DISSERTATION

THE GEOGRAPHY OF ARTIFICIAL LEVEES IN THE UNITED STATES

Submitted by

Richard Leo Knox

Department of Geosciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2022

Doctoral Committee:

Advisor: Ellen Wohl
Co-Advisor: Ryan Morrison

Melinda Laituri
Sara Rathburn

ABSTRACT

THE GEOGRAPHY OF ARTIFICIAL LEVEES IN THE UNITED STATES

Connectivity between different parts of the landscape is an important theme for river ecosystem functions. Recent advances in conceptual models of river ecosystems, computing power, and data availability, resolution, and extent have allowed the exploration of this theme at continental and global scales. However, these studies have not included the impacts of artificial levees on floodplain function and extent due to the lack of complete artificial levee databases. Local and regional studies have explored the harmful effects and identification of artificial levees. Several characteristics of artificial levees have inhibited the extension of these studies to greater spatial scales (e.g., artificial levees are shaped like other natural and anthropogenic features; artificial levee height and width are small compared to the vertical and horizontal resolution and accuracy of earth observation data available at continental and global scales; artificial levees have a long history of construction). I first present a methodology and data set for the identification of artificial levees in a case study of seven basins (ranging in size from 1,700 km^2 to 8,000 km^2 each) in the continental United States (CONUS) and then apply the methodology to the entire CONUS. This methodology, which includes a model that only uses land cover, distance from stream flow, and basin variables, detected over 182,000 km of artificial levees. Next, I use this dataset in combination with a pre-existing artificial levee database to determine how artificial levees influence floodplain extent, land cover, and association with stream order size in the CONUS. Surprisingly, this revealed that the 100-year CONUS floodplain was of greater extent with artificial levees than if they were not constructed. And not surprisingly, the 8,100 km^2 of CONUS floodplain that are disconnected by artificial levees are predominantly cultivated or developed land cover. Finally, I conduct a critical review of floodplain functions and analyze case studies of floodplain restoration involving the alteration of artificial levees. I define five interconnected floodplain functions that are vital to river ecosystems and are adversely impacted by artificial levee construction. Studies that analyze floodplain restoration are heavily concentrated in North America and Europe and evaluate effects within 30 years of restoration. In the United States, this type of restoration impacts less than 1% of river kilometers with artificial levees and 1-2% of disconnected floodplains. This dissertation provides an important advance in understanding the impacts of artificial levees on floodplain extent and function at a large spatial scale. It also provides several avenues for continued research.

ACKNOWLEDGEMENTS

I am particularly thankful for being alive, the wonderful people who have been put into my life, and for the orderliness of the created world. Specifically, and in no order of importance, I'd like to thank several people or groups of people. Thanks to Mom and Dad for giving me life, loving me, and teaching me so much throughout the years. Thank you to Tim and Marcia, my second set of parents, for their love and support. Thank you to my spouse, Michelle, for supporting me and being the best partner I can imagine. Thank you to my children, Leo, Bridget, Teresa, Isaac, and Charles, who are each wonderful in their own unique way and provide so much hope for the future. Thank you to the friends and families we've met in Fort Collins who have treated us with love, as if we've always lived here, and weren't leaving in just three short years. I am particularly grateful to Joseph and his family. I'm thankful for COL(R) Andrew Lohman who decided to send me to graduate school.

Thank you to my advisor, Dr. Ellen Wohl, who has been such a wonderful (and patient) advisor. I appreciate the bread crumbs she began leaving for me beginning fifteen years ago or earlier. Thank you to my co-advisor Dr. Ryan Morrison, whose attention to detail and striving for perfection inspires me. Thank you to my committee members, Drs. Sara Rathburn and Melinda Laituri, for their helpful input during key periods of my research. I'm thankful to the many members of the fluvial geomorphology lab who are always so willing to stop working on something to discuss a question I have.

I thank Tracie-Lynn Nadeau, Shane Cline, Doug Thompson, and Francis Magilligan for suggesting basins to consider for the case study in chapter 2. Kevin McNinch's assistance in providing a digital copy of the NHD plus was critical. I thank Oliver Wing for providing

MATLAB code to generate geomorphic variables. Antonio Annis's help in running GFPLAIN in ArcGIS Pro was extremely helpful. Paul Hudson's willingness to offer recommendations were very helpful. Anonymous reviewers improved Chapters 2 and 3 of this dissertation by providing helpful comments and suggestions. Ann Hess's statistical advice was invaluable.

DEDICATION

I dedicate this dissertation to my companions on this great adventure: Michelle, Leo, Bridget, Teresa, Isaac, and Charles.

TABLE OF CONTENTS

ABSTRACT ... ii
ACKNOWLEDGEMENTS ... iv
DEDICATION ... vi
LIST OF TABLES ... ix
LIST OF FIGURES ... x
Chapter 1 Introduction ... 1
Chapter 2 Identification of Artificial Levees in the Contiguous United States ... 4
2.1 Introduction ... 5
2.2 Materials and Methods ... 10
2.2.1 Overview ... 10
2.2.2 Site Descriptions ... 11
2.2.3 Hydrogeomorphic Floodplain Delineation ... 14
2.2.4 Data ... 15
2.2.5 Model Development and Implementation ... 16
2.3 Results ... 21
2.3.1 Case Studies ... 21
2.3.2 Leave-one-out cross-validation ... 22
2.3.3 National study ... 23
2.4 Discussion ... 30
2.4.1 Location and prevalence of artificial levees ... 30
2.4.2 Spatial and geographic implications of findings ... 31
2.4.3 Future directions ... 34
2.5 Conclusions ... 35
Chapter 3 A river ran through it: floodplains as America's newest relict landform ... 37
3.1 Introduction ... 37
3.2 Results ... 40
3.2.1 Area analysis ... 40
3.2.2 Land cover analysis ... 42
3.2.3 Stream order analysis ... 43
3.3 Discussion ... 44
3.4 Materials and Methods ... 50

3.4.1 Experimental Design 50
3.4.2 Topography Modification 50
3.4.3 Statistical Analysis 51
Chapter 4 Levees don’t protect, they disconnect: how artificial levees impact floodplain function 53
4.1 Introduction 54
4.2 Why it matters – floodplain functions 58
4.2.1 Material fluxes. 58
4.2.2 Enhanced spatial heterogeneity of hydrology and biogeochemistry. 61
4.2.3 Enhanced habitat abundance and diversity. 62
4.2.4 Enhanced biomass and biodiversity. 65
4.2.5 Hazard mitigation. 67
4.3 Synthesis discussion- What do case studies tell us? 68
4.3.1 Limits of selected case studies from English-language journals. 75
4.3.2 Reconnection and reconfiguration. 78
4.3.3 Restoration scale dilemma. 79
4.3.4 Unique place-based challenges. 83
4.4 Conclusion 84
Chapter 5 Conclusion 86
References 91
Appendix A: Supplemental information for Chapter 2: Identification of Artificial Levees in the Contiguous United States 127
Appendix B: Supplemental information for Chapter 3: A river ran through it: floodplains as America’s newest relict landform 133

LIST OF TABLES

Table 2.1. Attributes of each basin used in the case study. 13

Table 2.2. Variables and datasets used in the study... 16

Table 2.3. Model names and variables in the study.. 18

Table 2.4. Model performance in the national study 24

Table 2.5. NLD and potential levee lengths (km) by HUC2 basin... 30

Table 2.6. Model performance in the Lower Mississippi HUC2 basin 35

Table 4.1. Selected case studies of floodplain restoration in North America.............................. 70

Table 4.2. Selected case studies of floodplain restoration in Europe and Japan. 72

Table 4.3. Selected case studies of simulations 81

LIST OF FIGURES

Figure 2.1. Study location 12

Figure 2.2. Workflow for the case and national studies 17

Figure 2.3. Model performance 21

Figure 2.4. RF model performance 22

Figure 2.5. Results from the leave-one-out cross-validation. 23

Figure 2.6. Potential and NLD artificial levees 27

Figure 2.7. Potential and NLD artificial levee stream orders 27

Figure 2.8. Aerial and perspective images of potential artificial levees 29

Figure 3.1. Net disconnection and cumulative alteration in the CONUS 41

Figure 3.2. Land cover of anthropogenically connected and disconnected areas........................ 42

Figure 3.3. Stream order analysis 44

Figure 3.4. Two examples of floodplain alteration before and after artificial levee installation.. 47

Figure 4.1. Floodplain functions and selected rivers with case studies 57

Figure 4.2. Selected case study location. 76

Chapter 1 Introduction

The story of human civilization on floodplains is one that began at least 7,000 years ago (Butzer, 1976). Throughout our long history together, humans have modified floodplains in numerous ways through indirect alterations (outside the river corridor) and direct alterations (within the river corridor). There is a growing awareness that these alterations have resulted in simplified floodplains with reduced functions (Peipoch et al., 2015). There is also growing awareness that diminished floodplain functions are detrimental to humans because our involvement on floodplains is based on exploiting the very same functions (Tockner and Stanford, 2002).

Development in large floodplain rivers is in conflict with restoration and conservation efforts across the world (Sparks, 1995; Wohl, Bledsoe, et al., 2015). In the United States alone, floodplain development during the 20th century was so great that even after the expenditure of billions of dollars on flood defenses (mainly, artificial levees) in the second half of the century, flood-related economic losses more than doubled. Unfortunately, the loss of floodplain functions resulting from flood defenses is rarely calculated when considering economic losses related to flooding (Opperman et al., 2009; Jacobson et al., 2015). That the loss of floodplain function is little appreciated or understood stems in part from a lack of awareness about floodplain ecosystem complexity and floodplain connection to the atmosphere, channel, subsurface, and uplands (Gren et al., 1995).

Connectivity, the degree to which matter and organisms are able to move across areas in a landscape (Wohl, 2017), is a central theme for the role of floodplains in conceptual models of river ecosystems (e.g., the river continuum concept (Vannote et al., 1980); the flood pulse concept (Junk et al., 1989)). Understanding different aspects of connectivity (e.g., longitudinal, lateral, vertical) to off-channel environments is critical to understanding ecosystem functions (Ward, 1989; Kondolf et al., 2006; Harvey and Gooseff, 2015). Lateral connectivity describes the fluxes between channel and adjacent riparian and floodplain landforms (Ward, 1989; Covino, 2017).

Direct alterations, which include flow regulation and artificial levee construction (Wohl, 2018), either increase or decrease connectivity in river ecosystems. While most studies on connectivity focus on the local scale (e.g., Gergel et al., 2002), several national-scale studies have exposed the vast degree of harmful impacts to longitudinal connectivity from dams (e.g., Graf, 1999; Lehner et al., 2011; Grill et al., 2019; Jones et al., 2019) and to lateral connectivity from roads and railroads (e.g., Blanton & Marcus, 2009). Despite the cataloguing of harmful effects derived from artificial levees on floodplains and river ecosystems at local scales (e.g., increasing stage elsewhere (Criss and Shock, 2001), inducing bed incision (Frings et al., 2009), limiting nutrient exchange between floodplain and channel (Junk et al., 1989), and promoting development (White et al., 2001)), there has been no national-scale assessment of how artificial levees have altered lateral connectivity (Wohl, 2017) that is similar to Graf's national-scale assessment of the effects of dams on longitudinal connectivity (1999, 2001). Two reasons for this are the lack of a complete national-scale database of artificial levees and the difficult nature of identifying artificial levees.

This dissertation seeks to address the gaps in knowledge regarding artificial levee impacts to floodplain extent and function across the continental United States (CONUS) by addressing the following objectives: 1) identify the locations of potential artificial levees in the CONUS using existing levee databases as training data by first testing variables, sampling strategies, and model types in a case study before applying the best performing model to the entire CONUS (Chapter 2); 2) evaluate the impacts of artificial levees on floodplain extent in the CONUS by first deleting artificial levees from topography, then using two hydrogeomorphic flood models based on unmodified and modified topography to estimate flood extent, and finally comparing the resulting scenarios in relation to land cover, extent, and stream order (Chapter 3); and 3) define and review floodplain functions and understand how they are impacted by artificial levees by analyzing case studies of floodplain restoration involving artificial levee alteration (Chapter 4). In Chapter 5, I summarize the key findings of this work and suggest future research related to artificial levee impacts on floodplains.

Chapter 2 Identification of Artificial Levees in the Contiguous United States

Summary

Artificial levees are anthropogenic structures designed to hydrologically disconnect rivers from floodplains. The extent of artificial levees in the contiguous United States (CONUS) is unknown. To better estimate the distribution of artificial levees, I tested several different geomorphic, land cover, and spatial variables developed from the National Elevation Dataset, the National Land Cover database, and the National Hydrography Dataset HR Plus. I used known levee locations from the National Levee Database as training data. I tested machine learning and general logistic models' ability to detect artificial levees in a 100-year hydrogeomorphic floodplain of seven geographically diverse 8-digit HUC basins. Random forest models outperformed other models in predicting the location of levees using variables representing geomorphic attributes, land cover, and distance from streams ranging in size between stream order one through six. To demonstrate the ability of my approach to detect unknown levees, I conducted a leave-one-out cross-validation in the lower Mississippi Basin using approximately 1,100 artificial levees. This approach detected known levees constituting 94% of the total levee length in the basin. Scaling up to the CONUS, I applied a high performing (overall accuracy of 97%) random forest model using land cover and stream order variables. I detected 182,213 km of potential levees, mostly along streams of order 2-6 in the Mississippi and Missouri River Basins, indicating that the national levee database contains 20.4% of levee length. Potential levees and those documented in the national levee database modify 2% of the total length of streams in the contiguous United States (Knox et al., 2022a).

2.1 Introduction

John Barry described the Mississippi River as pulsing "like the artery of the American heartland" (Barry, 2007). Recognized as the world's most engineered megariver (Knox and Latrubesse, 2016), the Mississippi River is emblematic of alterations made to United States (U.S.) rivers during the last three hundred years. With an estimated 98% of the nation's 5.3 million km of rivers impacted by human activities (Graf, 2001), it is difficult to understate the degree of human modification to U.S. rivers. Direct alterations take place within river corridors (which I define as the active channel(s) and floodplain) and include flow regulation, channel engineering (e.g. straightening or dredging), placer and aggregate mining, beaver trapping, floodplain draining, and levee construction (Wohl, 2018).

During recent decades, rivers have been increasingly appreciated as ecosystems worthy of preservation and restoration (Graf, 2001, Bunn et al., 2010; Palmer et al., 2014; Castro and Thorne, 2019). Understanding the importance of river ecosystems across a broad spectrum of functions requires that we recognize the importance of longitudinal, lateral, and vertical connectivity to off-channel environments (Ward, 1989; Kondolf et al., 2006; Harvey and Gooseff, 2015). *Connectivity* describes the degree to which matter and organisms are able to move across areas in a landscape (Wohl, 2017). *Lateral connectivity*, which describes fluxes between the channel and adjacent riparian and floodplain landforms (Ward, 1989; Covino, 2017), reflects processes that create channel and floodplain topography and stratigraphy (Brooks, 2003), and thus aquatic and riparian habitats (Junk et al., 1989; Ward et al., 1999; Blanton & Marcus, 2009; Pennington et al., 2010).

Diverse human alterations of river form and processes either increase or limit connectivity in river ecosystems. Most studies on connectivity focus on the local scale (e.g., Briggs et al., 2013), although national-scale studies have assessed the effects of dams on longitudinal connectivity (e.g. Graf, 1999; Jones et al., 2019) and roads and railroads on lateral connectivity (e.g. Blanton & Marcus, 2009). One anthropogenic feature that adversely impacts lateral connectivity is *artificial levees*, which can be defined as raised linear features built between active channels and floodplains to contain peak flows in the channel (Tobin, 1995). Although artificial levees can strongly influence lateral connectivity within river corridors, their national- to global-scale effects have not been quantified in a manner similar to the effects of dams (e.g., Lehner et al., 2011; Grill et al., 2019).

Local and regional studies have found that artificial levees influence river hydrology by increasing stage at and upstream from levee locations and increasing downstream conveyance and flooding beyond the levees (Tobin, 1995; Criss and Shock, 2001; Heine and Pinter, 2012; Czech et al., 2015). Levees alter channel processes by inducing bed coarsening and incision caused by increased channel velocities (Frings et al., 2009). Levees also limit lateral connectivity and the exchange of nutrients, sediment, and organisms between the channel and floodplain, resulting in significant ecological harm (Blanton and Marcus, 2009; Sparks et al., 2017; Wohl, 2018). In the context of human societies, the presence of artificial levees can promote floodplain development and increase the vulnerability of populations and infrastructure to flood damage (White et al., 2001; Pinter, 2005). On the other hand, levees can be viewed as effective flood protection measures within their intended design standards and are uncomplicated and inexpensive to build (Tobin, 1995).

Artificial levee construction began in the U.S. in the early 1700s with landowners living alongside the lower Mississippi River (Wohl, 2005). By the 1800s, the majority of the basin was leveed with patchworks managed by individuals and communities in nascent levee boards

(Hudson et al., 2008). Beyond the lower Mississippi River basin, the 20th century became the era of federal levee construction with funding provided by the federal government and other entities (Wohl, Lininger, & Baron, 2017). Levee construction increased in scope well into the 20th century, especially in the Midwestern and Eastern US, as the federal government and the US Army Corps of Engineers (USACE) became more involved in mitigating natural disasters (Wohl, 2005). Many levees were built in the Mississippi basin after the 1927 and 1937 floods and in California after flooding in 1907 and 1909 (ASCE, 2017). The nation's focus on artificial levees as a prime flood protection tool, resulting in billions of dollars expended to construct thousands of kilometers of levee, has been described as a "levee love affair" (Tobin, 1995).

The length of artificial levees in the U.S. is unknown but estimates range between 48,000 and 167,000 km, corresponding to coverage of roughly 1% and 3% of total estimated river km in the contiguous US (Heine and Pinter, 2012; ASCE, 2017). The USACE started a national levee inventory in 2006, which resulted in the National Levee Database (NLD). Each levee is annotated by a line representing the levee crest and varying amounts of metadata. The NLD is currently estimated to be 30% complete (ASCE, 2017), but a comprehensive evaluation of the NLD's thoroughness has not been completed (Wing et al., 2017). Consequently, there is no national-scale assessment of how artificial levees have altered lateral connectivity on U.S. rivers (Wohl, 2017) analogous to Graf's national-scale assessments of the effects of dams on river longitudinal connectivity (1999, 2001).

