





Degradation of floodplain integrity within the contiguous United States

Ryan R. Morrison¹  , Kira Simonson¹, Ryan A. McManamay²  & Dan Carver³

Despite the numerous hydrological, geological, and ecological benefits produced by floodplain landscapes, floodplains continue to be degraded by human activities at a much higher rate than other landscape types. This large-scale landscape modification has been widely recognized, yet a comprehensive, national dataset quantifying the degree to which human activities are responsible for this degradation has not previously been evaluated. In this research, we analyze floodplain integrity for the contiguous United States by spatially quantifying the impact of anthropogenic stressors on almost 80,000 floodplain units. We demonstrate the prevalence of human modifications through widely available geospatial datasets, which we use to quantify indicators of floodplain integrity for five essential floodplain functions of flood attenuation, groundwater storage, habitat provision, sediment regulation, and organics and solute regulation. Our results show that floodplain degradation is spatially heterogeneous and that the integrity of nearly 70% of floodplains in the United States is poor. We highlight that quantifying the integrity of spatially explicit floodplain elements can allow for restoration efforts to be targeted to the areas in most desperate need of preservation.

¹Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, CO, USA. ²Department of Environmental Science, Baylor University, Waco, TX, USA. ³Geospatial Centroid, Colorado State University, Fort Collins, CO, USA. email: ryan.morrison@colostate.edu

Floodplains are critically important ecosystems responsible for numerous environmental benefits including flood reduction^{1,2}, groundwater storage^{3,4}, sediment regulation^{2,5}, organics and solutes regulation^{6,7}, and habitat provisioning^{8,9}. Floodplains also provide important ecosystem services, such as supplying groundwater for consumption and agriculture, ensuring productive soils for farming, supporting fisheries, providing recreational opportunities, and holding cultural importance^{10,11}. Not all floodplains provide the same environmental benefits¹¹, but the overall integrity of a particular floodplain can be defined as the ability of that floodplain to support essential geomorphic, hydrologic, and ecological functions that maintain biodiversity and ecosystem services¹². This definition focuses on the holistic integrity of a floodplain's condition and its ability to perform essential functions.

Though necessary for human and environmental wellbeing¹³, floodplains are amongst the most endangered ecosystems, disappearing at a rate much higher than other landscape types¹⁴. Key environmental functions that floodplains provide are often seen as incompatible with human development¹⁵, leading to major modifications of river corridors (which include floodplain ecosystems), such as the construction of dams and levees^{16,17}, the straightening and dredging of channels¹⁸, and the expansion of agricultural and urban land use¹⁹. In the United State (U.S.) alone, over 30% of all floodplains have been directly cultivated or developed¹⁷, and an estimated 98% of the nation's rivers have been impacted by human activities²⁰ with detrimental consequences to floodplain ecosystems.

Despite the importance of floodplains and a long history of human modification, a comprehensive analysis of human-caused floodplain degradation (i.e., loss of floodplain integrity) within the contiguous U.S. (CONUS) does not exist. Efforts have been made to assess floodplain integrity at catchment levels²¹, but these types of assessments typically focus on evaluating the biological condition of riparian habitat rather than the overall functionality of a floodplain. A more comprehensive assessment of floodplain integrity at larger scales is important to reveal the geographic distributions of floodplain degradation and opportunities for restoration¹³, especially assessments using mechanistic approaches that provide human-environmental linkages and lend themselves to management actions. Thus, we performed analyses to assess the influence of human stressors on floodplain integrity across the CONUS, including the relative impacts of stressors on five specific floodplain functions of (i) flood reduction, (ii) groundwater storage, (iii) sediment regulation, (iv) organics and solute regulation (e.g., water quality benefits), and (v) habitat provisioning. Our work identifies the functions most impaired by human activities, and therefore this study is unique among other broad geographic analyses of floodplains^{14,22}. We leveraged publicly available geospatial datasets representing human activities to quantify the degree and extent of human stressors on floodplain environments, including both hydrologic and terrestrial stressors.

Results

An index of floodplain integrity. We use a quantitative framework, referred to as the index of floodplain integrity (IFI)¹², for assessing the influence of human activities on floodplain ecosystems. The IFI approach is a numerical method for quantifying floodplain conditions using geospatial extents or occurrence of floodplain landscapes and overlapping human stressors. We estimated the spatial extent of floodplains within the CONUS using results from two-dimensional hydrodynamic modeling of the 1% recurrence interval (100-year) flood based on regionalized flood frequency estimates and 90 m resolution topography²³.

Individual floodplain units were created by dividing the floodplain extent along 12-digit hydrologic unit code boundaries (HUC12). Only HUC12 areas with boundaries completely within the U.S. border were included in the analysis due to the lack of availability of comparable datasets outside of the U.S. This process created 78,304 unique floodplain units within the CONUS, with a total area of 662,566 km² and a resulting average floodplain unit area of 8.46 km².

Furthermore, we compiled datasets relevant to human activities in floodplains that impact each of the five floodplain functions (Table 1) and scaled the datasets according to their coverage density within each floodplain unit. For point, polyline, and polygon, and raster datasets, we calculated the number (count/km²), length (km/km²), or area (km²/km²) of each stressor dataset within the floodplain unit. We then rescaled all stressor datasets from zero to one, with a value of zero indicating the absence of the stressor in the floodplain, and one representing the 90th percentile of the stressor amount in the CONUS. For the three datasets that had no stressor present at the 90th percentile (canals and ditches, leveed area, and groundwater wells), we scaled the density values based on the maximum observed value in the CONUS. This stressor rescaling was done to create a consistent scale of comparison for stressor prevalence amongst all types of stressor datasets¹².

Using the stressor densities within each floodplain unit, we calculated the IFI for each floodplain unit by assuming that floodplain integrity was inversely related to the cumulative density of stressors within a floodplain unit. This allowed us to calculate IFI on a scale between zero and one, where one indicates little floodplain degradation and zero indicates loss of one or more floodplain functions. We assessed the IFI values for each of the five previously mentioned floodplain functions, and we also calculated an overall IFI that accounts for the relative impact of each floodplain function.

Floodplain degradation in the CONUS is concentrated in the southeastern U.S., throughout major river corridors such as the Mississippi River, and within the Central Valley of California (Fig. 1a). The distribution of IFI values (Fig. 1b) indicate that 68% of the total floodplain area in the CONUS are in poor condition based on an IFI threshold value of 0.7. We recognize that a single IFI threshold value does not indicate the transformation of floodplains from healthy to degraded, so it is necessary to evaluate the full spectrum of values within the CONUS. For the entire CONUS, the median IFI value is 0.762 (mean = 0.740, standard deviation = 0.152). Furthermore, the function IFI values all show a high correlation, ranging from 0.7 to 0.91 (Supplementary Fig. 1), with sediment regulation having the lowest median IFI score and flood reduction having the highest score (Table 2).

