivity with respect to both major river basin and  $\omega$ . Here, it is important to highlight that both major river basin and  $\omega$  categorical variables only provide coarse representation of the observed variability in event floodplain volumes across gages. Major river basin subdivisions each represent multiple U.S. Geological Survey 2-digit Hydrologic Unit Codes (HUCs), and thus often encompass heterogeneous hydroclimatic regimes. Similarly, while  $\omega$  provides a useful categorical variable to describe variation along the river corridor<sup>11,70</sup>, regional variation in river network geometry results in wide variability of stream size within a given stream order.

**Application to the US river network.** In order to generalize the river flooding metrics to the U.S. river network, we separated the gaging stations by USGS HUC-2 regions (2-digit hydrologic units of approximately 460,000 km<sup>2</sup> on average and fitted models of the form:

$$\log_{10}(V_{fp}^e) = a_{V_{fp}^e} + b_{V_{fp}^e} \log_{10}(A), \quad (7)$$

$$\log_{10}(V_{fp}^a) = a_{V_{fp}^a} + b_{V_{fp}^a} \log_{10}(A), \quad (8)$$

where the coefficients  $a_{V_{fp}}$  and  $b_{V_{fp}}$  were estimated with a Robust Linear Regression function *rlm* implemented in R statistical software<sup>61</sup> (red lines in the example for the Ohio River Basin, Supplementary Fig. 9). These models allowed us to map these variables along the reaches of the National Hydrography Dataset (NHD Plus Version 2, <http://nhd.usgs.gov>).

For each reach  $i$  of the NHD Plus Version 2 river network, two metrics are of particular interest: the average floodplain exchange per event and per unit length of channel ( $E_{fp,i}^e$ ; [L<sup>3</sup>Event<sup>-1</sup>L<sup>-1</sup>]), and the average floodplain exchange per year and per unit length of channel ( $E_{fp,i}^a$ ; [L<sup>3</sup>T<sup>-1</sup>L<sup>-1</sup>]). These metrics were calculated for each reach as follows:

$$E_{fp,i}^e = \frac{V_{fp}^e(A_{d,i}) - V_{fp}^e(A_{u,i})}{L_{r,i}}, \quad (9)$$

$$E_{fp,i}^a = \frac{V_{fp}^a(A_{d,i}) - V_{fp}^a(A_{u,i})}{L_{r,i}}, \quad (10)$$

where  $A_{d,i}$  and  $A_{u,i}$  are the drainage areas of the upstream and downstream nodes of the NHD reach, respectively,  $L_{r,i}$  is the length of the NHD reach, and the floodplain inundation volumes were calculated with Eqs. (7) and (8).

Finally, using these estimates, we calculated key metrics to represent the cumulative effect of flooding along the river network. We define the cumulative floodplain exchange for all reaches of stream order  $\omega$  during a year for each major river basin as

$$V_{fp,\omega}^a = \sum_{i \in \omega} E_{fp,i}^a L_{r,i}, \quad (11)$$

where  $V_{fp,\omega}^a$  is the cumulative floodplain exchange for all reaches of stream order  $\omega$  within their respective major river basin during a year.

## Data availability

Data is publicly available from [<https://doi.org/10.5281/zenodo.3491240>].

## Code availability

Code is publicly available from [<https://github.com/gomezvelezlab/floodplainInundationCONUS>].

Received: 5 December 2017; Accepted: 23 October 2019;