Because artificial levees are constructed by people and share common morphological characteristics, many levees should be recognizable and distinguishable by their shape (Brown et al., 2017). Recent advances over the last two decades in the availability of high-resolution topography have revolutionized the ability to study landscapes (Passalacqua et al., 2015).

Accordingly, nearly every study on the identification of artificial levees has exclusively used topography or topographic-derived geomorphic variables with the exception of two studies that used spectral signatures (Steinfeld and Kingsford, 2013; Steinfeld et al., 2013). Identification of levees at regional scales has used maximum curvature, entropy, and residual topography from lidar digital terrain models (Sofia et al., 2014). At the national scale, Wing et al. (2019) used geomorphic variables derived from the National Elevation Dataset (NED) to determine which geomorphic features are important to retain in a hydrodynamic flood model during DEM coarsening, although the study was not specifically intended to recognize artificial levees.

Researchers have used different methods of analyzing variables to predict artificial levee location, including logistic regression and image segmentation (Steinfeld et al., 2013), statistical analysis (Sofia et al., 2014), hillshade and wavelet analysis (Czuba et al., 2015) and visual inspection (Steinfeld et al., 2013; Czuba et al., 2015). Although not yet applied to identifying artificial levees, the modeling capabilities of machine learning techniques make them a suitable application in the geosciences (Lary et al., 2016). Machine learning techniques, which include decision trees, neural networks, and support vector machines, are especially effective at recognizing patterns in complex data or in scenarios where the underlying principles are poorly understood (Valentine and Kalnins, 2016). Random forest modeling is a supervised machine learning technique that builds decision trees and predicts category labels, is relatively fast, and can result in high predictive accuracy compared to other machine learning techniques (Breiman, 2001; Choi et al., 2020). The high accuracy of random forest models compared to other machine learning techniques has made random forest a popular choice for detecting surface and subsurface anthropogenic features as well as hydrological predictions (Deines et al., 2017; Cho et al., 2019). However, the validity of methods used to detect artificial levees in previous studies and machine learning techniques to detect artificial levees at the regional or national scale are unknown.

I aim to improve upon these previous studies by employing and testing different categories of data (i.e., geomorphic, land cover type, and distance from stream) and different types of models (general logistic models (GLM), random forest models (RF), and support vector machine models (SVM)) to the specific problem of identifying artificial levee locations. I generated a calibrated 100-year hydrogeomorphic floodplain as the extent of analyses because this floodplain contains most known levee locations and at 10% of the contiguous U.S. (CONUS) area, reduces the computational requirements of the CONUS analysis. I use a case study of seven different 8-digit hydrologic unit code (HUC8) watersheds (Seaber et al., 1987), to test the effects of sample size, ratio of levee to non levee data, and model variables on model accuracy. I use known levee locations from the NLD to generate data from levee and non levee locations for training and validation. I conduct a leave-one-out cross-validation on 1,171 levees in the lower Mississippi Basin to understand how a highly accurate RF model detects undocumented levees. At the CONUS level, I generate two different sized datasets (n = 30,600 and 3,060,000) for training and validation and test the accuracy of a selection of models and variables based on the results of the case study. I apply the most accurate model (trained on data sampled from 2,142,000 locations across the CONUS) to the entire CONUS floodplain and analyze the results to determine the location, length, and stream order association of potential levees that are not identified by the NLD. My primary objective is to estimate the locations, spatial distribution, and stream order association of artificial levees across the contiguous U.S., especially as they relate to the completeness of the NLD.

I organize this work as follows. First, I provide a site description and justification of the seven HUC8 watersheds I chose as case studies along with information about our modeling approach and the data I used in the case study, in the leave-one-out cross validation, and in the CONUS study. I then present results using different accuracy measures. Finally, I discuss implications for understanding the completeness of the NLD and ideas for future research.

2.2 Materials and Methods

2.2.1 Overview

In general, modeling included the following major steps. First, I chose seven diverse 8-digit HUC basins for a case study (Figure 2.1). I then created a CONUS-scale 100-year hydrogeomorphic floodplain using GFPLAIN at a 30m resolution (Nardi et al., 2019) in ArcGIS Pro (ESRI Inc., 2021) (Figure A.1). I used this floodplain as the studies' geographic extent. For each case study basin, I created multi-layered raster files with each variable represented by a layer. I used R software (R Core Team, 2020) and an NLD shapefile to test different variables, sample sizes, machine learners, and GLMs. Then, using R and ArcGIS Pro I conducted a leave-one-out cross-validation using ~1,100 artificial levees in the HUC2 lower Mississippi River basin (Figure 2.1) to determine how the model predicts undocumented levees. I then applied my approach to the entire CONUS (Figure 2.1) and generated two differently sized data sets using Google Earth Engine (GEE) (Gorelick et al., 2017) and ArcGIS Pro. Finally, I applied the most accurate model trained from the larger CONUS data set to the CONUS floodplain, creating a prediction surface map. This map was then segmented and compared to the NLD to determine which artificial levees were not documented.

Throughout, I used a full suite of accuracy metrics to assess model performance such as true and false rates and confusion matrices, when possible. In the case study, I used Cohen's kappa coefficient, which assesses interclassifier agreement and is sensitive to class prevalence (Fitzgerald and Lees, 1994) to compare the performance of nearly 1,000 models. This coefficient, like all measures, is an imperfect index of overall accuracy (Foody, 2020). However, the comparison of many models using multiple measures of performance is unfeasible. To mitigate this, I test sample size, sampling ratio, model variables and type again in the national study, using multiple accuracy metrics to select which model to apply to the CONUS.

2.2.2 Site Descriptions

For the case studies, I chose seven distinct HUC8 watersheds across the 48 contiguous states to represent a wide variety of geographic, land cover, and hydrologic conditions (Figure 2.1, Table 2.1). The seven basins range in size from the 1,700 km^2 lower Columbia River in the Pacific Northwest to the 7,900 km^2 Little Snake River on the Colorado-Wyoming border. Climates of the seven locations range from the semi-arid Middle Kern River in southern California to the humid subtropical lower Red River in southern Louisiana and the humid continental Hudson River in upstate New York. Relief ratios, the dimensionless ratio of the total vertical elevation difference in a basin divided by the basin length (Schumm, 1956), are an especially important consideration because artificial levees are distinguishable from other locations as particularly steep and rough terrain in the alluvial plains of the Midwest (e.g., 0.001 relief ratio of the Middle Wabash along the Illinois – Indiana border and the 0.002 relief ratio of the Middle Arkansas in Kansas) but seem more flat in comparison in the mountains of California or New York (e.g., 0.031 and 0.013 relief ratios of the Middle Kern and the Hudson, respectively). The primary land cover of each basin (Table 2.1) varies between cultivated crops (e.g., the Middle Wabash and the Middle Arkansas), to herbaceous or shrub/scrub (e.g., the

Middle Kern and the Little Snake), to woody wetlands or forest (e.g., the Lower Red, the Hudson, the Columbia).

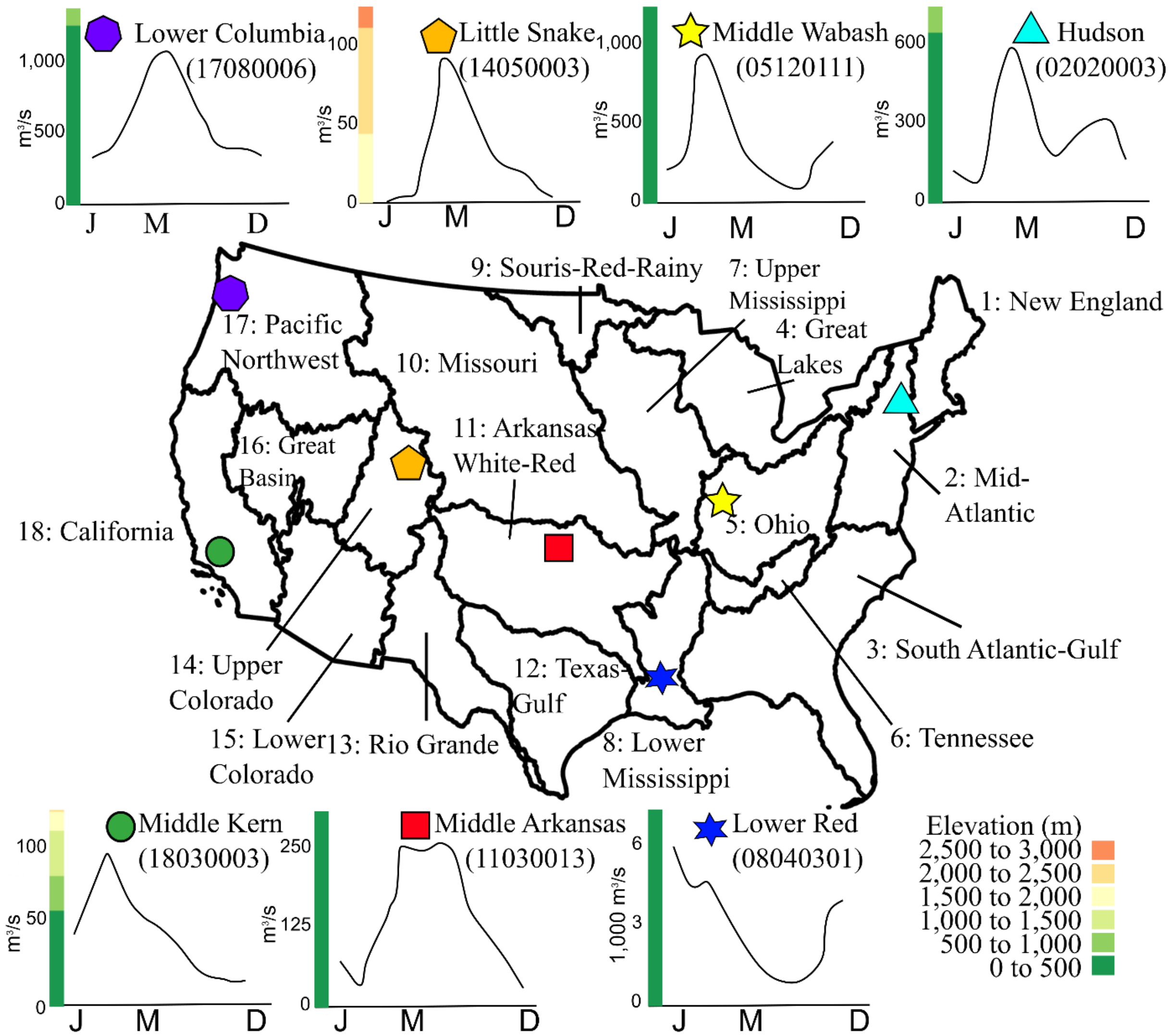


Figure 2.1. Location of the eighteen HUC2 watersheds and seven HUC8 watersheds (HUC8 number listed below each basin's name) selected for the case study with average annual hydrographs and elevation bars (on the hydrographs' y-axes) , representing the proportion of the basin in that elevation range.

Table 2.1. Attributes of each basin used in the case study.

HUC8	Basin name	Primary land cover[a]	Maximum Strahler stream order[b]	Mean annual discharge (m^3/s)	Mean peak annual discharge (m^3/s)	Area (km^2)	Relief ratio	Principal sources of flood-causing precipitation or runoff[c]	References
17080006	Lower Columbia	Evergreen forest	9	7,533	14,915	1,754	0.017	Rain from extratropical cyclone on snowmelt	Simenstad et al. (2011); Cannon (2015)
02020003	Hudson	Deciduous forest	6	404	2,767	4,936	0.013	Rain from extratropical cyclone on snowmelt	Jackson et al. (2005)
18030003	Middle Kern	Herbaceous	6	46	149	6,779	0.031	Extratropical cyclone or associated front	Katibah (1984)
05120111	Middle Wabash	Cultivated crops	7	343	1,761	5,253	0.001	Rain from extratropical cyclone on snowmelt	Scheel et al. (2019)
14050003	Little Snake	Shrub/Scrub	6	32	147	7,926	0.001	Rain from extratropical cyclone on snowmelt	Blaschak (2012); Caskey (2013); Caskey et al. (2015)
08040301	Lower Red	Woody wetlands	8	6,173	15,282	2,350	0.001	Extratropical cyclone or associated front	Knox and Latrubesse (2016)
11030013	Middle Arkansas	Cultivated crops	7	124	800	2,350	0.002	Rain from extratropical cyclone on snowmelt	Guilliams (1998); Matthew et al. (2005)

Note. [a]Source is the 2016 NLCD. [b]Stream order based on Strahler (1952). [c]Sources of flood-causing precipitation or runoff based on Hirschboeck (1991).

2.2.3 Hydrogeomorphic Floodplain Delineation

I established a CONUS floodplain calibrated to FEMA special flood hazard areas A and AE using the GFPLAIN algorithm (Nardi et al., 2019) and the 1 arc-second (~30 m resolution) NED (Table 2.2). The FEMA special flood hazard maps' coverage of the CONUS (~60% of the CONUS area) makes them suitable as validation for CONUS flood hazard models, but unsuitable as a stand-alone CONUS floodplain. Valued for its continental coverage, the accuracy of FEMA maps varies and can be less accurate than local high quality flood models (Blessing et al., 2017). Previous studies have used FEMA floodplain maps for either calibration or validation (e.g., Wing et al., 2017; Nardi et al., 2018; Annis et al., 2022), as the level of agreement between hydrogeomorphic models and other flood models indicates the suitability of this type of application, especially in data-poor areas (Lindersson et al., 2021). Lower model accuracy of hydrogeomorphic floodplains in certain areas (e.g., dry, steep, flat areas or those near the coast) (Annis et al., 2022; Lindersson et al., 2021) is mitigated by our calibration at the 2-digit HUC basin level. The GFPLAIN algorithm identifies geomorphic floodplains in two main steps: (1) terrain analysis of a DEM for basin drainage extraction and (2) floodplain delineation. It uses an adaption of a scaling regression from Leopold and Maddock (1953) to relate stage to upstream contributing area:

$$FH_i = aA^b$$

where FH_i is the maximum flow depth at a location for the recurrence interval i, a and b are dimensionless scaling parameters, and A is the contributing area for that location (Scheel et al., 2019). Calibration was conducted by comparing model performance with FEMA flood maps on stream segments from the first six stream orders in each 2-digit HUC basin using the F measure of fit, which is the ratio of floodplain area correctly modeled to the total area modeled and predicted by FEMA (Equation (4); Horritt and Bates, 2001). I kept parameter *a* constant at 0.0035 while varying *b* between 0.25 and 0.42, given the strong linear correlation of the two parameters (Annis et al., 2019). The value of *a* and the range of *b* was selected based on previous studies (Nardi et al., 2006; Nardi et al., 2018, Annis et al., 2019; Scheel et al., 2019). A third parameter, contributing area threshold, was kept constant at 50 km^2 based on previous studies (Annis et al., 2019; Scheel et al., 2019). Values for the *b* parameter 0.32, 0.34, and 0.36 resulted in the highest F measure of fit during floodplain calibration (Figure A.1). The complete results are located in Appendix A.

2.2.4 Data

I used national, publicly available data sources (Table 2.2) in our analyses. The geomorphic variables (slope, planform curvature, profile curvature, relative elevation, and aspect difference) were developed according to Wing et al. (2019) from the National Elevation Dataset (NED 10 with resolution of 1/3 arc-second and approximately 10 m) based on its higher vertical accuracy compared to other nationally available topographic data sets (Gesch et al., 2014). I included the 2016 National Land Cover Database (NLCD) due to its national coverage and the expected association of artificial levees with certain land covers. I developed the six variables according to the distance from stream order one through six using the National Hydrography Dataset Plus High Resolution (NHD Plus) to capture the expected proximity of artificial levees to streams.

Table 2.2. Variables and datasets used in the study

	Variable	Dataset	Type	Resolution
Geomorphic Variables	Slope	National Elevation Dataset (Gesch et al., 2002)	Raster	10 m
	Planform curvature			
	Profile curvature			
	Relative elevation			
	Aspect difference			
Land Cover Variable	Land cover	National Land Cover Database, 2016 (Jin et al., 2019)	Raster	30 m
Spatial Variables	Distance from stream order 1 stream	National Hydrography Dataset Plus High Resolution (Buto & Anderson, 2020)	Vector	-
	Distance from stream order 2 stream			
	Distance from stream order 3 stream			
	Distance from stream order 4 stream			
	Distance from stream order 5 stream			
	Distance from stream order 6 stream			
	Basin boundaries	USGS Hydrologic Unit Maps (Seaber et al., 1987)	Vector	-

2.2.5 Model Development and Implementation

2.2.5.1 Case Study Modeling

Similar data processing and modeling procedures were followed in the case and national studies (Figure 2.2). For the case studies, R software was used to process data sources (R Core Team, 2020). Twelve raster layers (all of the variables in Table 2.2 except "basin," which was added after sampling), each with a 10 m resolution, were created for the calibrated 100-year hydrogeomorphic floodplains in each of the seven basins (Table 2.1). I randomly sampled levee locations from the NLD and non-levee locations from other places within the GFPLAIN floodplain using over 480 general logistic models (GLMs), random forest models (RFs), and support vector machine models (SVMs) with various sample sizes, non-levee to levee sample size ratios (absence/presence), and different combinations of variables (Table 2.3). I defined and tested the full model as model 1 with all variables: the five geomorphic variables (slope, profile curvature, planform curvature, relative elevation, and aspect difference), the 2016 NLCD, the six distance-from-stream-order variables, and the HUC8 basin (Table 2.2). I also tested models with each of the variables removed from the full model and some other combinations.