To better understand how IFI values may vary on a regional scale, we analyzed function and overall IFI values for three subregions within the CONUS that represent a range of human alterations (Fig. 2; Supplementary Tables 2–4). First, the San Joaquin subregion, with a total floodplain area of 8248 km² and 435 individual floodplain units, has an average overall IFI value of 0.691. Spatial variability in average IFI values in the San Joaquin subregion can be seen in Fig. 2, with a concentration of low IFI values in the western section of the subregion contrasting against the moderate to high IFI values of the eastern section. This distribution of IFI values reflects the densely populated western portion of the subregion and the intense agricultural development in the Salinas Valley. Both high and low IFI values can be seen in the floodplain units of the subregion, showing how strongly human development is impacting floodplain integrity.

Second, the Missouri-White subregion spans Nebraska and South Dakota has a total floodplain area of 1740 km² with 594 individual floodplain units. The average overall IFI value within

Table 1 Floodplain functional stressors and corresponding datasets.

Function	Stressor	Dataset	Resolution/ Units Used in Analysis	Data attributes	
Flood reduction	Reduced storage volume	Buildings ⁵¹	km ² /km ²	Polygon, accessed Jan 20th, 2022	
	Floodplain disconnection	Leveed areas ²⁵	km ² /km ²	Polygon, accessed Jan 29th, 2021	
Groundwater storage	Overland flow interception	Roads and railroads ⁵¹	km/km ²	Polyline, accessed Jan 20th, 2022	
	Land cover change	Forest cover loss events ⁴⁴	30 m, km ² /km ²	Raster, loss events from 2000 to 2020	
	Land cover change	Developed area ²⁴	30 m, km ² /km ²	Raster, 2019 version	
	Impermeable surface	Percent imperviousness ²⁴	30 m, km ² /km ²	Raster, 2019 version	
	Channelized overland flow	Ditches and canals ⁵²	1:100,000, km/km ²	Polyline, NHD + V2	
	Colmation	Agricultural area ²⁴	30 m, km ² /km ²	Raster, 2019 version	
	Loss of wood and vegetation	Forest cover loss events ⁴⁴	30 m, km ² /km ²	Raster, 30 m, loss events from 2000-2020	
	Lowered water table	Groundwater wells ⁵³	count/km ²	Points, accessed Jan 20th, 2022	
	Sediment regulation	Land cover change	Agricultural area ²⁴	30 m, km ² /km ²	Raster, 2019 version
		Loss of wood and vegetation	Forest cover loss events ⁴⁴	30 m, km ² /km ²	Raster, loss events from 2000 to 2020
Organics and solute regulation	Overland flow interception	Roads and railroads ⁵¹	km/km ²	Polyline, accessed Jan 20th, 2022	
	Hydrologic alteration	Probability of change in peak flows ⁴⁵		Data for NHD + V2 polylines	
	Hydrologic alteration	Probability of change in peak flows ⁴⁵		Data for NHD + V2 polylines	
	Vertical connectivity	Percent imperviousness ²⁴	30 m, km ² /km ²	Raster, 2019 version	
	Overland flow interception	Roads and railroads ⁵¹	km/km ²	Polyline, accessed Jan 20th, 2022	
Habitat provision	Loss of wood and vegetation	Forest cover loss events ⁴⁴	30 m, km ² /km ²	Raster, loss events from 2000 to 2020	
	Land cover change	Agricultural area ²⁴	30 m, km ² /km ²	Raster, 2019 version	
	Land cover change	Developed area ²⁴	30 m, km ² /km ²	Raster, 2019 version	
	Loss of wood and vegetation	Forest cover loss events ⁴⁴	30 m km ² /km ²	Raster, 2019 version	
	Species invasion	Non-native introduced vegetation ⁴⁶	30 m, km ² /km ²	Raster, 2016 version	
	Overland flow interception	Roads and railroads ⁵¹	km/km ²	Polyline, accessed Jan 20th, 2022	
Hydrologic alteration	Hydrologic alteration	Probability of change in peak flows ⁴⁵		Data for NHD + V2 polylines	

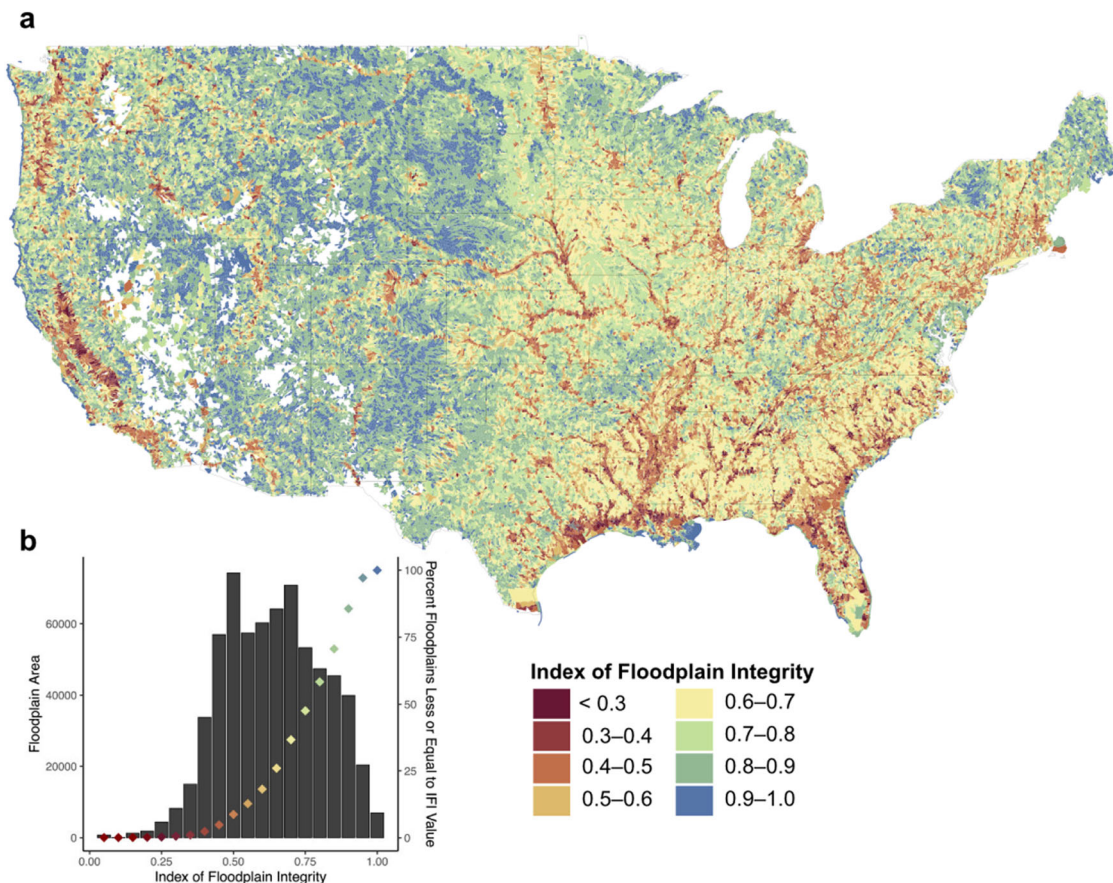


Fig. 1 Index of floodplain integrity (IFI) values in the CONUS. **a** IFI values are mapped to 12-digit HUC boundaries for easier visualization. Values near the blue end of the scale represent high integrity, and values near the red end of the scale represent low integrity. **b** Distribution of IFI values according to total floodplain area (gray columns) and percentage of all floodplains with IFI values less than or equal to a given value (diamond points).