Published online: 15 November 2019

## References

- Sparks, R. Need for ecosystem management of large rivers and their floodplains. *Bioscience* **45**, 168–182 (1995).
- Mallakpour, I. & Villarini, G. The changing nature of flooding across the central United States. *Nat. Clim. Change* **5**, 250–254 (2015).
- Hirsch, R. & Archfield, S. Flood trends: not higher but more often. *Nat. Clim. Change* **5**, 198–199 (2015).
- Maidment, D. Conceptual framework for the national flood interoperability experiment. *J. Am. Water Resour. Assoc.* **53**, 245–257 (2017).
- Ward, P. et al. A global framework for future costs and benefits of river-flood protection in urban areas. *Nat. Clim. Change* **7**, 642+ (2017).
- Clark, M. et al. Improving the representation of hydrologic processes in earth system models. *Water Resour. Res.* **51**, 5929–5956 (2015).
- Milly, P. & Dunne, K. Potential evapotranspiration and continental drying. *Nat. Clim. Change* **6**, 946+ (2016).
- Scott, D., Keim, R., Edwards, B., Jones, C. & Kroes, D. Floodplain biogeochemical processing of floodwaters in the Atchafalaya River basin during the Mississippi River flood of 2011. *J. Geophys. Res.—Biogeosciences* **119**, 537–546 (2014).
- Alexander, R., Smith, R. & Schwarz, G. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. *Nature* **403**, 758–761 (2000).
- Woodward, G. et al. Continental-scale effects of nutrient pollution on stream ecosystem functioning. *Science* **336**, 1438–1440 (2012).
- Gomez-Velez, J., Harvey, J., Cardenas, M. & Kiel, B. Denitrification in the Mississippi River network controlled by flow through river bedforms. *Nat. Geosci.* **8**, 941–U75 (2015).
- Acreman, M. & Holden, J. How wetlands affect floods. *Wetlands* **33**, 773–786 (2013).
- Burgess, O., Pine, W. & Walsh, S. Importance of floodplain connectivity to fish populations in the Apalachicola River, Florida. *River Res. Appl.* **29**, 718–733 (2013).
- Dunne, T. & Leopold, L. *Water in Environmental Planning*. (Freeman, New York, 1978).
- England, Jr., J. F. et al. *Guidelines for Determining Flood Flow Frequency—Bulletin 17c*. Technical Report (U.S. Department of the Interior, U.S. Geological Survey, Reston, VA, 2018).
- Stone, M. & Cohen, S. The influence of an extended atlantic hurricane season on inland flooding potential in the southeastern United States. *Natural Hazards Earth Syst. Sci.* **17**, 439–447 (2017).
- Williams, M., Knauf, M., Cory, R., Caine, N. & Liu, F. Nitrate content and potential microbial signature of rock glacier outflow, Colorado front range. *Earth Surf. Processes. Landforms* **32**, 1032–1047 (2007).
- Montgomery, D. Process domains and the river continuum. *J. Am. Water Resour. Assoc.* **35**, 397–410 (1999).
- Merz, R. & Blöschl, G. A process typology of regional floods. *Water Resour. Res.* **39**, 1340 (2003).
- Berghuijs, W., Woods, R., Hutton, C. & Sivapalan, M. Dominant flood generating mechanisms across the United States. *Geophys. Res. Lett.* **43**, 4382–4390 (2016).