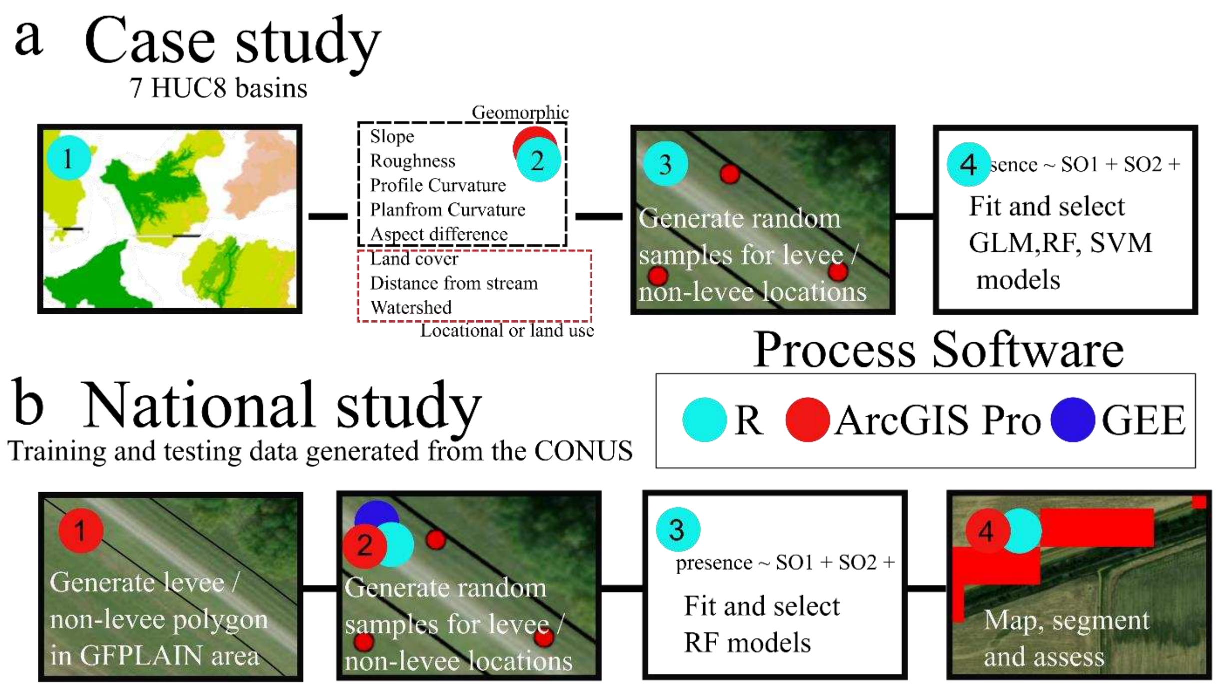


Figure 2.2. Workflow for the case and national studies. (a) In the case study I (1) generated 7 individual multi-layered rasters for each HUC8 basin hydrogeomorphic floodplain where each layer is a variable (2). (3) I combined random samples taken from levee (the NLD) and non-levee locations and (4) fit different GLM, RF, and SVM models as discussed above. I then assessed results. (b) In the national study, I (1) generated polygons around the NLD levees and all other areas in the hydrogeomorphic floodplain in each HUC2 basin to assist random sampling in ArcGIS Pro. (2) I generated random samples from the CONUS for two different sized data sets (n = 30,600 and 3,060,000) using GEE to generate the geomorphic variable values and ArcGIS Pro to generate the NLCD and distance from stream order 1-6 variables. (3) I fit different RF models to the data and applied the highest performing model to the contiguous U.S. hydrogeomorphic floodplain (4). I generated data from each HUC2 basin in ArcGIS Pro, imported and fit the data in R, and exported the model results back into ArcGIS Pro where I mapped, segmented, and then analyzed the results.

Table 2.3. Model names and variables in the study

Model	Total variables	Variables
1	13	Geomorphic variables (slope, profile curvature, planform curvature, relative elevation, aspect difference), NLCD, basin, distance from stream order 1-6
2	12	Model 1 without slope
3	12	Model 1 without profile curvature
4	12	Model 1 without planform curvature
5	12	Model 1 without relative elevation
6	12	Model 1 without NLCD
7	12	Model 1 without aspect difference
8	12	Model 1 without basin
9	7	Model 1 without distance from stream order variables
10	6	Distance from stream order variables only
11	11	Model 1 without distance from stream order variables and aspect difference
12	8	NLCD, basin, distance from stream order 1-6

I analyzed the impacts of sample size, ratio of levee to non levee (presence to absence) data, and model variables on GLM, RF, and SVM performance through separate comparisons of the performance of models generated from a large number of data sets in which only those (e.g. sample size) quantities varied. I used 113 independent data sets varying from 110 to 13,900 sampled locations per basin with model 1. Based on those results, in which the RF model outperformed SVM and GLM models at every sample size, I narrowed subsequent case study analyses to RF models only though I revisited all three model types at the CONUS level. The impacts of varying ratios of levee to non levee data to RF performance were analyzed using 93 independent data sets with absence to presence ratios ranging from 0.04 to 23.6. I attempted to hold sample size constant but experienced a range of 812 to 856 total sampled locations for each dataset due to varying "NA" value generation resulting from imperfect replacement of NA values sampled from masks. Absences were sampled everywhere in the GFPLAIN floodplain, excluding a 10 m buffer from the NLD centerlines. RF models with absence to presence ratios of 0.7 had the highest performance. Accordingly, I used this ratio for subsequent analyses although I tested ratios of 0.7 and 1.0 at the CONUS scale. I analyzed RF model performance with different variables using 50 independent data sets each consisting of data from 1,000 locations per basin with a 0.7 absence to presence ratio. Training data consisted of 70% of each data set with 30% used for validation. The results of these processes guided the CONUS study.

I conducted a leave-one-out cross-validation (Stone, 1974) with 1,171 NLD levees in the lower Mississippi Basin to understand how a parsimonious RF model (model 12) using the NLD and distance-to-stream-order variables behaved with undocumented levees that were not in the NLD. Given the close proximity of artificial levees (e.g., 74% of NLD levees in the LMR basin are within 5 km of each other) and the size of data sets (n > 3,000,000) used in the CONUS study, this cross validation best approximates a model's ability to detect undocumented levees. Using a 170,000-location sample with a 0.7 absence to presence ratio, I added a "levee ID" to each levee and wrote R code to iterate through each levee by first removing that levee from the data set, generating a model from 70% of the remaining data, applying the model to the withheld levee data, and recording the model accuracy. This process was conducted for each of the 1,171 artificial levees. I also recorded the shortest distance between each levee to understand how proximity impacted model performance.

2.2.5.2 National Modeling

In the national study, I tested different model types (GLM, RF, and SVM) and model variables using two different sized datasets generated from the CONUS floodplain for training and validation. I selected the most accurate model based on multiple accuracy metrics, a RF model trained on 2,142,000 sampled locations with land cover, HUC2 basin, and six distance from stream order 1-6 variables and applied it to the entire CONUS floodplain using ArcGIS Pro and R. More details about initial national modeling efforts are in Appendix A.

I used ArcGIS Pro to generate two polygons in each of the 18 HUC2 basins: a "levee" polygon within 10 m of the NLD, and a "non-levee" polygon for other areas in the hydrogeomorphic floodplain. Based on analyses in the case study, I randomly generated two differently sized data sets both with 0.7 absence to presence ratios: (1) 700 non-levee and 1000 levee locations in each HUC2 basin (total n ~ 30,600) and a much larger sample with (2) 70,000 non-levee and 100,000 levee locations in each HUC2 basin (total n = 3,060,000) for all 13 variables required for model 1. I used 70% of each data set for model training (n = 21,420 sampled locations for the small data set and 2,142,000 sampled locations for the large data set). I used GEE cloud computing and the Terrain Analysis in Google Earth Engine (TAGEE) script (Safanelli et al., 2020) to calculate the five geomorphic variables (Table 2.2). The land cover variable and the distance-from-stream-order variables were generated in ArcGIS Pro. These data were exported into R and fit to different models (Table 2.4).

I selected the best performing model based on a full suite of accuracy measurements (Table 2.4). I then generated a 30 m resolution raster from the calibrated CONUS 100-year hydrogeomorphic floodplain containing ~880,000,000 pixels. Working by HUC2 basin, I generated the dataset for each location in ArcGIS Pro, exported the data into R, applied the model, and exported the predicted values (in the case of RF models "1" or "0") back into ArcGIS Pro, where the predicted values were mapped and segmented. Segments were analyzed to determine whether they were already represented in the NLD, which stream order they were meant to "protect" against, and to estimate their length using segment attributes.

2.3 Results

2.3.1 Case Studies

RF models demonstrated the best predictive performance for identifying artificial levees at every sample size, followed by SVM models (Figure 2.3a). The GLMs demonstrated the worst performance of all three models. Because the RF models demonstrated the best performance, all our subsequent results in the case study focus on RF model outputs. An absence-to-presence ratio of 0.7 resulted in the best RF model performance, but data with ratios between 0.45 and 1.24 performed well (Figure 2.3b). Furthermore, RF performance continued to improve with larger sample sizes from each basin, exceeding a kappa of 0.8 for sample sizes larger than 1,000.

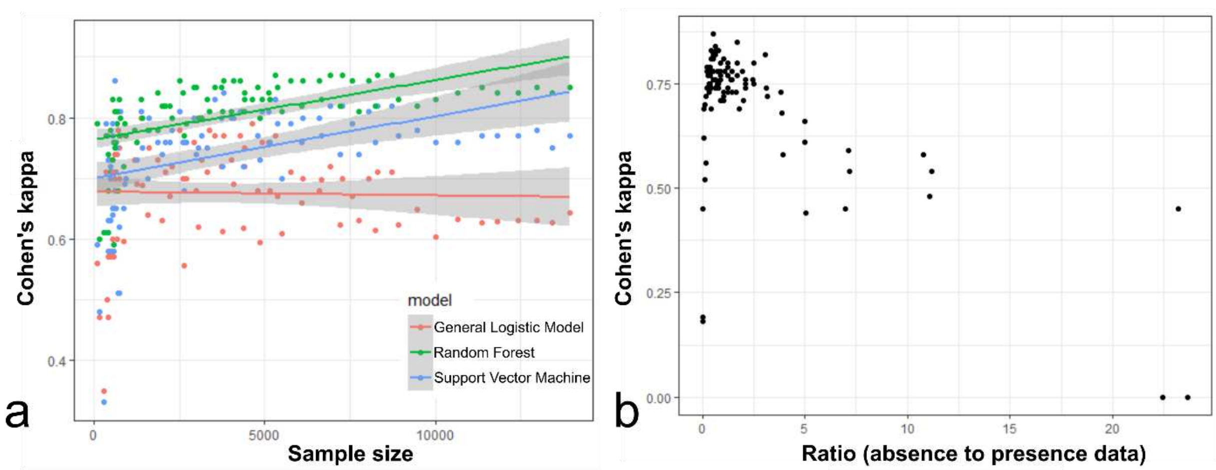


Figure 2.3. Model performance using Cohen's kappa in the case study. (a) GLM, RF, and SVM model performance with sample size varying from 110 to 13,900 sampled locations total in the seven basins for 113 independent samples. The grey envelopes are 95% confidence intervals for logistic models, depicted by a solid line, fit to the data (b) Performance of 93 independent RF models by varying absence/presence ratio of sampled locations while controlling for sample size (n ~ 832 sampled locations).

Different RF models, each with 100 trees of three variables sampled at each node, were applied to 50 different random samples from 1,000 sampled locations and a 0.7 absence/presence ratio (Figure 2.4). Model 1, with all variables, only slightly outperformed models with one less variable with kappas in the 0.75-0.8 range. A model without any geomorphic variables (model 12) performed almost as well as the full model.

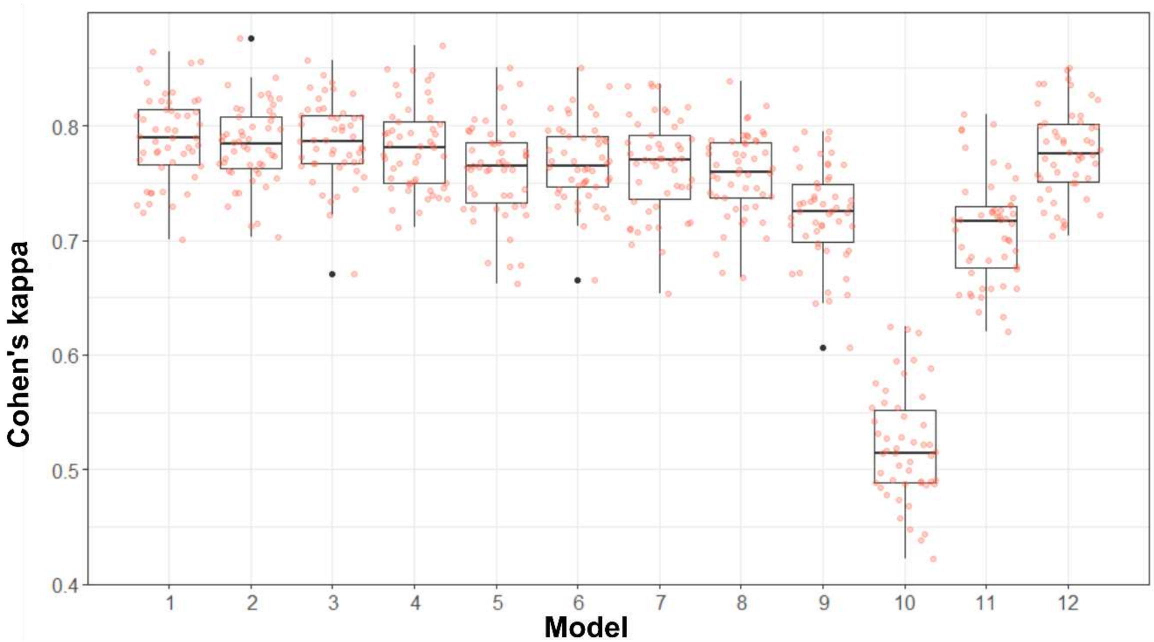


Figure 2.4. RF model performance by Cohen's kappa and variables for 50 data sets, each with a 0.7 ratio of absence to presence data and ~ 1,000 sampled locations. Boxplots are plotted along with individual model values. The model number on the x-axis corresponds to models listed in table 2.2.

2.3.2 Leave-one-out cross-validation

An RF model using model 12 detected 61% of levees when they were left out of the training dataset. Detected levees were longer than undetected levees such that sum of the length of detected levees (7,473 km) represented 94% of total levee length (7,910 km) (Figure 2.5a). Levees were close together, with 74% of levees within 5 km of each other and 94% of levees within 25 km (Figure 2.5b).

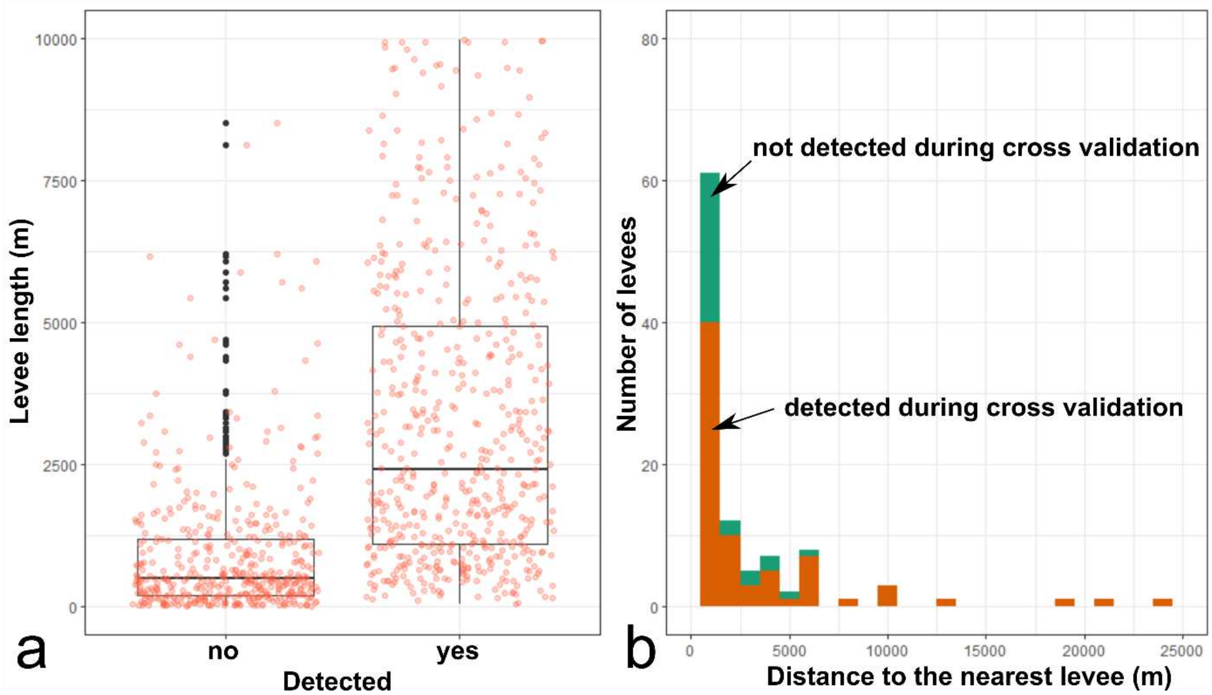


Figure 2.5. Results from the leave-one-out cross-validation. (a) Longer levees were detected more often than shorter levees so that detected levees represent 94% of total levee length. (b) Levees are close together, with 74% of levees within 5 km of each other and 94% within 25 km.

2.3.3 National study

I tested different variables, model types, sample sizes and absence/presence ratios using models trained at a national scale (Table 2.4). As in the case study, RF models outperformed SVMs and GLMs with RF model performance increasing with sample size. I was initially surprised to discover that model 12, without any of the five geomorphic variables, outperformed the model 1 by 0.1 kappa and other performance measures (Table 2.4, Figure A.2). Further, I did not initially test model 12 in the case study, and only added it after discovering it here.

The performance of model 12 corresponds to a 97% accuracy, meaning that 97% of the known non-levee and levee locations in the ~918,000 CONUS validation sample were predicted correctly, although the results of this study indicate that many false-positives may represent undocumented artificial levees (what I call potential levees). However, the model performance varied spatially, ranging from 0.813 kappa in the California HUC2 basin to 0.999 kappa in the Upper Colorado HUC2 basin, correlating to the inverse of the NLD levee total length in each HUC2 basin.