Table 2 Summary statistics of overall IFI values for the CONUS.

Statistical Measure	Overall IFI	Flood reduction	Groundwater storage	Sediment regulation	Organics and solutes regulation	Habitat provision
Minimum	0.000	0.013	0.363	0.015	0.000	0.074
Maximum	1.000	1.000	1.000	1.000	1.000	1.000
Median	0.762	0.893	0.867	0.697	0.712	0.712
Mean	0.740	0.831	0.862	0.676	0.675	0.699
Std. Dev.	0.152	0.168	0.109	0.176	0.179	0.159

the region is 0.860. Most of the land cover in this subregion is classified as grassland or herbaceous vegetation²⁴. This lack of major anthropogenic alteration to the subregion is reflected in the high IFI values observed in the region (Fig. 2).

Third, the Lower Mississippi-Yazoo subregion, which has a high percentage of leveed areas compared to other regions of the U.S., is in the lower Mississippi River Basin and spans Arkansas, Louisiana, Mississippi, and Tennessee. The total floodplain area within the subregion is 17,969 km² with 391 individual floodplain units. The average overall IFI value within the region is 0.561. This low IFI value reflects the major loss of floodplain connectivity in the basin. Within the Lower Mississippi-Yazoo subregion, more than 68% of the floodplain intersects with levees²⁵. The high density of levees in the lower Mississippi basin has been shown to impact the sediment transport²⁶ and the flood reduction²⁷ capacity of the system. The prevalence of this stressor is reflected in the spike of lower flood reduction and sediment regulation function IFI values seen in Fig. 2.

Distribution of IFI across stream and land attributes. Human alterations to floodplain ecosystems can vary across river size²⁸ and degree of land cover development²⁹, so we further analyzed the distributions of floodplain degradation across these domains. We assessed the overall IFI values in both urban ($n = 9352$) and rural ($n = 58,565$) areas³⁰ and found that floodplains are significantly ($P < 0.0001$) more degraded within urbanized regions (Fig. 3). The IFI values of floodplains that intersected urban areas (median = 0.557, mean = 0.561, standard deviation = 0.142) are lower than the IFI values of floodplains that did not (median = 0.790, mean = 0.769, standard deviation = 0.127).

We also compared overall IFI values between stream sizes (e.g., Strahler stream orders³¹) associated with each floodplain (Fig. 4), and we found there were not significant differences in mean IFI values between stream orders except between third- and fourth-order streams. In general, IFI values tend to decrease with increasing stream order, especially for stream orders larger than eight. It is worth noting that for the orders which do not exhibit a decrease in

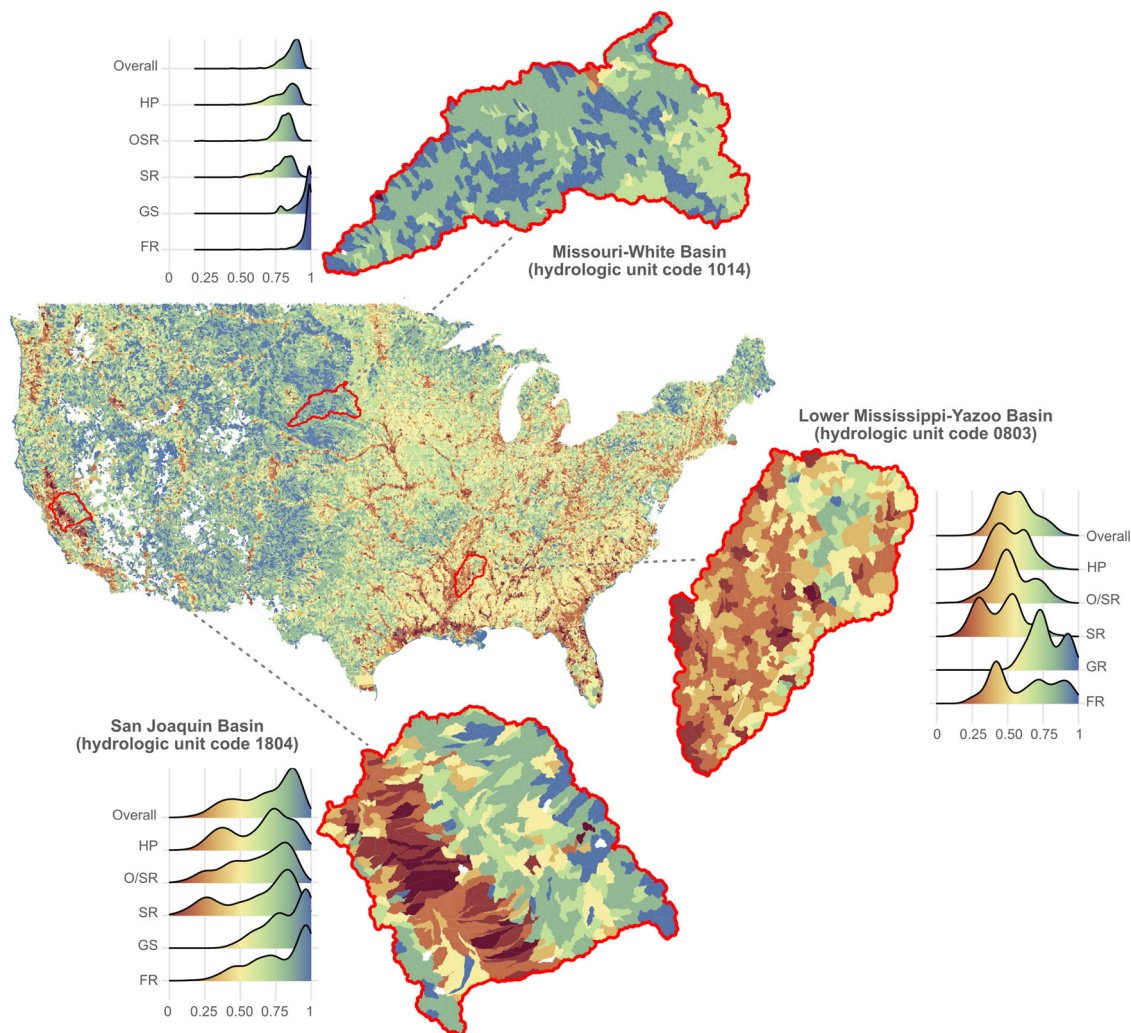


Fig. 2 Index of floodplain integrity values for subregions representing the Missouri-White Basin, Lower Mississippi-Yazoo Basin, and San Joaquin Basin. Values near the blue end of the scale represent high integrity, and values near the red end of the scale represent low integrity. Density plots for each functional IFI are shown next to each basin (Key: Overall = aggregate IFI value, HP habitat provisioning, OSR organics/solutes regulation, SR sediment regulation, GS groundwater storage, and FR flood reduction).

average IFI value as stream order increases, the sample size is small relative to those orders that follow the trend of decreasing IFI with increasing stream order (Supplementary Table 5).