- Villarini, G. On the seasonality of flooding across the continental United States. *Adv. Water Resour.* **87**, 80–91 (2016).
- Strahler, A. Quantitative analysis of watershed geomorphology. *Eos Trans. Am. Geophys. Union* **38**, 913–920 (1957).
- Wohl, E. Connectivity in rivers. *Prog. Phys. Geography* **41**, 345–362 (2017).
- Van Appledorn, M., Baker, M. E. & Miller, A. J. River-valley morphology, basin size, and flow-event magnitude interact to produce wide variation in flooding dynamics. *Ecosphere* **10**, e02546 (2019).
- Kroes, D. & Hupp, C. The effect of channelization on floodplain sediment deposition and subsidence along the Pocomoke River, Maryland. *J. Am. Water Resour. Assoc.* **46**, 686–699 (2010).
- Gergel, S., Carpenter, S. & Stanley, E. Do dams and levees impact nitrogen cycling? simulating the effects of flood alterations on floodplain denitrification. *Global Change Biol.* **11**, 1352–1367 (2005).
- Remo, J., Carlson, M. & Pinter, N. Hydraulic and flood-loss modeling of levee, floodplain, and river management strategies, middle Mississippi River, USA. *Natural Hazards* **61**, 551–575 (2012).
- Izumi, T. & Ramankutty, N. How do weather and climate influence cropping area and intensity? *Global Food Security* **4**, 46–50 (2015).
- Lesk, C., Rowhani, P. & Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **529**, 84+ (2016).
- Turner, R., Rabalais, N. & Justic, D. Predicting summer hypoxia in the northern Gulf of Mexico: Riverine N, P, and Si loading. *Marine Pollution Bull.* **52**, 139–148 (2006).
- Allen, S., Krauss, K., Cochran, J., King, S. & Keim, R. Wetland tree transpiration modified by river-floodplain connectivity. *J. Geophys. Res.—Biogeosciences* **121**, 753–766 (2016).
- Sommer, T., Nobriga, M., Harrell, W., Batham, W. & Kimmerer, W. Floodplain rearing of juvenile chinook salmon: evidence of enhanced growth and survival. *Can. J. Fisheries Aquatic Sci.* **58**, 325–333 (2001).
- Vorosmarty, C. et al. Anthropogenic sediment retention: major global impact from registered river impoundments. *Global Planet. Change* **39**, 169–190 (2003).
- Averyt, K. et al. Sectoral contributions to surface water stress in the conterminous United States. *Environ. Res. Lett.* **8**, 035046 (2013).
- Bieger, K., Rathjens, H., Allen, P. M. & Arnold, J. G. Development and evaluation of bankfull hydraulic geometry relationships for the physiographic regions of the United States. *J. Am. Water Resour. Assoc.* **51**, 842–858 (2015).
- McManamay, R. A. et al. US cities can manage national hydrology and biodiversity using local infrastructure policy. *Proc. Natl Acad. Sci.* **114**, 9581–9586 (2017).
- Kelly, S. A., Takbari, Z., Belmont, P. & Foufoula-Georgiou, E. Human amplified changes in precipitation-runoff patterns in large river basins of the Midwestern United States. *Hydrol. Earth Syst. Sci.* **21**, 5065–5088 (2017).