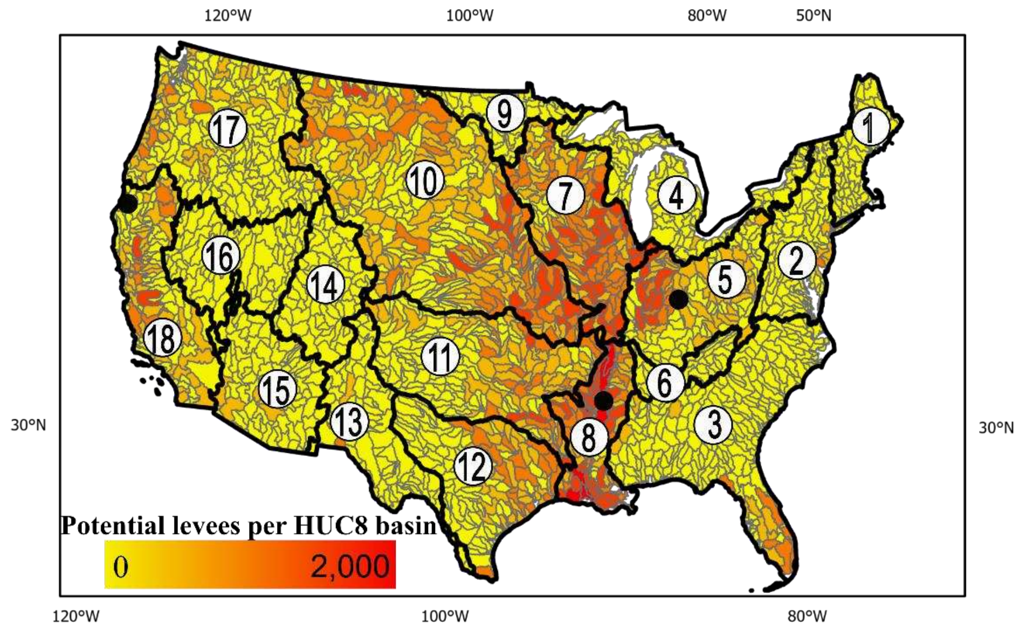
Table 2.4. Model performance in the national study

Model[a]	ML[b]	Size[c]	Ratio[d]	k[e]	TP Rate[f]	TN Rate[g]	FP Rate[h]	FN Rate[i]
1	RF	1700	0.7	0.69	0.87	0.82	0.18	0.13
1	RF	1700	1	0.69	0.82	0.88	0.12	0.18
1	GLM	1700	1	0.47	0.68	0.80	0.20	0.32
1	SVM	1700	1	0.55	0.83	0.74	0.26	0.17
12	RF	1700	0.7	0.65	0.87	0.77	0.23	0.13
1	RF	170,000	0.7	0.85	0.95	0.90	0.10	0.05
12	RF	170,000	0.7	0.94	0.99	0.95	0.05	0.01
12 + relative elevation	RF	170,000	0.7	0.89	0.97	0.92	0.08	0.03
12 + profile curvature	RF	170,000	0.7	0.90	0.98	0.93	0.07	0.02
12 + aspect difference	RF	170,000	0.7	0.87	0.97	0.91	0.09	0.03
12 + slope	RF	170,000	0.7	0.89	0.97	0.92	0.08	0.03

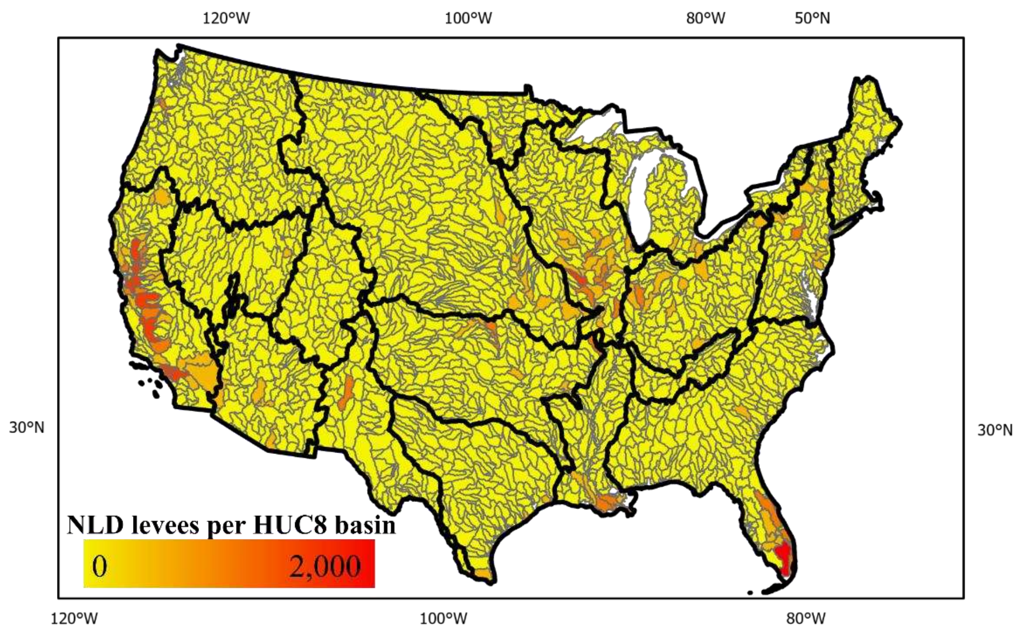
Note. [a]"Model" corresponds to the variables listed in table 2.3. [b]The "ML" column denotes the machine learning or statistical model used. [c]"Size" denotes the total sample size taken from each HUC2 basin for both model training and testing. [d]"Ratio" denotes the ratio of absence to presence in the sample. [e]The result denotes the Cohen's kappa of the model on the testing sample, where I used a 70/30 random split for training and validation in all models. [f]True positive rate. [g]True negative rate. [h]False positive rate. [i]False negative rate.

Potential artificial levees, those areas identified that may be artificial levees and are not identified by the NLD, are listed in Table 2.5 (which also lists NLD levees separately). Potential levees were concentrated in the upper and lower Mississippi and the Missouri basins (basins 7,8,10 in Table 2.5 and Figure 2.6). Potential levees were also concentrated along streams of order 2 to 6, constituting 75% of total levee length (Figure 2.7). There were 146,404 potential levees identified constituting a total length of 182,213 km (Table 2.5). Normalized artificial levee length (the total length of NLD or potential levees associated with stream order X divided by the combined length of streams with stream order X) gives a sense of how streams of a particular order are impacted by artificial levees. Artificial levees provide greater "protection" along streams of greater order, with normalized length approaching 0.20 for stream order 10 (Figure 2.7). Potential levees and those documented in the NLD represent coverage of 2% of the total length of streams in the contiguous United States.

a


120°W
100°W
80°W
50°N
30°N
Potential levees per HUC8 basin
0
2,000
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

b


120°W
100°W
80°W
50°N
30°N
NLD levees per HUC8 basin
0
2,000

c

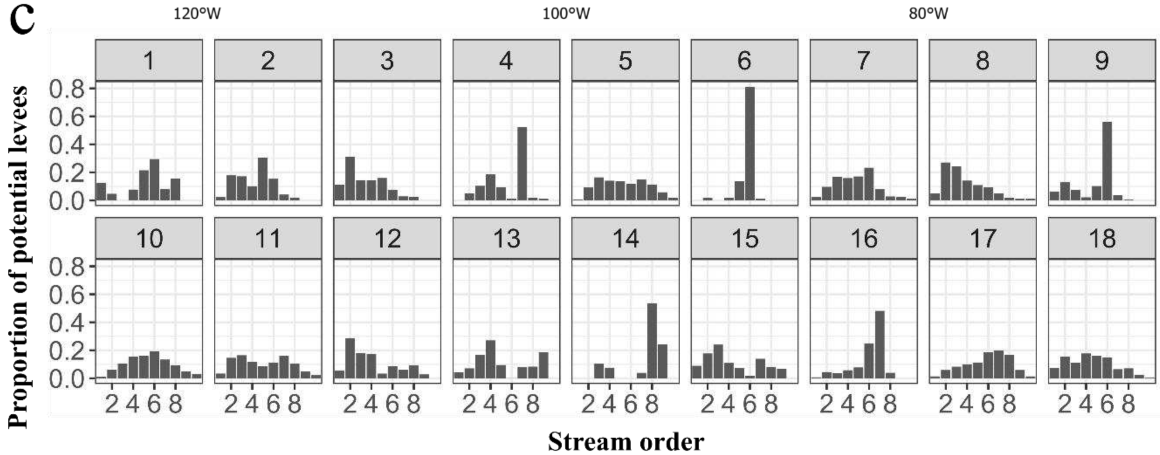

Proportion of potential levees
Stream order
0.0
0.2
0.4
0.6
0.8
2 4 6 8
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

Figure 2.6. A spatial and stream order representation of potential and NLD artificial levees by HUC8 and HUC2 basin. (a) The number of potential levees per HUC8 basin. HUC2 basin boundaries, in bold, are denoted by number. Three black dots indicate potential levees examined in Figure 2.8. (b) The number of NLD levees per HUC8 basin with HUC2 basin boundaries in bold. (c) The proportion of potential artificial levee length along each stream order in 18 HUC2 basins.

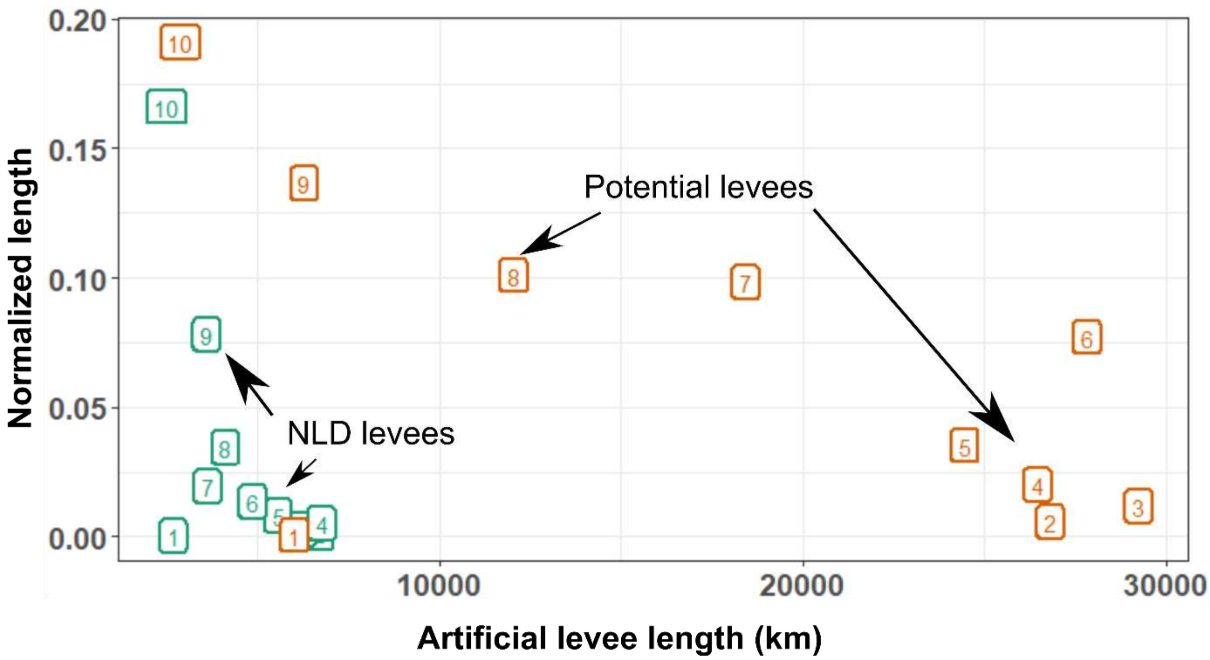


Figure 2.7. Potential and NLD artificial levee stream orders by normalized length and the sum of levee length for that order. NLD and potential levees are annotated by green and orange colors, respectively.

To illustrate a few locations where I identified levees not present in the NLD and to discuss how the model works, I highlight three potential levees and the raw model results (prior to segmentation) that I was able to ascertain are definitely levees (Figure 2.8). Brookville, Indiana sits at the confluence of the East Fork Whitewater River (stream order 7) and the Whitewater River (stream order 8) which is a tributary to the Ohio River (Figure 2.8a). This artificial levee was built along the East Fork Whitewater River and is more than 850 m long, 1-2 m tall, and has been the subject of a riverwalk project (Norwood, 2020). This illustrates how undocumented artificial levees influence hydrogeomorphic floodplains, as the 100-year floodplain and detection are immediately next to the levee. Elaine, Arkansas is situated several kilometers from the Mississippi River next to an old meander scar and is “protected” from Mississippi River floods by kilometers of massive artificial levees. One 400 m section of artificial levee is not documented in the NLD (Figure 2.8b). Arcata, California is a tidally influenced community situated next to Arcata Bay and the Pacific Ocean in northern California. The area was subject to salt marsh reclamation for pasturage during the last 150 years (Murray & Wunner, 1980). This 200 m length of artificial levee, visible from U.S. Route 101, was likely built to reclaim floodplain along Jacoby Creek for pasturage (Figure 2.8c). Detected in the raw model results, this large segment was not included in the final results because the segment overlapped a NLD levee to the north. These results indicate the value of buffering the hydrogeomorphic floodplain by 100m and considering how to separate model results from the NLD.

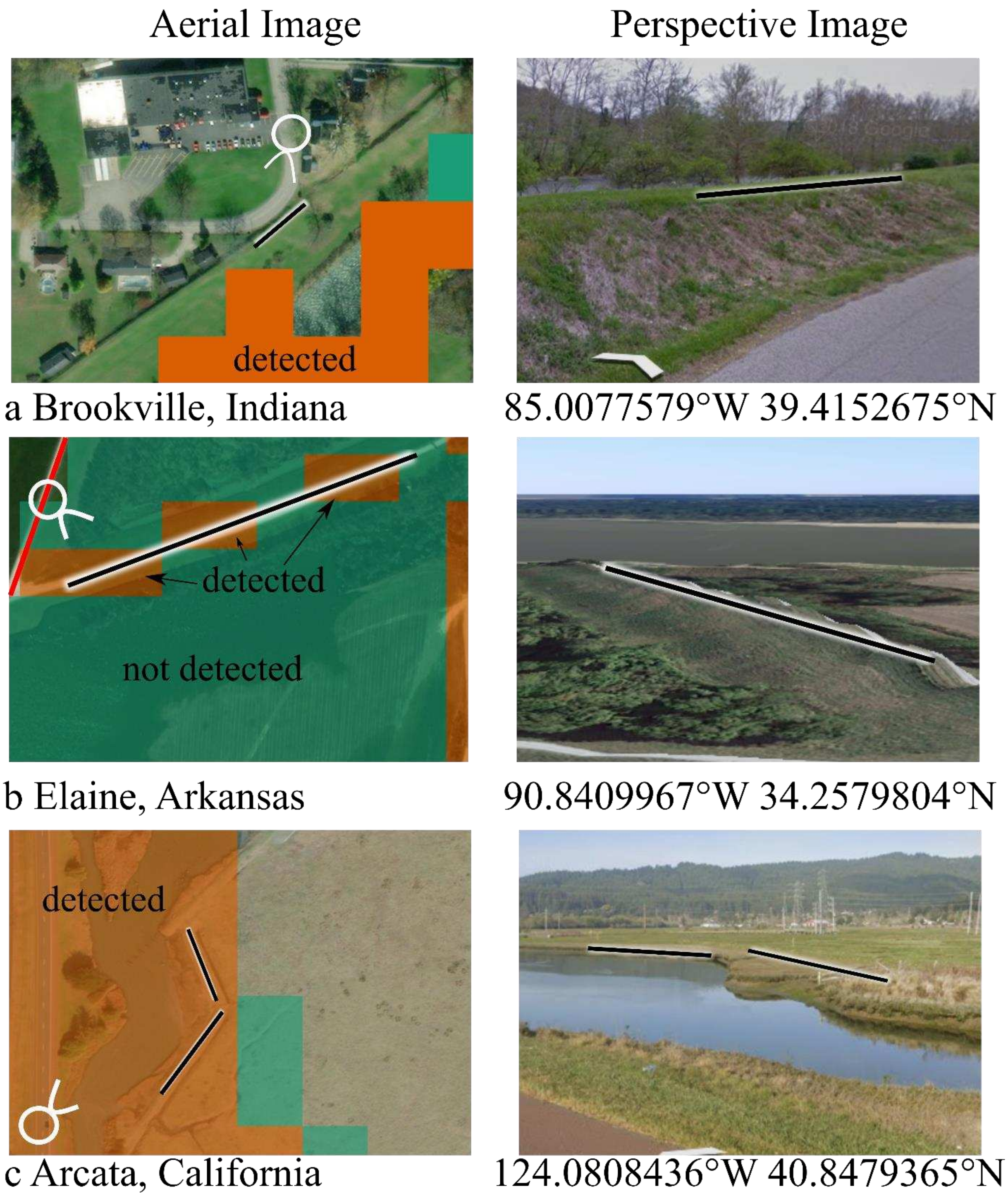


Figure 2.8. Aerial and perspective images of potential artificial levees discovered during this process. The white circle and arrow on the aerial image indicates the perspective location for the perspective image. Black lines correlate locations between the two images. Raw model results, prior to segmentation, are plotted in orange (detected) and green (not detected). (a) Brookville, Indiana levee is visible as a long linear feature in aerial imagery, on Google Street view, and written about in an online news article (Norwood, 2020). (b) Elaine, Arkansas levee

on the Mississippi River is connected to a NLD levee (indicated by a red line) but remains undocumented. (c) Arcata, California levee along the Gannon Slough and U.S. Route 101 is likely related to salt marsh reclamation for pasturage.

Table 2.5. NLD and potential levee lengths (km) by HUC2 basin. Potential levees were identified by this study and are not part of the NLD.