Studies have shown distinct differences in watershed integrity across ecoregions³², but similar distinctions in floodplain landscapes are unclear. Therefore, we compared IFI values between nine aggregated Omernik Level III ecoregions in the U.S.³³, based on similar analyses from other studies³². The highest IFI values occurred in the Northern Plains ecoregion (median = 0.878, mean = 0.855, standard deviation = 0.079), while the lowest IFI values occurred in the Coastal Plains ecoregion (median = 0.625, mean = 0.607, standard deviation = 0.142). Differences between mean overall IFI in each ecoregion is statistically significant for all ecoregions (Fig. 4; Supplementary Table 6). We observed similar spatial patterns of overall IFI and a metric for watershed integrity³² within the nine aggregated Omernik Level III ecoregions. The three regions with the highest IFI and watershed integrity values (Northern Plains, Xeric, and Western Mountains) are consistent among datasets. The two ecoregions with the lowest IFI and watershed integrity values are also the same (Temperate Plain and Coastal Plain). This similarity in spatial distribution by ecoregion between IFI values and watershed integrity values suggests that, although watershed integrity may

not be a good way to validate floodplain functionality, the IFI methodology is successful in identifying the spatial patterns of broader river corridor conditions across the CONUS.

All three of the ecoregions with the highest IFI values have high percentages of grassland/shrub land cover type (NPL = 68%, XER = 76%, WMT = 37%). Fifty-four percent of the Western Mountains ecoregion is characterized as forested. This high proportion of undeveloped land may be responsible for the higher IFI values seen in the three regions. Of the two regions with the lowest IFI values, 69% of the Temperate Plain ecoregion and 26% of the Coastal Plain is classified as cultivated/pasture. The high percentage of agricultural development reflected by this land use is potentially driving the lower IFI values seen in the region. The Coastal Plain ecoregion is made up of many different land use types, as it encompasses the Mississippi Delta, Gulf Coast, the entire state of Florida, a portion of eastern Texas, and the Atlantic coast from Florida to New Jersey. The large variety in land use types and geospatial variability in the degree of anthropogenic modification seen through the Coastal Plain ecoregion may explain the low IFI values observed.

Comparison of IFI results to other datasets. We found it particularly challenging to assess the accuracy of our results for

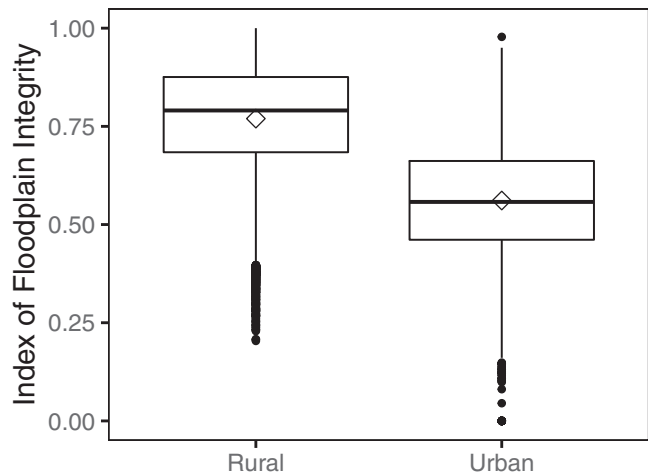


Fig. 3 Comparison of IFI values within urban and rural areas. The bottom of the boxes represents the 25th percentile and the top of the boxes represents the 75th percentile. The line through the box represents the median value, and the diamond points represent the mean value. Error bars above and below the boxes represent the 90th and 10th percentiles, respectively, and black dots represent data outliers.

numerous reasons including, (1) lack of quantitative assessments of floodplain conditions around the U.S., (2) incongruities in dataset scales, and (3) scarcity of published data. The lack of such information was the primary motivation for this study.

However, we found two datasets that were best suited to compare to our IFI result. First, we used a 2018 index of watershed integrity (IWI) dataset produced by Thornbrugh³² because it is similarly geographically distributed throughout CONUS at a catchment scale and implicitly includes floodplain landscapes in its assessment of watershed processes. Second, we used a 2015 geographic study by Konrad³⁴, which evaluated similar floodplain functions in the Pudget Sound basin of Washington State using discrete numerical categories.

A comparison between overall IFI and IWI values³² yielded no meaningful relationships. However, after comparing our IFI functions of “flood reduction”, “sediment regulation”, and “habitat provisioning” to data from Konrad³⁴, we found that our method estimated higher integrity values for “flood reduction” and “sediment regulation” and similar values for “habitat provisioning” compared to similar floodplain functions in floodplains along major rivers of the Pudget Sound basin (Supplementary Figs. 3–5). This comparative analysis was comprised of approximately 2338 km² of floodplain area (from the Konrad dataset) depending on the particular floodplain function we evaluated.