38. Gomez-Velez, J. D. & Harvey, J. W. A hydrogeomorphic river network model predicts where and why hyporheic exchange is important in large basins. *Geophys. Res. Lett.* **41**, 6403–6412 (2014).
39. Raymond, P., Saiers, J. & Sobczak, W. Hydrological and biogeochemical controls on watershed dissolved organic matter transport: pulse-shunt concept. *Ecology* **97**, 5–16 (2016).
40. Covino, T. Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. *Geomorphology* **277**, 133–144 (2017).
41. Hughes, D. Floodplain inundation—processes and relationships with channel discharge. *Earth Surf. Processes Landforms* **5**, 297–304 (1980).
42. James, W., Richardson, W. & Soballe, D. Effects of residence time on summer nitrate uptake in Mississippi River flow-regulated backwaters. *River Res. Appl.* **24**, 1206–1217 (2008).
43. Harvey, J. et al. How hydrologic connectivity regulates water quality in river corridors. *J. Am. Water Resour. Assoc.* **55**, 369–381 (2019).
44. Jones, C., Scott, D., Guth, C., Hester, E. & Hession, W. Seasonal variation in floodplain biogeochemical processing in a restored headwater stream. *Environ. Sci. Technol.* **49**, 13190–13198 (2015).
45. Noe, G. & Hupp, C. Seasonal variation in nutrient retention during inundation of a short-hydroperiod floodplain. *River Res. Appl.* **23**, 1088–1101 (2007).
46. Kreiling, R. et al. Wetland management reduces sediment and nutrient loading to the upper Mississippi River. *J. Environ. Quality* **42**, 573–583 (2013).
47. Roley, S. S. et al. Floodplain restoration enhances denitrification and reach-scale nitrogen removal in an agricultural stream. *Ecol. Appl.* **22**, 281–297 (2012).
48. Tockner, K. & Stanford, J. A. Riverine flood plains: present state and future trends. *Environ. Conservation* **29**, 308–330 (2002).
49. Opperman, J. J. et al. Sustainable floodplains through large-scale reconnection to rivers. *Science* **326**, 1487–1488 (2009).
50. Czuba, J. A., David, S. R., Edmonds, D. A. & Ward, A. S. Dynamics of surface-water connectivity in a low-gradient meandering river floodplain. *Water Resour. Res.* **55**, 1849–1870 (2019).
51. Knight, D. & Demetriou, J. Flood plain and main channel flow interaction. *J. Hydraulic Eng.* **109**, 1073–1092 (1983).
52. Jones, C. N., Scott, D. T., Guth, C., Hester, E. T. & Hession, W. C. Seasonal variation in floodplain biogeochemical processing in a restored headwater stream. *Environ. Sci. Technol.* **49**, 13190–13198 (2015).
53. Falcone, J., Carlisle, D., Wolock, D. & Meador, M. Gages: a stream gage database for evaluating natural and altered flow conditions in the conterminous United States. *Ecology* **91**, 621–621 (2010).
54. Muggeo, V. M. R. Estimating regression models with unknown break-points. *Stat. Med.* **22**, 3055–3071 (2003).
55. Davies, R. B. Hypothesis testing when a nuisance parameter is present only under the alternatives. *Biometrika* **74**, 33–43 (1987).
56. Bacon, D. & Watts, D. Estimating the transition between two intersecting straight lines. *Biometrika* **58**, 525–534 (1971).
57. Hupp, C. Hydrology, geomorphology and vegetation of coastal plain rivers in the south-eastern USA. *Hydrol. Processes* **14**, 2991–3010 (2000).
58. Domeneghetti, A., Castellarin, A. & Brath, A. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. *Hydrol. Earth Syst. Sci.* **16**, 1191–1202 (2012).
59. Kennedy, E. *Discharge Ratings at Gaging Stations*. Report (U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A10, 1984).
60. Slater, L., Singer, M. & Kirchner, J. Hydrologic versus geomorphic drivers of trends in flood hazard. *Geophys. Res. Lett.* **42**, 370–376 (2015).
61. Team, R.D.C. R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna Austria* **0** (2016).
62. Muggeo, V. segmented: An {R} package to Fit Regression Models with Broken-Line Relationships. *R News* **8**, 20–25 (2008).
63. Bieger, K., Rathjens, H., Allen, P. M. & Arnold, J. G. Development and evaluation of bankfull hydraulic geometry relationships for the physiographic regions of the United States. *J. Am. Water Resour. Assoc.* **51**, 842–858 (2015).
64. Haldar, A. & Mahadevan, S. *Probability, Reliability, and Statistical Methods in Engineering Design* 1st edn (Wiley, New York ; Chichester England, 1999).
65. Smith, R. C. *Uncertainty Quantification: Theory, Implementation, and Applications*. (SIAM-Society for Industrial and Applied Mathematics, Philadelphia, 2013).
66. Gilgen, H. *Univariate Time Series in Geosciences: Theory and Examples* 2006 edn (Springer, Berlin ; New York, 2005).
67. Chow, V. *Open Channel Hydraulics* (McGraw Hill, 1959).
68. Bjerklie, D. M. Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information. *J. Hydrol.* **341**, 144–155 (2007).
69. Wilkerson, G. et al. Continental-scale relationship between bankfull width and drainage area for single-thread alluvial channels. *Water Resour. Res.* **50**, 919–936 (2014).
70. Butman, D. & Raymond, P. Significant efflux of carbon dioxide from streams and rivers in the United States. *Nat. Geosci.* **4**, 839–842 (2011).

## Acknowledgements

Support was provided by the John Wesley Powell Center for Analysis and Synthesis and NSF Geosciences. D.T.S. is supported by the Virginia Agricultural Experimental Station and the Hatch Program of the National Institute of Food and Agriculture (U.S. Department of Agriculture) and the Institute for Critical Technology and Applied Science at Virginia Tech. J.D.G.V. is funded by the U.S. National Science Foundation (award EAR 1830172) and U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Subsurface Biogeochemistry Research Program (SBR). This contribution originates from the SBR Scientific Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). J.W.H. was also supported by the USGS Water Resources Availability Program and Water Observing Systems Program. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government

## Author contributions

C.N.J. and D.T.S. conceived the original study, with input from J.D.G.V. and J.W.H. All authors wrote the paper. The data compilation and modeling were performed by J.D.G.V. and C.N.J. All authors interpreted results.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-019-13184-4>.

**Correspondence** and requests for materials should be addressed to D.T.S.

**Peer review information** *Nature Communications* thanks Timothy Covino and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019