HUC2 basin	NLD (km)	Potential levees (km)
New England (1)	89	25
Mid-Atlantic (2)	617	2,220
South Atlantic-Gulf (3)	2,410	5,921
Great Lakes (4)	41	277
Ohio (5)	1,148	12,216
Tennessee (6)	45	40
Upper Mississippi (7)	4,804	35,374
Lower Mississippi (8)	7,912	38,657
Souris-Red-Rainy (9)	466	459
Missouri (10)	4,438	39,221
Arkansas-White-Red (11)	2,939	16,073
Texas-Gulf (12)	2,403	9,230
Rio Grande (13)	1,074	965
Upper Colorado (14)	154	9
Lower Colorado (15)	1,582	1,471
Great Basin (16)	133	392
Pacific Northwest (17)	2,082	10,747
California (18)	14,306	8,916
Total Length (km)	46,643	182,213

2.4 Discussion

2.4.1 Location and prevalence of artificial levees

This analysis indicates that the NLD may only show 20% of the artificial levees in the CONUS. Over 62% of potential levees are concentrated in the upper and lower Mississippi basins and the Missouri basins with potential levee length exceeding documented levee length by factors of seven, five, and nine, respectively (Figure 2.6, Table 2.5). Potential levee length in the Ohio basin exceeds NLD levees by a factor of 11. These details, combined with the concentration of potential levees along smaller stream orders (Figure 2.7), seems to reflect the long history of artificial levee construction in the lower Mississippi River basin and early interventions along the Mississippi and Ohio Rivers, such as the 1824 Rivers and Harbors Bill (Wohl, 2005). I suspect that NLD and potential levees represent a continuum in which NLD levees represent one pole with larger levees more recently built with state or federal funds, and designed to "protect" against streams of a larger order. The other pole is represented by potential levees which are smaller, built longer ago and by landowners, and situated next to smaller streams. I see this dynamic illustrated in the California basin where potential levee length is half that of NLD levee length (Table 2.5), reflecting the later period (i.e., early 20th century) of artificial levee construction in that basin (ASCE, 2017) and better documentation that comes with more recent construction.

2.4.2 Spatial and geographic implications of findings

Characterizing the shape of Earth's surface is considered the primary method for quantitative land-surface analysis (Sofia, 2020). Our understanding of landscapes arises from human cognition of spatial patterns in the field (Roering et al., 2013), so it is reasonable that geomorphic variables are nearly the exclusive set of variables used in recent efforts to detect artificial levees and other earthworks. Nonetheless, topographic patterns are only one way to parameterize geomorphic features. A quick review of efforts to model other physical phenomena with less pronounced topographic profiles, such as wetlands, indicates a willingness of researchers to explore other ancillary variables such as "distance to stream" (e.g. Golden et al., 2016; Berhane et al., 2018; Ansari et al., 2019). Unlike other geomorphic features of mostly natural origin, artificial levees are constructed solely by humans to protect infrastructure from river floods. They are anthropogenic features that are intimately tied to human land cover and hydrologic features. Consequently, the spatial patterns humans have created in building levees are real and useful for modeling.

I am interested in the causes of the difference in accuracy between models that use land cover and stream order variables (e.g., model 12) and those models using geomorphic variables (e.g., model 1). Detecting artificial levees presents significant technical challenges due to their small size, geographic ubiquity, and varied morphology (Steinfeld and Kingsford, 2013). Artificial levees can be massive structures or features nearly invisible to both the eye and topographically based analyses (Figure A.3), with the height of some artificial levees less than the vertical error of topographic datasets (e.g., the mean relative vertical accuracy of the NED is 0.81 m with the accuracy of 95% of locations within 2.93 m (Gesch et al., 2014)). Furthermore, the resampling process of digital elevation models tends to smooth topographic crests (such as those of levees) making the features even more topographically indistinguishable or even invisible (Wing et al., 2019). Consequently, it is not surprising that spatial and land cover patterns seem to be more useful than geomorphic patterns in a national study given the diverse geomorphic signatures of both documented artificial levees (such as those in figure A.3, which can be used as training data) and undocumented levees. Other researchers working at the national scale or larger (e.g., Grill et al., 2019) recognized the strong correlation of lateral discontinuity structures with human development and employed nightlight intensity as a proxy for lateral discontinuity in the absence of global records of artificial levees and other structures.

Recent investigations have raised concerns over validation strategies for large scale modeling studies where the employment of spatially autocorrelated training and validation data leads to inflated estimates of model accuracy (e.g., Ploton et al., 2020). The issue is covered extensively, especially in the ecological literature (see Roberts et al., 2017), but I consider it appropriate to discuss the suitability of the validation techniques employed here. Karasiak et al. (2021) explains how the winner of a land cover classification contest was able to employ geographic pixel location only due to the ostensible spatial autocorrelation of the training and validation data sets. Unlike that example, and as previously discussed, I consider the spatial patterns expressed by the distance from stream order and land cover variables to be real patterns created by humans because land cover and stream flow were primary factors in the decision process that led to artificial levee construction. In addition, our method of mapping model 12 over the GFPLAIN floodplain is considered interpolation, not extrapolation, because I am applying the model in the same domain (i.e., the same geographic extent and variable domain) as that from which the training data are generated. Validation error of random samples is considered accurate in models with applications in similar geographic and variable domains (Roberts et al., 2017). Unlike the problematic models discussed by Ploton et al. (2020), our training and validation samples (n ~ 3,060,000) are drawn from the same geographic and variable space as the model application area (the full 100-year GFPLAIN floodplain). I am not applying the model in a different geographic area. The detection of unknown levees representing 94% of total levee length in the leave-one-out cross-validation substantiates these claims. Initially, our use of the 2-digit HUC code, instead of the 8-digit code, as a variable was driven by the RF package's (Liaw and Wiener, 2002) limit of 32 levels for a factor variable. Our experience indicates that the use of the 18 levels of the 2-digit HUC code strikes the right balance between harnessing RF's ability to model large samples without overfitting (Breiman, 2001) and the need to apply spatial patterns locally. After even a casual perusal of the USACE NLD website (https://levees.sec.usace.army.mil), it is reasonable to suspect that different areas of the United States might exhibit different spatial patterns of artificial levees and thus employing the natural basin configuration at the regional scale is appropriate. An example of this is the artificial levee density difference between basins such as the lower Mississippi Basin and California Basin (with a relatively high density of artificial levees) and the New England basin (with a relatively low density). Another consideration for modeling is that the 2-digit HUC basin level is the most specific HUC code designation so that all basins contain known levees. CONUS models using HUC4 basins or larger HUC code designators would have basins without training data.

2.4.3 Future directions

Several areas appear promising for future research. First, artificial levee detection from a truly object-based approach could allow for the introduction of several object-based variables such as levee length and volume, mimicking the methods that experienced engineers or geomorphologists might employ to judge whether a structure is an artificial levee. Second, including additional variables, such as stream order 7-10 variables and spectral properties, could improve model accuracy. To test additional variables and the idea of using a larger sample size, I added a larger training sample size and a distance from stream order 10 variable in the lower Mississippi Valley (Table 2.6, Figure A.4). Both ideas improved model accuracy with increased sample size resulting in the greatest gains.

Table 2.6. Model performance in the Lower Mississippi HUC2 basin with a larger training sample size and an additional variable

Model	n	Kappa	TP Rate[a]	TN Rate[b]	FP Rate[c]	FN Rate[d]
12	170,000	0.846	0.94	0.91	0.09	0.06
12	1,700,000	0.958	0.97	0.99	0.01	0.03
12 + SO10	1,700,000	0.974	0.98	0.99	0.01	0.02

Note. [a]True positive rate. [b]True negative rate. [c]False positive rate. [d]False negative rate.

Third, I suggest coupling a geomorphic post-processing technique using high resolution topographic data with this model to determine whether the areas identified as potential-levees are shaped like levees. Although I found that including geomorphic variables reduced predictive power at the national scale, assessing how "levee like" potential levees are geomorphically in a subsequent step could help prioritize ground-truthing efforts. Fourth, the incorporation of field validation of potential levees similar to the validation of barriers conducted by Jones et al. (2019) would improve model certainty. Actual ground-truthing of potential levees identified here could complement the methods currently in use to document artificial levees for the NLD resulting in a more complete and certain database. Fifth, revisiting this analysis in a decade could benefit from expanded computing ability and more open source options for object based classification, a more thorough NLD as training data, and a more accurate NHD Plus HR. Sixth, the causes of the spatial patterns of potential and NLD levees, illustrated in figure 2.6, deserve more detailed exploration.

2.5 Conclusions

Our exploration of different variables and models to detect artificial levees led to a random forest model with land cover and stream order variables. Applying this model in a 100-year geomorphic floodplain in the contiguous U.S. indicated the potential for 182,000 km of artificial levees that are not included in the national levee database, suggesting that the database contains only 20% of artificial levee length in the continental U.S.. These levees missing from the national database were concentrated in the lower and upper Mississippi and Missouri basins and mostly along streams of order 2 through 6. When normalized for total stream length, larger stream orders were more impacted than smaller streams, with more than a third of stream order 10 streams impacted by NLD or potential levees. Ideas for future directions include methods that could further improve model performance and validation.

Data Availability

Potential levee location, estimated length, and HUC2 basin are available as supporting data: https://doi.org/10.4211/hs.729c0aea00bb48d6b6814c147e4318c4

Chapter 3 A river ran through it: floodplains as America's newest relict landform

Summary

Artificial levees are a major human modification of river corridors, but we still do not have a clear understanding of how artificial levees impact floodplain extent at regional and larger scales. I estimated changes in river-floodplain connectivity due to artificial levees in the contiguous United States (CONUS) using a combination of artificial levee databases, delineations of floodplain areas, and the deletion of artificial levees from topography. Our results indicate that artificial levees do not just decrease floodplain extent, but also alter locations of floodplain connectivity. Anthropogenically connected and disconnected locations are similar in land cover and are predominantly, in decreasing order of extent, cultivated, wetland, forested, and developed land cover types, with over 30% of the entire floodplain area in the CONUS cultivated or developed. This study indicates that artificial levees cause complex changes in river-floodplain connectivity and can increase flooded areas in some rivers (Knox et al., 2022b).

3.1 Introduction

River corridors include the active channel(s), floodplain, and underlying hyporheic zone. I define the floodplain as a frequently flooded, low-relief landform created by erosional and depositional processes under the contemporary hydrologic regime (Dunne and Alto, 2013). River scientists and engineers emphasize the importance of three-dimensional connectivity within river corridors (Harvey and Gooseff, 2015). Examination of lateral connectivity between the channel and floodplain can focus on water in association with flooding hazards (Dingman and Platt, 1977), flood peak attenuation (Woltemade and Potter, 1994), floodplain inundation (e.g., perirheic zones; Mertes, 1997), ecological considerations (e.g., flood-pulse concept; Junk et al., 1989), sediment fluxes (Wolman and Leopold, 1957; Woman and Miller, 1960; Nanson and Croke, 1992; Dunne et al., 1998), or other processes, but the commonality is that alteration of natural levels of lateral connectivity influences diverse river corridor functions.

American floodplain development kept pace with flood protection efforts during the 20th century, resulting in the constant rise of average flood-related economic losses (White, 2000). Worldwide, the restoration, rehabilitation, and conservation of large floodplain rivers are increasingly in conflict with development (Sparks, 1995; Wohl, Bledsoe, et al., 2015). Managing these conflicts requires an understanding of floodplain location and extent, as well as the water and sediment interactions between floodplain and channel (Wohl, Bledsoe, et al., 2015; Nardi et al., 2018). A rapid increase in the availability of Earth observation datasets and computational power has created new opportunities for the evaluation of floodplain mapping models (Annis et al., 2019), including hydrodynamic models at the continental scale (Wing et al., 2017) and hydrogeomorphic models at basin, continental, and global scales (Nardi et al., 2018; Annis et al., 2019; Nardi et al., 2019; Scheel et al., 2019; Annis et al., 2022). Hydrodynamic models are the state-of-the-art method for flood-hazard analysis and include backwater effects, flood wave attenuation, and urban interactions (Annis et al., 2019). Hydrogeomorphic models make efficient use of topographic data and are based on the natural depiction of floodplain topography resulting from recurring floods (Nardi et al., 2006). The level of agreement between hydrogeomorphic models and other flood hazard models indicates the suitability of hydrogeomorphic modelling, especially in data-poor areas (Lindersson et al., 2021). However, one of the sources driving model disagreement and inaccuracy is infrastructure, including artificial levees (Nardi et al., 2018; Annis et al., 2022).

Diverse human activities alter flow regime, floodplain morphology, and channel-floodplain connectivity (Hudson et al., 2008; Sofia et al., 2014). Artificial levees, for example, are built to inhibit lateral connectivity and are associated with significant ecological harm (Blanton and Marcus, 2009; Wohl, 2018). Surprisingly, there are few studies that evaluate the impact of artificial levees on floodplain extent at large watershed scales (Scheel et al., 2019). One example of such an evaluation employed the hydrogeomorphic GFPLAIN flood model (Nardi et al., 2019) on two versions of a DEM (digital elevation model), one with artificial levees removed, in the 100,000 km^2 four-digit hydrologic unit code (HUC) (Table B.1) Wabash basin (Scheel et al., 2019). At the continental scale, however, it remains unknown to what extent floodplains have been disconnected from channels in the USA or elsewhere in the world.

This is in striking contrast to knowledge of longitudinal disconnectivity created by dams (e.g., Graf, 1999, 2001; Nilsson et al., 2005). Dams are more readily detected in remote imagery and there are more likely to be systematic records of dam construction and the dimensions of individual dams (e.g., the US Army Corps of Engineers' National Inventory of Dams, or Global Dam Watch's global dam database). Increasing recognition of the intensity and spatial extent of river longitudinal disconnection by dams has been accompanied by a growing scientific literature on the environmental hazards created by this disconnectivity (e.g., Poff et al., 1997; Vörösmarty et al., 2003; Syvitski et al., 2005; Hecht et al., 2019). Extensive networks of artificial levees may be creating a similar amount of riverine degradation, but remotely delineating natural floodplains remains difficult, especially on smaller rivers (e.g., Stout and Belmont, 2014; Jafarzadegan and Merwade, 2017), and efforts to quantify the lateral disconnection of floodplains by artificial levees at regional to continental scales have been limited by lack of systematic records and inability to detect levees in remote imagery.

The United States Army Corps of Engineers (USACE) and the Federal Emergency Management Agency (FEMA) maintain a national levee database (NLD) for the United States, but it has not been evaluated for completeness until recently. In chapter 2, I estimated the completeness of the NLD to be approximately 20%, with over 182,000 km of undocumented potential levees identified in the contiguous U.S. (CONUS).

Here, I explore the spatial extent of lateral disconnectivity caused by artificial levees, called "anthropogenically disconnected" floodplains, as well as areas that levees cause to flood, called "anthropogenically connected" floodplains, in the CONUS. I apply a GFPLAIN flood model calibrated with FEMA flood-hazard maps (Table B.1) to two digital elevation models: one unmodified and one with artificial levees removed. My primary objectives are to determine the spatial distribution and stream order patterns of floodplain disconnection by artificial levees in the CONUS.

3.2 Results

3.2.1 Area analysis of anthropogenically connected and disconnected floodplain areas

The net effect of artificial levees varies by HUC8 basin with anthropogenically connected (areas flooded by artificial levees) exceeding anthropogenically disconnected floodplains (floodplains separated from rivers by artificial levees) CONUS-wide (Figure 3.1a, Table B.2). At the larger HUC2 basin scale, the Lower Mississippi River (HUC2 no. 8, 6,714 km^2), California (HUC2 no. 18, 2,043 km^2), and Missouri Basins (HUC2 no. 10, 2,016 km^2) had the greatest total artificially flooded and disconnected floodplains (Figure 3.1b). These basins have the greatest (46,569 km), fourth greatest (23,222 km), and second greatest (43,659 km) lengths, respectively, of known and potential artificial levees (chapter 2). Notably, the anthropogenically connected area was ~70% larger than the anthropogenically disconnected floodplain area in the

Lower Mississippi River basin (HUC2 no. 8) (Table B.2), reflecting the more confined topography of disconnected floodplain areas when compared to anthropogenically connected areas in that basin (Table B.3).

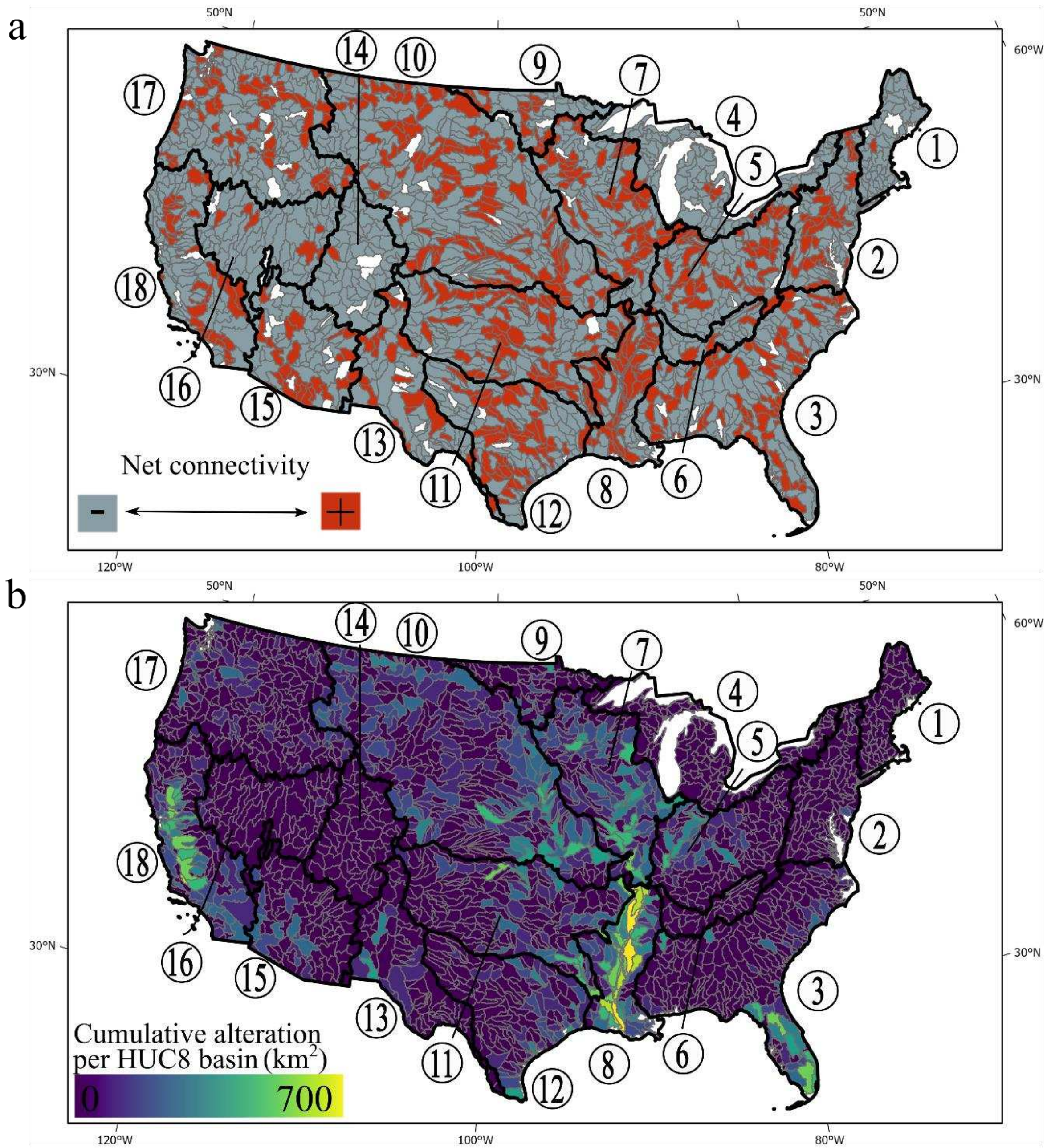


Figure 3.1. Net connectivity and cumulative alteration in the CONUS HUC8 basins. (A) Net connectivity compares whether each HUC8 basin has more anthropogenically connected or more anthropogenically disconnected floodplain area. Basins with no change in connectivity are

indicated by white. (B) Cumulative alteration by anthropogenically connected and disconnected floodplain areas. The 18 HUC2 basins are annotated in each figure with black lines and by numbers.