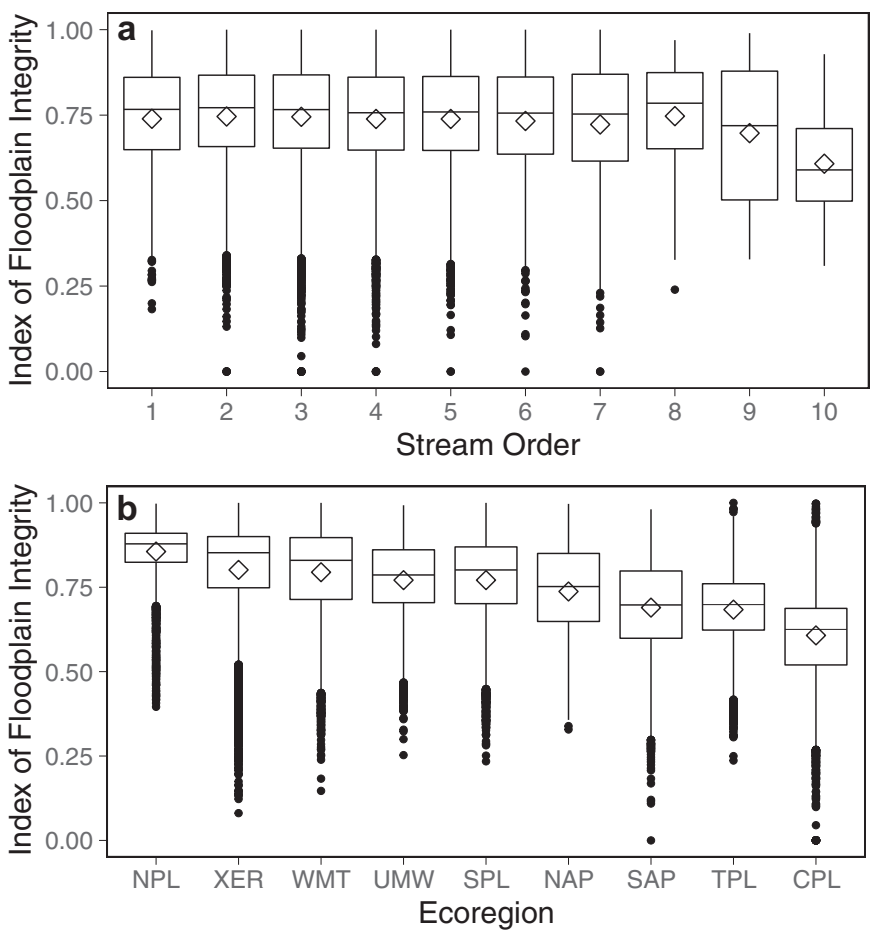


Fig. 4 Overall IFI values for across different stream orders and ecoregions. **a** Analysis of IFI by maximum stream order. **b** Analysis of IFI by aggregated Level III Ecoregions. (Key: NPL Northern Plains, XER Xeric, WMT Western Mountains, UMW Upper Midwest, SPL Southern Plains, NAP Northern Appalachians, SAP Southern Appalachians, TPL Temperate Plains, CPL Coastal Plain). The bottom of the boxes represents the 25th percentile and the top of the boxes represents the 75th percentile. The line through the box represents the median value, and the diamond points represent the mean value. Error bars above and below the boxes represent the 90th and 10th percentiles, respectively, and black dots represent data outliers.

Discussion

Applying the IFI methodology at the national scale allows for a quantitative measure of individual floodplain integrity relative to other floodplains across the country. The geospatial results of the integrity index provide a visual indication of the degree to which anthropogenic activity has impeded floodplain functionality. Since the IFI methodology provides quantitative information on floodplain conditions compared to other floodplains within the scope of study, those areas most in need of restorative efforts nationally can be identified. Additionally, the flexibility of the IFI methodology means that it can be repeated at smaller scales to gain a better understanding of the variability in floodplain functionality at a more localized level.

The observed spatial heterogeneity in IFI across the CONUS (Fig. 2) suggests that the IFI methodology is successful in identifying the floodplains most impacted by human development. It is not surprising to see that IFI values tend to be lower in higher order streams (Fig. 4), as this is historically where human development and anthropogenic activity are concentrated³⁵. This negative relationship between human development and floodplain integrity^{36–39} is also seen when comparing floodplains in rural areas to floodplains in urban areas (Fig. 3). The lower average IFI values observed in urban areas confirm that floodplain condition is degraded by human modifications.

The IFI methodology is limited by the availability, validity, and scope of the stressor datasets included in the index of integrity calculation. There are many other anthropogenic stressors that may exist within the floodplain not included in IFI methodology due to dataset unavailability or uncertainty driven by legacy effects. Such stressors may include the presence of pesticides, non-native vegetation, extirpation of beavers, historical removal of large wood, bank stabilization, and watershed land cover changes. The IFI methodology also only accounts for the presence of anthropogenic stressors within the floodplain itself and does not consider the impact of surrounding stressors that may impact floodplain functions. The IFI methodology provides a broad quantitative assessment of floodplain conditions on a large-scale, but the limitations of the methods used must be acknowledged when assessing the results of the integrity index.

The application of the IFI methodology at the CONUS scale allowed for a better understanding of the spatial variability in floodplain integrity across the CONUS. However, a more robust understanding of floodplain integrity at the localized level will be necessary before explicitly selecting the floodplains most in need of restorative efforts. The CONUS IFI results are intended to be used as a tool to understand the heterogeneity in floodplain functionality for the nation as a whole and do not provide a detailed picture of floodplain conditions on a localized level. However, since the methodology was designed to be adaptable and flexible, it can be repeated at these smaller scales to analyze floodplain functionality at a finer resolution.

In our comparison of CONUS IFI results to an independent floodplain functional analysis in the Pudget Sound basin, Washington, we found that our results may be overpredicting floodplain integrity for some functions. For instance, compared to Konrad's estimates of "flood storage" and "sediment regulation" in major rivers throughout the Pudget Sound basin, our functional IFI values were nearly always larger for all floodplain areas that overlapped in the two studies. However, comparable datasets for the floodplain function of "habitat provisioning" were more linearly related (Supplementary Table 4). Because we were required to lump discrete classification data from the Konrad study before making a comparison to our results, interpreting these findings can be challenging. But our comparison indicates that, depending on the floodplain function of interest, our results may be adequate to assess floodplain integrity, yet finer spatial

data of human stressors may be required in areas where our results predict especially high IFI values.

Our approach is also based on a negative linear relationship between stressor density and critical floodplain function, which is an oversimplification of the complex relationship between anthropogenic development and floodplain integrity. Non-linear floodplain stressor-function relationships may change functional and overall IFI results¹². Unfortunately, given the current literature on floodplain responses to human activities, we simply do not have the information necessary to determine the appropriate non-linear relationships. In addition, our understanding of floodplain processes is still limited regarding functional thresholds and alternate states that may occur due to human stressors in river corridors⁴⁰. Capturing these thresholds would require better knowledge of legacy effects, threshold criteria, and functional-stressor relationships. Ultimately, we chose to use a linear relationship to relate stressors to floodplain integrity for two reasons: (1) a linear relationship provides an unbiased estimate of declines in floodplain integrity such that it will estimate intermediate floodplain integrity values compared to values calculated using various non-linear relationships¹²; and (2) similar linear relationships were used in studies that evaluated watershed integrity^{32,41}. Our methodology could be revised with a more robust stressor density to floodplain integrity relationship¹². The type of relationship established between stressor prevalence and floodplain integrity may change the range and magnitude of values reflected in the resulting index, but the IFI methodology was designed to be iteratively improved upon and revised in this manner.

Methods

Floodplain delineation. We obtained floodplain boundaries from a 30 m resolution shapefile dataset of the 100-year undefended (without levees) floodplain across the U.S.²³. The floodplain boundaries in this shapefile were developed using a 2D hydrodynamic model and regionalized flood frequency estimates²³. We processed the floodplain shapefile by clipping it to the 2-digit hydrologic code Watershed Boundary Dataset for the CONUS⁴², removing isolated pixel groups less than 2700 m² (3 pixels), and filling gaps of 2700 m² or less. If the subdivision of the floodplain map resulted in an area of less than 2700 m², the area was removed. This reduced the overall floodplain area from 740,967 km² to 721,799 km² (approximately a 2.5% reduction).

We further divided the floodplain map along 12-digit hydrologic unit code boundaries (HUC12) located within the U.S. border. This process created 78,304 unique "floodplain units" within the CONUS, with a total area of 662,566 km² and a resulting average floodplain unit area of 8.46 km².

We determined the maximum stream order within each floodplain unit by using the National Hydrography Dataset Plus Version 2 (NHDplusV2) flowlines⁴³. For each streamline that intersected the floodplain, a unique COMID for that streamline was also associated with a floodplain unit. Due to discrepancies between the NHD flowlines dataset and the floodplain delineation developed by Wing et al.²³, 2,855 floodplain units do not intersect with any flowlines. This means that for these units there is not an associated maximum stream order.

Selection of anthropogenic stressor datasets. Of the datasets used by Karpach et al.¹² for the state of Colorado, we confirmed that the resolution of the data was at least that of the floodplain map (30 m) and publicly available at a national level. Only one of the previously selected datasets, groundwater wells, was unavailable at the national level. For this missing dataset, we identified a similar dataset at the national scale. It is important to note that only some datasets are direct measurements of their associated stressor, such as the representation of loss of wood and vegetation being measured directly by the forest loss cover events dataset⁴⁴. For other stressors, we used highly representative datasets due to the limited availability of direct measurement data at the national scale. An example of this is using the prevalence of groundwater wells in the floodplain as a representation of groundwater depletion.

A key stressor dataset used to numerically quantify floodplain integrity is a collection of data estimating the degree of hydrologic alteration for a range of metrics to the NHDplus V2 streamlines dataset. We associated the hydrologic alteration metrics for each maximum order streamline with the floodplain dataset. To account for hydrologic alteration within the index of integrity calculation we used the hydrologic alteration metric "alteration to mean annual maximum flows divided by catchment area" (MH20) as a representation of the peak flow conditions most likely to activate floodplains⁴⁵.

Calculation of stressor dataset densities. After the identification and evaluation of the representative datasets (Table 1), the density of each of these stressors within the floodplain unit had to be calculated. For point, polyline, and polygon datasets, the number (count/km²), length (km/km²), and area (km²/km²) of each stressor dataset within the floodplain unit were calculated. For raster data, the process for calculating density was dataset specific. We computed agricultural area as the percentage of cells in the floodplain reported as pasture/hay or cultivated crops (NLCD classes 81 and 82)²⁴. We computed developed area as the percentage of cells in the floodplain reported as low, medium, and high intensity development (NLCD classes 22, 23, and 24)²⁴. We calculated forest cover loss events as the percentage of cells in the floodplain that reported forest loss events between 2000 and 2020. Percent imperviousness was computed by averaging the percent imperviousness values reported for each 30 m cell for all cells in the floodplain unit. We quantified the prevalence of invasive species by computing the percentage of cells in the floodplain reported as non-native, introduced vegetation (LANDFIRE Existing Vegetation Type groups 701-709, 711, and 731)⁴⁶. We averaged the hydrologic alteration MH20 values⁴⁵ for all the maximum stream order segments within each floodplain to aggregate the hydrologic alteration metrics for the floodplain units. All stressor densities within the floodplain were calculated using R programming scripts (Code Availability).

Once we calculated the prevalence of stressors in the floodplain by the methods outlined above, the density values needed to be rescaled to comparable metrics across each dataset. Although the density values reported as percentages (e.g., km²/km²) have a potential maximum of one, the datasets measured by count and length have no theoretical maximum value. To address this issue, all stressor datasets were rescaled from zero to one, with a value of zero indicating the absence of the stressor in the floodplain, and one representing the 90th percentile of the stressor amount in the CONUS. For the three datasets that had no stressor present at the 90th percentile (canals and ditches, leveed area, and groundwater wells), the density values were scaled to the maximum observed value in the US. This stressor rescaling was done to create a consistent scale of comparison for stressor prevalence amongst all types of stressor datasets. Rescaling the stressor datasets relative to either the 90th percentile or the highest maximum observed value allows for a consistent comparison of floodplain integrity relative to other floodplains across the CONUS (Supplementary Fig. 2).

Calculation of functional IFI. We calculated functional IFI values for each of the five critical floodplain functions based on the rescaled stressor densities. Before computing the IFI value, we performed a Pearson correlation analysis⁴⁷ between each stressor dataset to avoid overweighing any individual data source (Supplementary Fig. 1). We found correlated datasets were linearly related, which is a required assumption of Pearson correlation analyses. For any functional IFI calculation that included two datasets with a correlation of over 0.7, only one dataset was included in the computation.

We computed the function IFI value using the following equation:

$$IFI_{i,k} = 1 - \sum_{j=1}^{n_{j,k}} \frac{S_{i,j}}{n_{j,k}} \quad (1)$$

where $IFI_{i,k}$ denotes the integrity value of the i^{th} floodplain unit for the k^{th} function, $S_{i,j}$ is the scaled stressor value in the i^{th} floodplain unit for the j^{th} stressor, $n_{j,k}$ is the number of stressor datasets, j , that impact the k^{th} function.

The results of the functional IFI computation produce a negative linear relationship between stressor density and function floodplain integrity. This method assumes an equal impact of each stressor dataset on floodplain functionality.

Calculation of overall IFI. We computed the overall integrity values for each floodplain unit by the geometric mean of the function specific IFI values:

$$IFI_i = \left(\prod_{k=1}^5 IFI_{i,k} \right)^{\frac{1}{5}} \quad (2)$$

where IFI_i denotes the overall integrity value of the i^{th} floodplain unit, and $IFI_{i,k}$ is the integrity value for the k^{th} function in the i^{th} floodplain unit.

Computing the overall IFI value by geometric mean of the function specific IFI values reflects the importance of each individual critical floodplain function to floodplain health. By this method, a function IFI value of zero as produces an overall IFI value of zero. This emphasizes that each of the five functions is essential to overall floodplain functionality. See Supplementary Table 1 for sample calculations.

Statistical analyses. We tested for normality of IFI distributions across categories (e.g., stream orders) using two tailed Kolmogorov-Smirnov tests ($\alpha = 0.05$) and found all groups of data were non-normal. We used non-parametric tests to evaluate significance between data categories, including Kruskal-Wallis and Dunn tests⁴⁷ (stream order and ecoregion datasets) and Mann-Whitney tests⁴⁷ (urban datasets). Significance results from the Dunn tests did not change when we applied Holm or Bonferroni adjustments. See the Supplementary Tables 5–6 for the complete results for our statistical analyses.