3.2.2 Land cover analysis

Land cover patterns of anthropogenically connected and disconnected floodplains are similar but with some notable differences (Figure 3.2, Table B.4). By far, cultivated land cover (cultivated crops and hay/pasture) makes up the largest proportion (55% for artificially flooded and 47% for disconnected floodplain) of each type of area. Wetlands (15% artificially flooded and 11% disconnected floodplain), forested (11% and 16%), and developed (11% and 12%) categories constitute progressively smaller proportions of land cover.

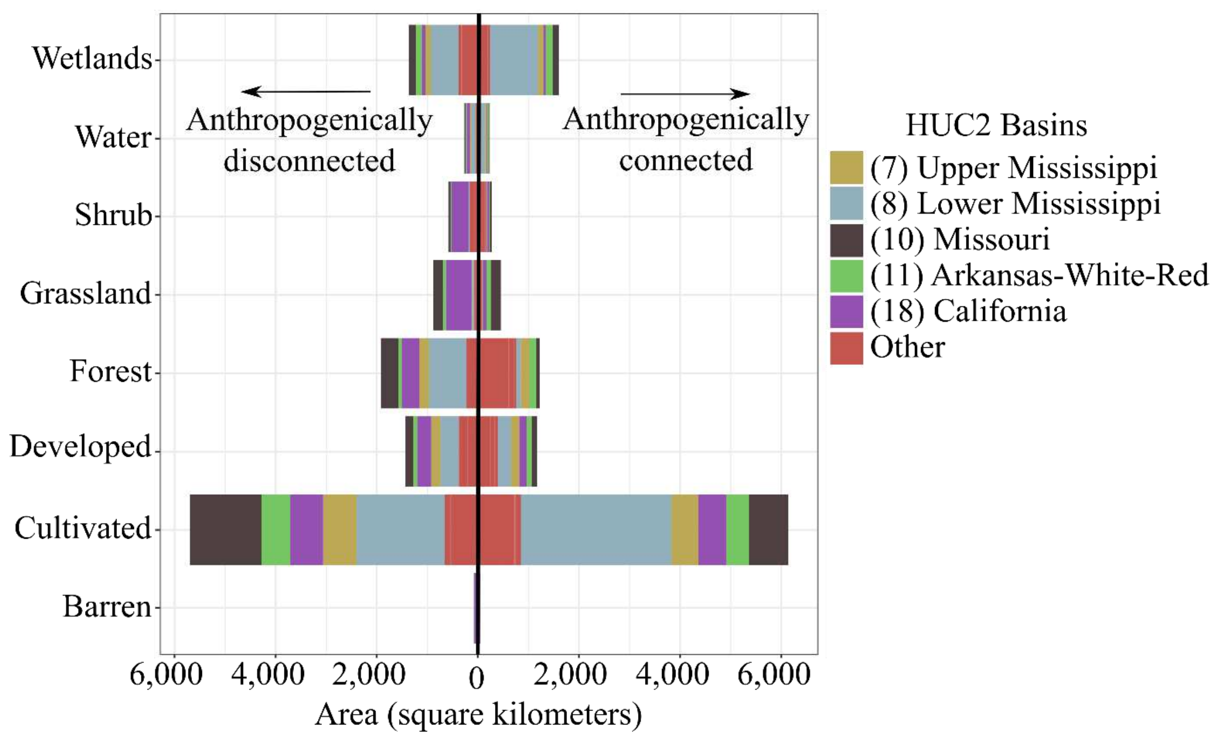


Figure 3.2. Land cover of anthropogenically disconnected floodplain and anthropogenically connected areas. CONUS land cover area (square kilometers) of each type of

area with HUC2 basin contributions annotated by color. HUC2 basins contributing less than 1,000 km^2 cumulative alteration were combined as “Other” for clarity.

There are several notable differences in the artificially flooded and disconnected floodplains (referred to as “disagreement areas”) when compared to the agreement areas (Table B.4). Cultivated land cover constitute twice the size of disagreement areas (55-47%) when compared to agreement areas (24%). Forested and developed areas experience similar trends. Agreement areas include more wetlands, open water, and shrub cover.

3.2.3 Stream order analysis

Stream order is a metric used to classify streams: a first order stream has no tributaries, and stream order increases downstream from the confluence of two streams of equal order (Strahler, 1957). Artificial levees are more likely to disconnect floodplains in first- to third-order streams, whereas the levees are more likely to enhance floodplain inundation in streams of fourth and higher orders (Figure 3.3). Stream order contribution patterns vary widely by HUC2 basin (Figure B.1). When compared to stream order contributions to agreement areas, disagreement areas peak in order two streams and then decrease with increasing stream order, indicating the effects of artificial levees on smaller order streams (Figure 3.3).

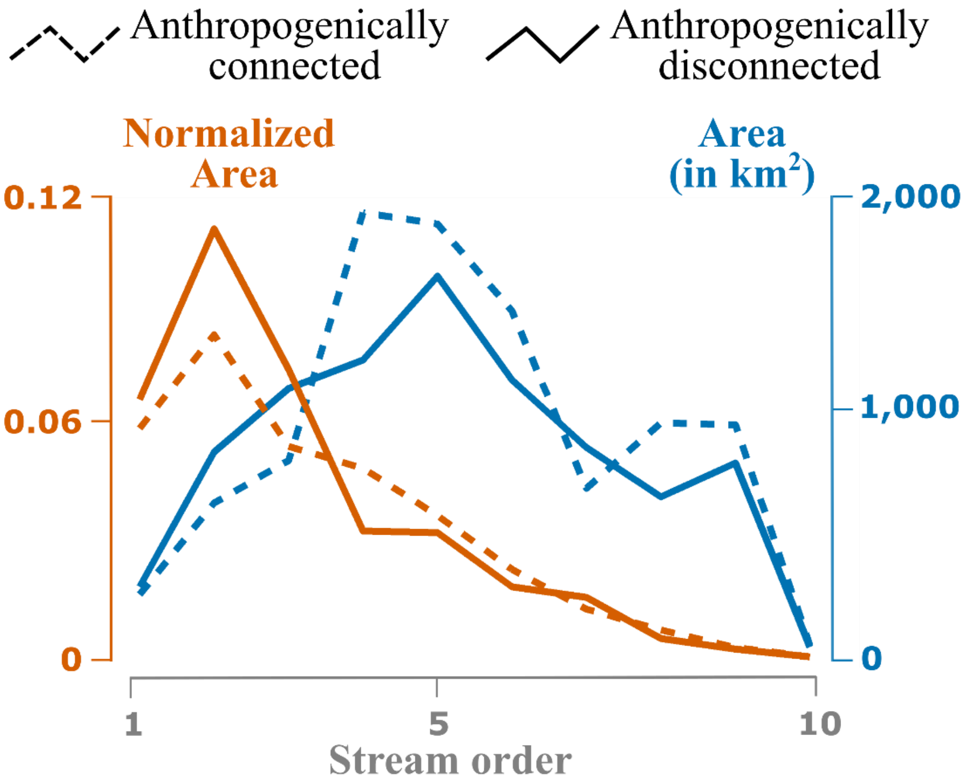


Figure 3.3. Stream order analysis of anthropogenically connected and disconnected areas. Actual and normalized areas in the CONUS, distinguished by stream order. Areas are normalized by stream order contributions to the agreement areas.

3.3 Discussion

Our finding that the anthropogenically connected extent was larger than the anthropogenically disconnected floodplain extent (Table B.2) was unexpected, although the 811 km^2 difference was much less than 1% of the agreement area floodplain. This corroborates other research illustrating the unintended upstream and downstream flooding caused by artificial levees (e.g., Tobin, 1995; Criss and Shock, 2001; Heine and Pinter, 2012; Czech et al., 2016).

Where artificial levees disconnect floodplains, their presence impacts active floodplain area through two processes; simple floodplain disconnection and lateral flowline alteration (Figure 3.4). The former occurs when artificial levees disconnect floodplains and river channels, especially along larger stream systems (Figure 3.4a). The end result of this process is a reduction in active floodplain area. The latter, lateral flowline alteration, involves an adjustment of the direction of flood waters and shifts the location of flooding (Figure 3.4b). Instead of decreasing the active floodplain, floodplain location is shifted from one location to another. In this example, the course of the river channel is adjusted and channelized through an artificial channel with levees. The result is the disconnection of the former channel and floodplain from floodwaters. Floodwaters are conveyed to the bottom of the figure where the channel is leveed on one side only, resulting in both anthropogenically connected and disconnected floodplains. With the exception of one other study (Scheel et al., 2019), this effect of artificial levees on floodplain extent has gone unreported until now, despite the well-known ability of levees to increase stage height (Heine and Pinter, 2012). This type of alteration is a result of the massive degree of topographic adjustment represented by the construction of enough artificial levees to wrap around the Earth six times (Knox et al., 2022a). The concentration of artificial levees along smaller streams (73% of artificial levees are along streams of orders 2 through 6) (Knox et al., 2022a) indicates the ability of this process to affect floodplain connectivity in ways that do not fit the normal conceptual model of artificial levees, which is based on larger stream systems (e.g., Heine and Pinter, 2012). The discovery of lateral flowline alteration in addition to the traditional understanding of simple floodplain disconnection is the latest facet of our understanding of the Anthropocene.

Gilbert White noted that the main policy aim of the last century was to minimize losses on floodplains instead of maximizing social benefits (White, 2000). In spite of that aim and the expenditure of billions of dollars on flood protection projects, flood losses in the US continued to rise and were 2.5 times higher during the period 1951-1985 than 1916-1950 (Tobin, 1995).

What insight can this study provide to this problem? I found that if I considered cultivated (cultivated crops and hay pasture) and developed land covers as those susceptible to economic losses, those areas cover 297,794 km^2, which is 66%, 59%, and 30% of the anthropogenically connected, disconnected floodplain, and agreement area floodplains, respectively. These estimates corroborate recent research indicating the large scale conversion of Mississippi River basin floodplains to cultivated and developed land covers during the last 60 years (Rajib et al., 2021). The preeminence of cultivated land covers impacted by artificial levees in the CONUS reflects the intersection of the huge concentration of levees in the Mississippi basins (40% of levee length in the CONUS is in the Lower and Upper Mississippi basins) (Knox et al., 2022a) with the degree of agricultural intensification in the same basins (Wohl, 2018). The association of wetland drainage with cultivation (Sparks, 1995; Wohl, 2018) indicates the reason for the disconnection of more than 1,500 km^2 of wetlands by artificial levees (Figure 3.2). These trends also reflect artificial levee association with certain land covers, with 67% of levees situated on developed or cultivated land covers in the CONUS (Knox et al., 2022a). Cultivated and developed land covers constitute 30.6% of floodplain areas and 3.7% of the entire CONUS. That nearly one-third of floodplain areas in the CONUS are used for some sort of economic purpose likely explains at least one of the causes for the trend noted by Tobin (1995) and White (2000).

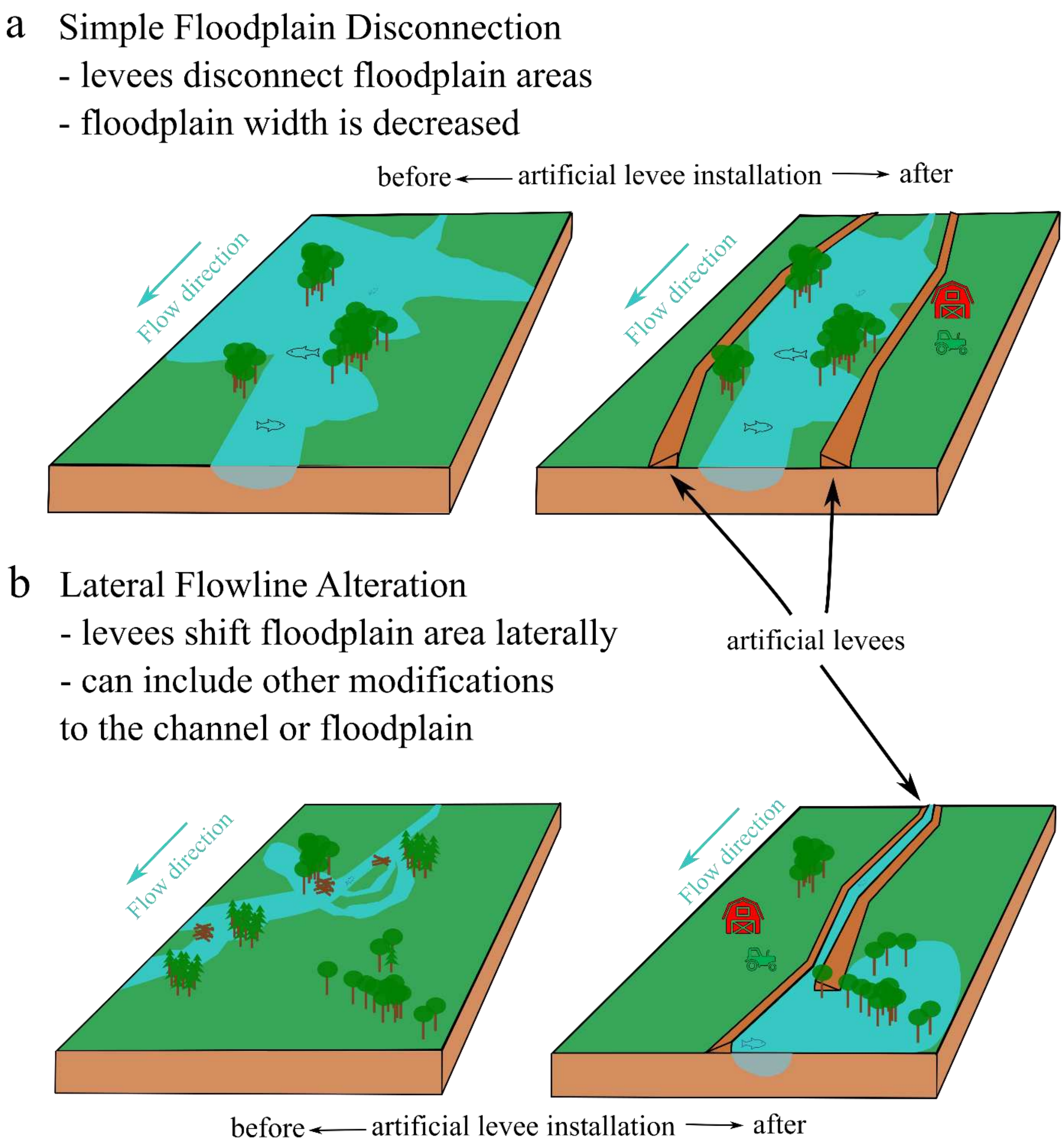


Figure 3.4. Two examples of floodplain alteration before and after levee installation. (A) In simple floodplain disconnection, levees disconnect floodplains and rivers. (B) Lateral flowline alteration occurs when levees alter the spatial extent of floodwaters, causing areas to flood (called "anthropogenically connected") and areas to disconnect (called "anthropogenically disconnected"). This type of alteration can occur with other modifications to include

channelization, rerouting of tributary inputs, levee construction on one side of the stream only, and cut and fill from channel or levee construction.

The anthropogenically connected and disconnected floodplain areas in the Lower Mississippi basin are notable given their large magnitude and the size difference, with the anthropogenically connected areas ~70% larger than the anthropogenically disconnected floodplain areas (Table B.2). This estimate, that the area flooded by artificial levees is 70% larger in extent than areas “protected” by levees, deserves some exploration. Each of these areas are created by floodwaters with the same upstream contributing area. An analogy is pouring one cup of water into a shallow bowl and then again into a tall, narrow vase. The same amount of water results in a different cross sectional area. Therefore, we tested the idea that floodplain geometry differences are responsible for the seemingly large difference in extent. We generated slope maps for the unmodified and modified DEMs and calculated the max and median values in each floodplain segment. The anthropogenically disconnected floodplain segments experienced greater slope, despite having artificial levees removed from their margins (Table B.3).

This supports the idea that, in the LMR basin, more confined anthropogenically disconnected floodplain areas result in a smaller floodplain extent given the same contributing area and reflects the different processes that formed each area. Even in a dynamic system such as the LMR, anthropogenically disconnected floodplains are formed by fluvial and floodplain processes operating over hundreds of years. Anthropogenically connected areas have only recently experienced the same processes. We contend that similar dynamics, with artificial levees altering flow paths across a heterogeneous topography, result in the differences apparent in Table B.2 between anthropogenically connected and disconnected floodplain areas.