Comparison of IFI across the CONUS. Once we calculated the functional and overall IFI values for each floodplain, we associated each floodplain unit with a

variety of spatial attributes. Specifically, we associated each floodplain unit with three geospatial categories: (i) urban vs rural land cover³⁰; (ii) maximum Strahler stream order^{31,43}; and (iii) U.S. ecoregions³³.

We selected these three characteristics to analyze how IFI values vary regarding anthropogenic, hydrological, and ecological features in the CONUS. For the analysis of IFI by ecoregion, we used nine aggregated Omernik Level III ecoregions in the U.S.^{33,48}. These aggregated ecoregions were used in the National Rivers and Streams Assessment⁴⁹ and similar studies³², and the ecoregions were developed to minimize biological and hydrogeological differences within each region. We selected the aggregated ecoregions so that I could compare IFI values for regions with similar watershed ecology in the CONUS. Comparing IFI values by ecoregion allowed us to analyze the relationship between anthropogenic stressor prevalence, watershed ecology, and floodplain integrity.

We determined the intersection between each floodplain unit and the selected spatial attributes in ArcMap⁵⁰. We then analyzed the geospatial distribution of these attributes and the IFI results for the CONUS.

Comparison of IFI results to other datasets. Our IFI results were compared to two independent datasets: (1) index of watershed integrity (IWI) dataset produced by Thornbrugh et al.³², and (2) a geographic study by Konrad³⁴, which evaluated similar floodplain functions in the Pudget Sound basin of Washington State using discrete numerical categories.

We compared our overall IFI results to similar overall IWI results produced by Thornbrugh³² by graphically assessing linear relationships between the two datasets (see Karpack et al.¹² for a similar comparison).

We compared three functional IFI results to Konrad³⁴: (i) flood reduction; (ii) sediment regulation; and (iii) habitat provisioning. For this comparison we re-projected our results to match the coordinate reference system used by Konrad and cropped our study extent to the area evaluated by Konrad. Because Konrad identified discrete classifications of floodplain functions rather than functional gradients, such as this study, we grouped functional classifications based on descriptions provided by Konrad³⁴. For example, to compare flood reduction, we grouped classes 1 (connected, undeveloped high floodplain), 2 (connected, undeveloped low floodplain), and 3 (connected, undeveloped river area) of the “store and convey floods” function from Konrad. Similarly, to compare sediment regulation, we summed classes 1, 2, 3, and 4 of the “regulate sediment and wood supplies in river networks” function. To compare habitat provisioning, we grouped classes 1 and 2 of the “support forest ecosystems” function from Konrad. We also attempted to compare organic and solution regulation scores to the “retain and transform nutrients and contaminants” from Konrad, but we were not successful in accessing the data.

We derived a comparable score for each function based on the Konrad dataset by summing the total floodplain area associated with the appropriate classification and dividing by the total floodplain area within each HUC12. For instance, if half the floodplain area for the “store and convey floods” function was classified as a 1, 2, or 3, the total score for that floodplain was 0.5. This approach allowed us to compare our functional IFI scores to an equivalent score between 0 and 1 based on the Konrad data. We only included floodplain data from Konrad in our comparison if it overlapped fully or partially with our floodplain unit delineations. This resulted in the exclusion of approximately 370 floodplain units from the Konrad dataset. The total area of floodplain remaining for comparison was approximately 2338 km².

We compared our functional IFI scores for the three functions previously noted to the scores (ranging between 0 and 1) based on the Konrad data by graphically assessing their fit about a 1:1 equivalence line.

Data availability

All data from this study, including human stressor datasets and IFI calculations, are available in csv format at <https://doi.org/10.5061/dryad.prr4xgrk>.

Code availability

R code used in this study are available at <https://doi.org/10.5061/dryad.prr4xgrk>. Code was run using version 4.1.3 and the R packages sf (v. 1.0), terra (v. 1.7), dplyr (v. 1.1.2), tictoc (v. 1.2), readr (v. 2.1.4), and purrr (v. 1.0.1).

Received: 26 October 2022; Accepted: 2 June 2023;

Published online: 15 June 2023

References

1. Opperman, J. J., Luster, R., McKenney, B. A., Roberts, M. & Meadows, A. W. Ecologically functional floodplains: Connectivity, flow regime, and scale. *JAWRA J. Am. Water Resour. Assoc.* **46**, 211–226 (2010).
2. Wohl, E. An integrative conceptualization of floodplain storage. *Rev. Geophys.* **59**, e2020RG000724 (2021).