Limitations of these results include the application of the hydrogeomorphic floodplain model in areas with characteristics that can lead to lower model accuracy (e.g. dry, steep, flat areas or those near the coast) (Annis et al., 2022; Lindersson et al., 2021). Calibration of the floodplain model at the 2-digit HUC basin level provides some mitigation. Another limitation stems from the current inability to ground-truth potential levees identified in chapter 2 and the absence of a stream order-dependent buffer size for topography modification.

I removed known and potential artificial levee locations from a modified 1-arc second DEM of the contiguous United States. I then generated two hydrogeomorphic floodplains using the modified and unmodified DEM and compared the location and area, land cover, and the stream order of rivers associated with each floodplain segment. The overall effect of artificial levee removal was not to just extend the floodplain, but rather to shift the location of flooding. The massive extent and length of artificial levees, especially along smaller streams (Knox et al., 2022a), requires us to realize that floodplain alteration by artificial levees extends beyond normal conceptions of embankment. Constructed by individual farmers, municipal boards, and state and federal agencies over a 300 year period (Hudson et al., 2008), artificial levees constitute a massive topographical alteration at the CONUS level that alters floodwater flow paths, especially along smaller streams. This previously unknown dimension of artificial levee impacts to floodplains illustrates that we have massively underestimated the ecological and hydrological damage of levees. Anthropogenically disconnected floodplain (protected from flooding) and anthropogenically connected (induced to flood by artificial levees) areas each accounted for about 1% of the total CONUS floodplain, which was more than 960,000 km^2. More than 60% of the disagreement areas (mapped floodplain that differed with and without artificial levee presence) were cultivated, forested, wetland, or developed land cover. More than 30% of the

CONUS floodplain was either cultivated or developed. These results corroborate, on a national scale, previous local-scale investigations of the unintended consequences of artificial levees.

3.4 Materials and Methods

3.4.1 Experimental Design

Our analyses included the following major steps. First, I generated GFPLAIN floodplain areas for each of the 18 two-digit HUC (HUC2) basins in the CONUS using the 30-m resolution USGS National Elevation Dataset (NED) (Appendix B). Then, I altered the topography in each basin by deleting known and potential levees from the topography and applied GFPLAIN to the modified topography. Finally, I analyzed the differences in floodplain extent for the unmodified topography and the modified topography by stream order, land cover, and area.

3.4.2 Topography Modification

This procedure is similar to DEM modification by Scheel et al. (2019) in which the topographic effect of levees are removed from the DEM (Figure B.2). I developed an ArcGIS Pro (Esri, Inc., 2021) model that separately modifies topography near NLD levees and near potential levees before combining results into one DEM. The same procedure is applied to both types of levees. First, the centerline of each levee is identified. Then, the centerline is buffered by 90 m. This area, within 90 m of the centerline, is the only area in which topography is adjusted during the process. The 90 m buffered area is deleted from a second larger buffered area of 150 m beyond the original 90 m buffer, creating a ring of unmodified topography varying in distance of 90 to 240 m from the levee centerline. The focal mean tool, with radius of 120 m, is then applied to the area of the original 90 m centerline buffer, using the values of the ring of unmodified topography. Finally, three separate DEMs are combined together using the minimum value and the mosaic tool: the unmodified DEM, the modified DEM using NLD centerlines, and the modified DEM using potential levee centerlines.

3.4.3 Statistical Analysis

I developed custom ArcGIS Pro (Esri Inc., 2021) and R (R Core Team, 2020) scripts to analyze the differences between the GFPLAIN floodplain extent developed from unmodified and modified topography. Working by HUC2 basin, I identified areas of agreement and disagreement. Analysis focused mainly on the latter because areas of disagreement are created solely by the presence or removal of artificial levees.   Areas of disagreement between the two floodplains were classified as either anthropogenically disconnected floodplain or anthropogenically connected and were analyzed using ArcGIS Pro. Anthropogenically disconnected floodplain are those separated from overbank flow by the installation of artificial levees. Anthropogenically connected areas are those that are caused to flood by the installation of artificial levees. These areas were measured in terms of square kilometers and their coverage in the 2016 National Land Cover Database (NLCD) (Table B.1) was determined in ArcGIS Pro. I determined the largest stream order associated with each floodplain segment by searching in ArcGIS Pro within 500 m of each segment for every stream segment in the National Hydrography Dataset (Table B.1). I used R to process these data and select the largest stream order per floodplain segment using the map_dfr function in the purrr package (Henry and Wickham, 2020) and the group_by and summarise functions in the dplyr package (Wickham et al., 2020). I chose 500 m as the search radius after using several smaller values in the Lower Mississippi River HUC2 basin and determining that this search radius connected NHD segments with most floodplain segments (n ~ 60,000 of 66,000 total segments) without connecting unrelated stream and floodplain segments.

Data and materials availability: Polygon files for the agreement, anthropogenically disconnected floodplain, and anthropogenically connected areas are available publicly at https://doi.org/10.4211/hs.15c4ab0ebfe7447298b18af37caf4e0e.

Chapter 4 Levees don't protect, they disconnect: how artificial levees impact floodplain function

Summary

Despite the recognition of floodplain importance in the scientific community, floodplains are not afforded the same legal protection as river channels. In the United States alone, flood-related economic losses were much higher in the second half of the 20th century than the first half despite the expenditure of billions of dollars on flood defenses. Partially to blame are the low appraisal and understanding of human impacts to floodplain functions. Here, I explore the impacts of levees on floodplain functions and analyze case studies of floodplain restoration through levee removal. Floodplain functions include (1) fluxes of water, solutes, and particulate materials; (2) enhanced spatial heterogeneity of hydrology and biogeochemistry; (3) enhanced habitat abundance and diversity; (4) enhanced biomass and biodiversity; and (5) hazard mitigation. Case studies of floodplain restoration involving artificial levee adjustment are heavily concentrated in North America, Europe, and Japan, and those case studies assess floodplain functions within 30 years of restoration. In the United States, restoration through levee removal comprises less than 1% of artificial levee length and 1-2% of disconnected floodplains. Most case studies were impacted by stressors outside the study site, such as flow regulation or artificial levees, and took place in lowland alluvial rivers. Reconfiguration was successful at achieving limited aims while reconnection set floodplains on a trajectory to more fully restore floodplain functions. Case studies illustrated the tension between restoration scale and study resolution in time and space as well as the role of site-specific characteristics in determining restoration outcomes (Knox et al., 2022c).

4.1 Introduction

Floodplains are recognized by river scientists as a critical component of river ecosystems. Conceptual models emphasizing the role of floodplains include the flood pulse concept (Junk et al., 1989), riverine productivity model (Thorp and Delong, 1994), shifting habitat mosaic (Tockner et al., 2010), and the river wave concept (Humphries et al., 2014). Nevertheless, floodplains are not afforded the same legal protection as active channels (e.g., US Clean Water Act, EU Water Framework Directive) despite a long history of human alteration of floodplain forms and processes. Predictable flood pulses on large rivers in Egypt and Mesopotamia led to the development of some of the world's first complex societies (Butzer, 1976; Sparks, 1995), and subsequent cultures focused development on floodplains and learned to exploit floodplain functions (Tockner and Stanford, 2002). Collectively, human activities have resulted in simplified floodplains with substantially reduced functions along many rivers (Peipoch et al., 2015).

Despite a policy of minimizing flood losses during the 20th century (White, 2000), the United States (U.S.) experienced 2.5 times the economic losses ($3 billion annual flood losses) in the second half of the century after the expenditure of billions of dollars on flood protection projects (Tobin, 1995). An estimated 98% of the 5.3 million km of U.S. rivers are impacted by human activities (Graf, 2001). Some components of these activities, such as flow regulation, have been documented nationally and globally (e.g., Graf, 1999; Lehner et al., 2011; Grill et al., 2019), as have the impacts of roads and railroads (e.g., Blanton and Marcus, 2009). One key component of modern floodplain management is the construction of artificial levees (Hudson et al., 2008). However, the national or global impacts of artificial levees have not been evaluated until recently (Knox et al., 2022a, 2022b), mainly because of incomplete databases of artificial levees.

Natural fluvial levees are long, ribbon-like bodies of sediment deposited at river channel-floodplain margins when floodwaters lose competence (Brierley et al., 1997), whereas artificial levees are human-made linear features constructed between channels and floodplains to contain peak flows in the channel (Tobin, 1995). I define the floodplain as a frequently flooded, low-relief landform created by erosional and depositional processes under the contemporary hydrologic regime (Dunne and Alto, 2013). In previous chapters, I identified over 182,000 km of undocumented artificial levees in the contiguous U.S. (CONUS) and determined that the overall effect of artificial levees on flooding in the CONUS was to shift the location of inundation (Knox et al. (2022a), Knox et al. (2022b)). I also determined that over 30% of the CONUS 100-year floodplain was either cultivated or developed land cover. This prior work emphasized the problematic effects of artificial levees at a national scale rather than local scales, which has been the focus of most studies assessing the impacts of levees.

Local and regional studies have found that artificial levees shift the location of flooding by increasing stage upstream from levees and increasing downstream conveyance (Tobin, 1995; Criss and Shock, 2001; Heine and Pinter, 2012; Czech et al., 2015; Knox et al., 2022b). This can lead to increased channel velocities, bed coarsening, and incision (Frings et al., 2009). Levees also limit lateral connectivity and the exchange of nutrients, sediment, and organisms between the channel and floodplain, resulting in significant ecological harm (Blanton and Marcus, 2009; Sparks et al., 2017; Wohl, 2018). The presence of artificial levees encourages human development of floodplains and increases the vulnerability of populations and infrastructure to flood damage (White et al., 2001; Pinter, 2005).

The loss of floodplain functions resulting from artificial levee installation is seldom considered when calculating economic losses from flooding (Opperman et al., 2009; Jacobson et al., 2015). That accounting requires a better understanding of floodplain functions, which is a non-trivial pursuit given floodplain ecosystem complexity and floodplain connection to the atmosphere, channels, and uplands (Gren et al., 1995). Although there are a myriad of floodplain functions, I group them into the following major categories (Figure 1):

(i) fluxes of water, solutes, and particulate materials;

(ii) enhanced spatial heterogeneity of hydrology and biogeochemistry processes;

(iii) enhanced habitat abundance and diversity;

(iv) enhanced biomass and biodiversity; and

(v) hazard mitigation.

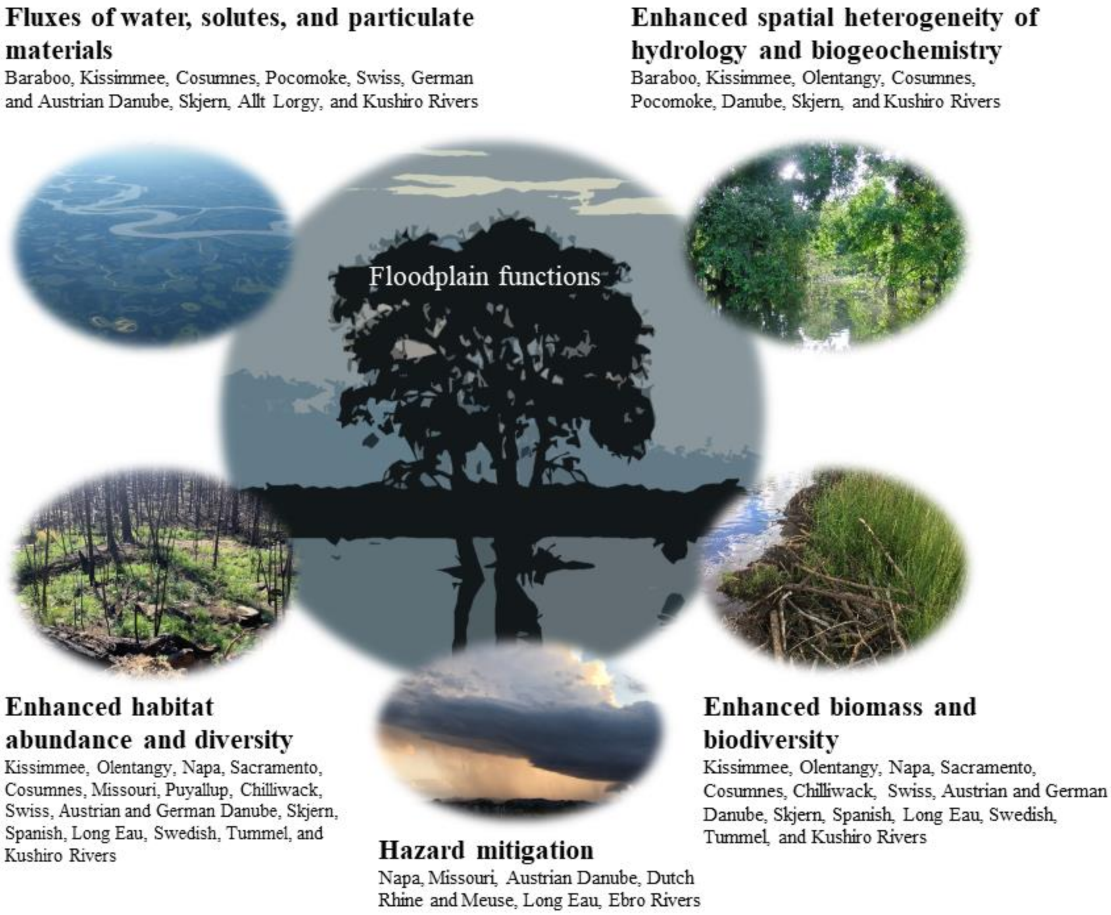


Figure 4.1. Floodplain functions and selected rivers with case studies (Tables 4.1 and 4.2) analyzing the impacts of restoration impacting that function.

This study sets out to address knowledge gaps in understanding the cumulative effects of artificial levees on floodplain functions; understanding floodplain functions across diverse river settings in an integrative manner; and the effectiveness of floodplain restoration involving artificial levee alteration. Our previous work in chapters 2 and 3 illustrated the effects of more than 228,000 km of artificial levees (almost enough levees to wrap around Earth six times) on floodplain extent in the U.S. (Knox et al., 2022a, 2022b). If the cumulative effects of artificial levees are anything close to those documented for dams (which store a year's worth of runoff in the U.S. (Graf, 1999)), then the consequences are of massive ecological significance. To grasp this significance, I present floodplain functions in an integrative way rather than the common approach that separates closely linked physical, chemical, and biological processes (Wohl, 2021). The high degree of floodplain functions' interdependence (discussed in the next section) magnifies the damaging impacts of artificial levees. The urgency to understand floodplain functions from an integrative perspective is that much more acute given the role of artificial levees as just one of many stressors that impact over 98% of river kilometers in the U.S. (Graf, 2001). Given that, I wanted to assess the role of floodplain restoration involving artificial levee removal or alteration to understand how this damage can be ameliorated.

Here, I review these functions and how they are compromised by artificial levees, using case studies from North America, Europe, and Japan as examples. I draw on this review to support the contention that artificial levees do not just protect human infrastructure, they also disconnect a vitally important component of the landscape that is of great value to humans and the environment.

4.2 Why it matters – floodplain functions

The importance of floodplain functions to ecological and human wellbeing is multifaceted and has been documented in the scientific literature for the past half century. I briefly highlight the current knowledge of floodplain functions and their importance below.

4.2.1 Material fluxes.

Non-living material stored on floodplains includes water, solutes, sediment, particulate organic matter (POM; > 0.45 μm in diameter), and large wood (≥ 10 cm diameter and 1 m length) (Wohl, 2021). Inundation hydrology describes the many sources of surface and subsurface water present in floodplains including groundwater, tributaries, overbank flow, overland flow from adjacent uplands, and precipitation (Mertes, 2000). Solutes stored on floodplains include dissolved forms of nitrogen (Noe et al., 2013), phosphorus (Records et al., 2016), and organic matter (Cuffney, 1988) present in surface and subsurface waters. Sediment is stored in floodplain features through vertical accretion from overbank flows, lateral accretion and channel-fill deposits from channel migration and avulsion, and colluvial and eolian deposits (Allen, 1965; Meade and Moody, 2010). POM such as leaf litter is heavily influenced by the type of riparian vegetation and the season (Tank et al., 2010) and can enter floodplains from channels, adjacent uplands (transport by wind, overland flow, and tributaries), and direct litterfall from floodplain vegetation. Reduced by orders of magnitude by human influence in many river systems, dead biomass in the form of large wood creates physical and ecological functions on floodplains (Wohl et al., 2019). Large wood can also enter floodplains from adjacent uplands, overbank flow from channels, or direct recruitment from floodplain forests (Wohl, 2020).

The most widespread human alteration to floodplains is disconnection from stream flow, which alters the volume and duration of storage of all materials by severing transport onto the floodplain from the channel; altering surface and subsurface water storage in the floodplains and associated biogeochemical processes and decay; or reducing floodplain erosion and deposition (Wohl, 2021). Although artificial levee installation is one of several ways to disconnect streams and floodplains, artificial levees can also drastically change floodplain storage through several mechanisms leading to terrestrialisation (Tena et al., 2020). In terms of water storage, the degree to which disconnected floodplains become like the adjacent uplands is dependent on other water inputs from tributaries and precipitation (e.g. Park and Latrubesse, 2017), groundwater inflow (e.g. Burt, 1996), and subsurface connection to the channel (e.g. Kupfer et al., 2015). Although the elimination of overbank flooding and lateral migration can decrease water and sediment storage, within-channel fluctuations in discharge can still influence low-lying floodplain areas such as secondary channels and floodplain wetlands (Tockner et al., 2000; Hudson et al., 2012;

Lininger and Latrubesse, 2016). Additionally, reduced lateral channel movement and reworking of floodplain sediments associated with artificial levee installation and bank stabilization may increase POM storage in some cases (e.g. Sutfin et al., 2021). Levees may be much less influential for organic carbon storage in riparian forests compared to the degree of forestation and groundwater fluctuations (Rieger et al., 2014).