3. Boulton, A. J., Detry, T., Kasahara, T., Mutz, M. & Stanford, J. A. Ecology and management of the hyporheic zone: stream–groundwater interactions of running waters and their floodplains. *J. North Am. Benthol. Soc.* **29**, 26–40 (2010).
4. Brunke, M. & Gonsler, T. The ecological significance of exchange processes between rivers and groundwater. *Freshw. Biol.* **37**, 1–33 (1997).
5. Fryirs, K. (Dis)Connectivity in catchment sediment cascades: a fresh look at the sediment delivery problem. *Earth Surf. Process. Landf.* **38**, 30–46 (2013).
6. Hopkins, K. G. et al. A method to quantify and value floodplain sediment and nutrient retention ecosystem services. *J. Environ. Manage.* **220**, 65–76 (2018).
7. Olde Venterink, H. et al. Importance of sediment deposition and denitrification for nutrient retention in floodplain wetlands. *Appl. Veg. Sci.* **9**, 163–174 (2006).
8. Amoros, C. & Bornette, G. Connectivity and biocomplexity in waterbodies of riverine floodplains. *Freshw. Biol.* **47**, 761–776 (2002).
9. Tockner, K., Pennetzdorfer, D., Reiner, N., Schiemer, F. & Ward, J. V. Hydrological connectivity, and the exchange of organic matter and nutrients in a dynamic river–floodplain system (Danube, Austria). *Freshw. Biol.* **41**, 521–535 (1999).
10. Fischer, C. et al. The “Habitat Provision” index for assessing floodplain biodiversity and restoration potential as an ecosystem service—method and application. *Front. Ecol. Evol.* **7**, 483 (2019).
11. Tomscha, S. A., Gergel, S. E. & Tomlinson, M. J. The spatial organization of ecosystem services in river–floodplains. *Ecosphere* **8**, e01728 (2017).
12. Karpach, M. N., Morrison, R. R. & McManamay, R. A. Quantitative assessment of floodplain functionality using an index of integrity. *Ecol. Indic.* **111**, 106051 (2020).
13. Johnson, K. A. et al. A benefit–cost analysis of floodplain land acquisition for US flood damage reduction. *Nat. Sustain.* **3**, 56–62 (2020).
14. Tockner, K. & Stanford, J. A. Riverine flood plains: present state and future trends. *Environ. Conserv.* **29**, 308–330 (2002).
15. Opperman, J. J., Moyle, P. B., Larsen, E. W., Florsheim, J. L. & Manfree, A. D. *Floodplains: Processes and Management for Ecosystem Services*. (Univ of California Press, 2017).
16. Knox, R. L., Wohl, E. E. & Morrison, R. R. Levees don’t protect, they disconnect: A critical review of how artificial levees impact floodplain functions. *Sci. Total Environ.* **837**, 155773 (2022).
17. Knox, R. L., Morrison, R. R. & Wohl, E. E. A river ran through it: Floodplains as America’s newest relict landform. *Sci. Adv.* **8**, eabo1082 (2022).
18. Wohl, E., Lininger, K. B. & Baron, J. Land before water: The relative temporal sequence of human alteration of freshwater ecosystems in the conterminous United States. *Anthropocene* **18**, 27–46 (2017).
19. Rajib, A. et al. The changing face of floodplains in the Mississippi River Basin detected by a 60-year land use change dataset. *Sci. Data* **8**, 271 (2021).
20. Graf, W. L. Damage control: Restoring the physical integrity of America’s rivers. *Ann. Assoc. Am. Geogr.* **91**, 1–27 (2001).
21. Munné, A., Prat, N., Solà, C., Bonada, N. & Rieradevall, M. A simple field method for assessing the ecological quality of riparian habitat in rivers and streams: QBR index. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **13**, 147–163 (2003).
22. Ward, J. V., Tockner, K. & Schiemer, F. Biodiversity of floodplain river ecosystems: ecotones and connectivity. *Regul. Rivers Res. Manag.* **15**, 125–139 (1999).
23. Wing, O. E. J. et al. Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resour. Res.* **53**, 7968–7986 (2017).
24. Dewitz, J. National Land Cover Database (NLCD) 2019 Products. <https://doi.org/10.5066/P9KZCM54> (2021).
25. U.S. Army Corps of Engineers. National Levee Database. (2019).
26. Kesel, R. H. Human modifications to the sediment regime of the Lower Mississippi River flood plain. *Geomorphology* **56**, 325–334 (2003).
27. Remo, J. W. F., Ickes, B. S., Ryherd, J. K., Guida, R. J. & Therrell, M. D. Assessing the impacts of dams and levees on the hydrologic record of the Middle and Lower Mississippi River, USA. *Geomorphology* **313**, 88–100 (2018).
28. Knox, R. L., Morrison, R. R. & Wohl, E. E. Identification of artificial levees in the contiguous United States. *Water Resour. Res.* **58**, e2021WR031308 (2022).
29. Mazzoleni, M. et al. Floodplains in the anthropocene: A global analysis of the interplay between human population, built environment, and flood severity. *Water Resour. Res.* **57**, e2020WR027744 (2021).
30. US Census Bureau. TIGER/Line Shapefiles. <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html> (2019).
31. Strahler, A. N. Hypsometric (area–altitude) analysis of erosional topology. *Geol. Soc. Am. Bull.* **63**, 1117–1142 (1952).
32. Thornbrugh, D. J. et al. Mapping watershed integrity for the conterminous United States. *Ecol. Indic.* **85**, 1133–1148 (2018).
33. US EPA. Ecoregions used in the National Aquatic Resource Surveys. <https://www.epa.gov/national-aquatic-resource-surveys/ecoregions-used-national-aquatic-resource-surveys> (2016).
34. Konrad, C. P. *Geospatial assessment of ecological functions and flood-related risks on floodplains along major rivers in the Puget Sound Basin, Washington*. 1–28 <https://pubs.er.usgs.gov/publication/sir20155033> (2015).
35. Böck, K., Polt, R. & Schülting, L. Ecosystem Services in River Landscapes. in *Riverine Ecosystem Management: Science for Governing Towards a Sustainable Future* (eds. Schmutz, S. & Sendzimir, J.) 413–433 (Springer International Publishing, https://doi.org/10.1007/978-3-319-73250-3_21 2018).
36. Hupp, C. R., Pierce, A. R. & Noe, G. B. Floodplain geomorphic processes and environmental impacts of human alteration along Coastal Plain rivers, USA *Wetlands* **29**, 413–429 (2009).
37. Peipoch, M., Brauns, M., Hauer, F. R., Weitere, M. & Valett, H. M. Ecological simplification: Human influences on riverscape complexity. *BioScience* **65**, 1057–1065 (2015).
38. Bunn, S. E. & Arthington, A. H. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environ. Manage.* **30**, 492–507 (2002).
39. Smucker, N. J. & Detenbeck, N. E. Meta-analysis of lost ecosystem attributes in urban streams and the effectiveness of out-of-channel management practices. *Restor. Ecol.* **22**, 741–748 (2014).
40. Livers, B., Wohl, E., Jackson, K. J. & Sutfin, N. A. Historical land use as a driver of alternative states for stream form and function in forested mountain watersheds of the Southern Rocky Mountains. *Earth Surf. Process. Landf.* **43**, 669–684 (2018).
41. Flotemersch, J. E. et al. A watershed integrity definition and assessment approach to support strategic management of watersheds. *River Res. Appl.* **32**, 1654–1671 (2016).
42. Seaber, P. R., Kapinos, F. P. & Knapp, G. L. *Hydrologic unit maps*. <http://pubs.er.usgs.gov/publication/wsp2294> <https://doi.org/10.3133/wsp2294> (1987).
43. Moore, R. B. et al. *User’s guide for the national hydrography dataset plus (NHDPlus) high resolution*. 66 (2019).
44. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
45. McManamay, R. A., George, R., Morrison, R. R. & Ruddell, B. L. Mapping hydrologic alteration and ecological consequences in stream reaches of the conterminous United States. *Sci. Data* **9**, 450 (2022).
46. Rollins, M. G. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *Int. J. Wildland Fire* **18**, 235 (2009).
47. Zar, J. H. *Biostatistical analysis*. (Prentice Hall, 1999).
48. Herlihy, A. T. et al. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. *J. North Am. Benthol. Soc.* **27**, 860–877 (2008).
49. National Rivers and Streams Assessment 2013–2014: A Collaborative Survey. <https://www.epa.gov/national-aquatic-resource-surveys/nrsa> (2020).
50. Environmental Systems Research Institute (ESRI). ArcMap Release 10.x. (2011).
51. OpenStreetMap. (2022).
52. Horizon Systems Coporation. NHDPlus. (2012).
53. US Geological Survey. Groundwater Watch. (2022).

Acknowledgements

This work was supported by funding from the National Science Foundation (awards 2115169 and 2142761) and Colorado Water Center.

Author contributions

R.R.M.: conceptualization, methodology, formal analysis, writing, reviewing, editing. K.S.: formal analyses, writing, reviewing, editing. R.A.M.: conceptualization, reviewing. D.C.: formal analysis, reviewing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43247-023-00877-4>.

Correspondence and requests for materials should be addressed to Ryan R. Morrison.

Peer review information *Communications Earth & Environment* thanks Maurizio Mazzoleni, Wendy Monk and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Claudia Zoccarato and Joe Aslin. A peer review file is available.

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) 2023