Floodplain storage contributions to base flow in un-altered tropical streams (e.g. Lininger and Latrubesse, 2016) indicate the potential impacts to river systems from disruptive interventions, such as levees. Artificial levees reduce dissolved and particulate carbon, nitrogen, and phosphorus input from channels to floodplains (Noe and Hupp, 2005). Artificial levees can decrease floodplain sedimentation to 0 mm/yr, with accidental breaches providing the only sediment supply in the form of sand-splays (Florsheim and Mount, 2002; Florsheim and Mount, 2003). Sediment storage loss can occur during levee breaches when simultaneous sand splays and scour reposition sediment covering more than hundreds of thousands of hectares on large rivers, resulting in anthro-geomorphic pond features called “wielen” in Dutch (Galat et al., 1998; Hudson et al., 2008).

Even though artificial levee installation impedes the influence of overbank flow on floodplain storage, the results can be far reaching due to secondary effects such as terrestrialisation and organism extinction; the interconnectedness of floodplains with the atmosphere, channel, and uplands; and other human activities common on floodplains with levees, such as land cover changes. In the CONUS, artificial levees are most associated with land cover changes to cultivated land covers (Rajib et al., 2021; Knox et al., 2022b). The installation of agricultural drain tiles can result in sediment compaction, reduced recharge to floodplain aquifers, and loss of storage through increased drainage (Blann et al., 2009). The link between hydrologic connectivity and spatial diversity of geomorphic units (Hudson and Colditz, 2003; Hudson et al., 2012; Park and Latrubesse, 2017) indicates that floodplain simplification brought about from levee construction can lead to storage alteration. Conversion to agricultural activities can increase nitrate storage because of application of artificial fertilizers (Wang et al., 2013). Phosphorus will follow a similar trend given the agricultural sources of phosphorus and the role of floodplains as a phosphorus sink (Sharpley et al., 2013). Complex organic carbon dynamics rely on many different factors (Sutfin et al., 2016), with different case studies illustrating the varying impacts of levees, degree of forestation, and groundwater fluctuations on carbon storage (Hanberry et al., 2015; Wohl, Hall, et al., 2017).

4.2.2 Enhanced spatial heterogeneity of hydrology and biogeochemistry.

Floodplains are landforms in which the mixing of waters with different sources leads to increased biogeochemical reactions (Harvey and Gooseff, 2015). Considered ecosystem control points, floodplains host important biogeochemical processes that greatly impact ecosystem dynamics (Appling et al., 2014; Bernhardt et al., 2017). Water is the medium that transports energy, solutes, and particulates between hillslopes, floodplains, channels, and other areas (Covino, 2017) and is considered the master driver of floodplain structure and functions (Wohl, 2021). Variations in the spatial mixing of surface waters in the perirheic zone can strongly influence floodplain geomorphology, biogeochemical reactions, and habitat (Mertes, 1997). Productivity is related to the juxtaposition of heterogeneous flow paths and microbial activity (Dwivedi et al., 2018). Floodplains are commonly depositional, low-energy locations with high primary productivity where large amounts of organic matter mix with reactive nitrogen and phosphorus (Noe, 2013). When not inhibited by artificial levees, this productivity often coincides with flood pulses (Tomasek et al., 2019). Repeated flooding and associated recycling of organic matter and nutrients is the principal driver of productivity in a river-floodplain system (Junk et al., 1989). Flood pulses enable the creation of a stream-soil interface on floodplains where numerous biogeochemical reactions occur due to the abundance of electron donors and acceptors (Hedin et al., 1998). The specific type of floodplain inundation highly affects biogeochemical processing by altering the relationships between vegetation, microbial activity, and chemical reactions (Baldwin and Mitchell, 2000). Inundation dynamics impact nutrient decomposition and mineralization rates (Brinson et al., 1981) as well as translocation of nitrogen and phosphorus within floodplain vegetation (Clawson et al., 2001). Phosphorus dynamics are complex and involve the interplay between soil, hydrologic conditions, and climate (Records et al., 2016). Phosphorus processes include biological assimilation, sorption to sediment, and precipitation reactions of inorganic salts (House, 2003). Soil nutrient mineralization rates of nitrogen and phosphorus are enhanced by greater inputs of sediment and water to floodplains (Noe et al., 2013). Artificial levees lead to decreased nutrient exchange between floodplains and rivers (Jenkins and Boulton, 2003). Disconnection of floodplains leading to terrestrialisation completely changes the floodplain inundation hydrology, significantly impacting biogeochemical reactions (Sanchez-Perez and Tremolieres, 2003). Decreased sediment storage on disconnected floodplains impacts biogeochemical reactions involving the sorption of phosphorus to sediment (Darke et al., 1996). Disconnection can also lead to a large decrease in organic matter delivery and production on floodplains, thus impacting multiple biogeochemical reactions and food webs (Heiler et al., 1995).

4.2.3 Enhanced habitat abundance and diversity.

Floodplain habitat diversity derives from different patterns of hydrologic, solute, and sediment connectivity interacting over spatially heterogeneous landscapes (Bayley, 1995). Junk et al. (1989) emphasized the seasonal flood pulse as a primary driver of these dynamic interactions. Applying the process domain concept (Montgomery, 1999) in this context, floodplains are the dynamic canvas on which flooding provides a natural and biologically advantageous spatial and temporal habitat disturbance (Junk et al., 1989; Tockner and Ward, 1999; Arscott et al., 2002) and sufficient space to accommodate diverse habitats (Bellmore and Baxter, 2014).

Disturbance can be considered a physical force or process that stresses an ecological system relative to its reference state (Rykiel, 1985) and can be of natural (e.g. flood, wildfire, drought) or anthropogenic origin (Magoulick and Kobza, 2003). Anthropogenic disturbances alter the floodplain physical landscape so that fundamental geomorphic thresholds are broken, making floodplains less resilient to future disturbance (Brierley et al., 2005; Karpack et al., 2020). I define resilience as the degree to which a system can persist by absorbing disturbance and maintaining similar relationships between populations and driving variables (Holling, 1973). In floodplains, resilience derives from the pathways by which hydrobiogeomorphic complexity and nested feedback loops are able to absorb disturbance and maintain equilibrium (Wohl et al., 2021). Fully functioning floodplains possess multiple process domains and diverse biota that facilitate small adjustments to hydrologic, geologic, biologic, or anthropogenic disturbances (Castro and Thorne, 2019). Biogeochemical cycling of nutrients provides a number of dynamic feedback loops that mediate ecological and hydrogeomorphic disturbance across multiple temporal and spatial scales (Atkinson et al., 2018).

The spatiotemporal heterogeneity and connectivity of floodplains enhance not only floodplain resilience but the resilience of the entire basin (McCluney et al., 2014; Wohl et al., 2022). In the same sense, fully functioning floodplains can be considered as river beads, which are spatially heterogeneous locations within the river network whose ability to store water and organic material, facilitate biogeochemical reactions, and enhance biodiversity lead to greater resilience in the entire network (Hauer et al., 2016; Wohl et al., 2018). River beads were originally described for mountain stream networks in which river segments with floodplains alternate downstream with laterally constrained segments with little to no floodplain development (Stanford et al., 1996). Along lowland rivers with laterally extensive and longitudinally continuous floodplains, the entire length of the river corridor acts as a bead. Resilience is enhanced by hierarchically organized physical, chemical, and biological processes operating across overlapping habitat scales (Beechie et al., 2010). Contextualizing floodplains within the basin and emphasizing connection as far as the ocean (Wohl and Iskin, 2021) indicates the potential role floodplains can play in resilience.

Artificial levees ultimately decrease floodplain habitat diversity and complexity through the elimination of hydrobiogeomorphic pathways and feedback loops by which floodplains can respond to disturbance. This begins with the reduction of water and sediment resulting from disconnection from stream flow. Levees disrupt the flood pulse that is the driving force of floodplain productivity. This disruption leads to the elimination of numerous biogeochemical pathways by which the floodplain can respond to disturbance. Cascading effects within food webs ensure the subsequent elimination of numerous biota that can no longer contribute to floodplain functions. The end state of artificial levee impacts to floodplains is a spatially and temporally homogenized floodplain with reduced biogeochemical activity that has minimal ability to absorb disturbance (Poff et al., 2007; McCluney et al., 2014; Wohl, Lininger, et al., 2017; Bouska et al., 2019).

4.2.4 Enhanced biomass and biodiversity.

The documented high biomass and high biodiversity of floodplains (Naiman et al., 1993) result at least in part from habitat diversity. High biodiversity in floodplains is based on organic matter productivity and food webs for numerous fish and other organisms (Opperman et al., 2017). Floodplains provide the habitat availability and connectivity needed for fish at different times in their life cycles (Schiemer, 2000). Floodplains can contain a great diversity of aquatic invertebrates and greater community respiration because of microbial activity (Bellmore and Baxter, 2014). The role of microbial communities in biogeochemical reactions in floodplain soils are directly impacted by hydrologic connectivity and organic matter availability (Argiroff et al., 2017). Organic carbon stocks strongly predict bacterial production (Cole et al., 1988), whereas hydrologic connectivity exerts an equally important control on bacterial community composition and the degree of enzymatic activity (Mayr et al., 2020). The key control on denitrification rates are microbial processes dictated by nitrate and oxygen concentrations (Berhardt et al., 2017). Microbial activity and different conditions imposed by hydrologic connectivity are two factors that influence mineralization of nitrogen and phosphorus, a key bottleneck geochemical process (Noe et al., 2013).

Connectivity between channels and floodplains enlarges habitats and biological productivity (Jenkins and Boulton, 2003). Amoros and Bornette (2002) emphasize the importance of connectivity operating at different spatial and temporal scales. They recognize four major habitat components (water temperature, suspended solids/turbidity, nutrient content, and substrata composition) that heavily influence biodiversity operating across the scales of a hydrologically connected floodplain and individual waterbodies. Amoros and Bornette (2002) describe a different set of processes occurring across the two spatial scales at two different time scales. The flood pulse drives different amounts of connectivity on monthly to yearly time scales, influencing productivity, nutrient exchange, biogeochemical processing, and the exchange of organisms whose life cycles are dependent on varying environmental conditions. This mosaic of patterns can result in antagonistic processes whereby complex responses drive gradients in different directions at varying times and locations (Amoros and Bornette, 1999). Examples of these processes, such as ecological succession, lateral channel migration, and river bed incision, enhance biodiversity at decadal and longer time scales by balancing the trend towards terrestrialisation with the formation and rejuvenation of water bodies (Ward and Stanford, 1995; Amoros and Bornette, 2002).

Biogeomorphic agents lend functional floodplains a self-healing capacity (Johnson et al., 2020). Riparian vegetation acts as a buffer between floodplains and streams and helps to trap and store particulate matter and facilitate biogeochemical uptake of solutes (Schlesinger et al., 1996). Riparian vegetation and large wood (LW) mediate disturbance events by providing localized resistance (Brooks and Brierley, 2002). Wetlands in functional floodplain lakes exhibit greater resilience to drought (Shi et al., 2017). Beaver (*Castor* spp.) increase resilience to drought and wildfire (Hood and Bayley, 2008; Fairfax and Whittle, 2020) through the creation of spatially and environmentally complex beaver meadows (Westbrook et al., 2011).

Artificial levee installation disrupts the flood pulse that is the driving force for productivity in formerly connected channels and floodplains. This disruption alters the disturbance regime around which most floodplain processes are based. Levees disrupt every aspect of biologically complex floodplains as conceived by Amoros and Bornette (2002). Levees decrease hydrologic connectivity important at the floodplain scale and can alter the individual waterbody through terrestrialisation. Temporally, levees and other engineering disturbances homogenize the natural rhythm of flood pulses (Moyle and Mount, 2007; Poff et al., 2007). At larger temporal scales, artificial levees and associated engineering works, such as bank stabilization and dams, either freeze floodplain processes (e.g., channel migration) or completely alter processes (e.g., ecological succession and incision). Artificial levees decrease edge habitat and ecosystem diversity (Florsheim and Mount, 2003). Reviews of floodplain habitat restoration efforts indicate the deleterious effect of artificial levees on species diversity (Roni et al., 2019). The terrestrial and disconnecting effects of artificial levees are especially deleterious to floodplains because floodplain foodwebs are based on allochthonous and autochthonous carbon sources (Opperman et al., 2017). The role that connectivity plays in the bottleneck processes of N and P mineralization indicates another fundamental way that levees alter floodplain ecosystems.

4.2.5 Hazard mitigation.

Fully functional floodplains offer mitigation against a wide range of natural and anthropogenic hazards (Sheaffer et al., 2002). They store flood water and attenuate peak flows (Woltemade and Potter, 1994; Škute et al., 2008; Lininger and Latrubesse, 2016); sequester contaminants from non-point sources and provide a site for biogeochemical remediation of some types of contaminants (Marron, 1992; Dennis et al., 2009; Gordon et al., 2020); attenuate downstream fluxes of sediment following upland disturbance (Poeppl et al., 2017; Wohl et al., 2022); and provide refugia for organisms during natural and anthropogenic disturbances (Sedell et al., 1990; Stella et al., 2011).

Artificial levees are effective up to their design standard (Tobin, 1995). Consequently, adjacent floodplains that have been disconnected by the levees offer little mitigation against flood waters and peak flow attenuation, except during extreme floods that cause levees to fail (e.g., the 1993 Mississippi River flood, Galloway (1995)). Impacted floodplains can only process and remove pollution from non-point sources to the degree that connection exists between the non-point sources and the floodplain (e.g., are the sources upstream or on the floodplain itself?) and the degree to which the necessary biogeochemical processes remain functional despite disconnection. Disconnected floodplains cannot attenuate fluxes of sediment or waste products from point sources, such as metal mining, unless those fluxes are delivered by hillslopes fringing the floodplain.

4.3 Synthesis discussion- What do case studies tell us?

A thorough literature review of floodplain restoration studies published in English-language journals indicates that the great majority of floodplain restoration projects have been undertaken in North America and Europe. I used Google Scholar and began with key word searches to include "levee setbacks", "setback levee", "restoration levee", "levee removal", and "reconnected floodplain". I also used similar terms with geographic place names or programs from known or suspected restorations (e.g., "RFR levee removal"). Google Scholar returns thousands to tens of thousands of articles for most of these searches, so I quickly transitioned to using Google Scholar to data-mine references and citing papers, searching out an increasingly larger web of related research. Works that were especially helpful in this regard include Gumiero et al. (2013), González et al. (2015), and Opperman et al. (2017). Selected case studies of floodplain restoration that involve artificial levee alteration (removal, notching, lowering, or setting back) in North America (Table 1) and Europe and Japan (Table 2) illustrate the wide-ranging impacts of artificial levees on floodplain functions. Here, I discuss the major implications of the case studies. I separate North American and European/Japanese case studies based on the much longer period of human involvement in Old World (Europe) river and floodplain management (e.g., Hudson et al., 2008) and the more gradual geomorphic readjustments of Old World river systems compared to New World (colonial) river systems (e.g., Brierley et al., 2005). I hypothesize that, given these differences, the response of floodplain functions to restoration efforts will be noticeably different. The degree of artificial levee adjustment (low, medium, high) is categorized in the "magnitude" column of each case study, with the range set by the case studies. "Low" indicates alterations made to single levees in one or several places with impacts that can be described along a river reach less than 10 kilometers long. "High" indicates alterations made to artificial levees along river lengths measured in the hundreds of kilometers. "Medium" alterations fall between low and high and the alterations to artificial levees are measured in tens of kilometers.

I organized the case study synthesis around four main themes but was able to discuss other ideas as well. The main themes are:

(i) Limits of selected case studies;
(ii) Reconnection and reconfiguration;
(iii) Restoration-scale dilemma; and
(iv) Unique place-based challenges.

Table 4.1. Selected case studies of efforts to restore floodplains by levee alteration in North America.

Location	Floodplain functions analyzed	Summary	Elapsed time[a] (years)	Magnitude[b]	Other stressors[c]	Reference
Baraboo River, Wisconsin, US	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry	controlled reconnected floodplain (a gate was installed in the levee) experienced water storage flux based on weather and high temporal and spatial denitrification rates	1-2	low	none known	Orr et al., 2007
Kissimmee River, Florida, US	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat diversity and biodiversity	reconnection of river and floodplain through canal and levee modification	1-20	high	headwater lakes are managed for flood control and biodiversity	Toth et al., 1998; Toth, 2010; Toth and van der Valk, 2012; Koebel and Bousquin, 2014; Jones, 2017; Koebel et al., 2021
Olentangy River, Ohio, US	enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat diversity and biodiversity	invasive species removal and levee breaches partially reconnect floodplain and decrease vegetation biodiversity while improving exchanges of total N and C	3-4	low	discharge is controlled by Delaware Lake releases; research site is in an urban watershed	Zhang and Mitsch, 2007; Swab et al., 2008
Napa River, CA, US	enhanced habitat diversity and biodiversity, hazard mitigation	varying responses of vegetation to levee removal in straightened reaches	8-30	medium	urban watershed	Bechtol and Laurian, 2005; Diggory and Parker, 2011.
Sacramento River, California, US	enhanced habitat diversity and biodiversity	levee setbacks resulted in greater abundance and diversity with larger elapsed time	3-12	high	discharge regulated by dams and diversions	Golet et al., 2008
Cosumnes River, California, US	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat diversity and biodiversity	levee breaches reconnected the river and floodplain; sand splay complexes added topographic variability; induced high levels of anaerobic ammonium oxidation, denitrification, and primary productivity; plant communities responded more stochastically	3-20	medium	discharge is unregulated by dams	Florsheim and Mount, 2002; Swenson et al., 2003; Ahearn et al., 2006; Sheibley, 2006; Trowbridge, 2007; Hoagland et al., 2019
Pocomoke River, Maryland, US	fluxes, enhanced spatial heterogeneity of	levee breaches improve trapping of P, N and sediment on floodplains, thereby improving water quality	1	low	watershed is heavily altered by human use	Noe et al., 2019

Location	Floodplain functions analyzed	Summary	Elapsed time[a] (years)	Magnitude[b]	Other stressors[c]	Reference
	hydrology and biogeochemistry					
Missouri River, Iowa, US	enhanced habitat diversity and biodiversity, hazard mitigation	levee setback resulted in reduced flood stages and improved biodiversity	5	low	discharge is heavily regulated for navigation, flood control, and power generation	Smith et al., 2017
Puyallup and Carbon Rivers, Washington, US	enhanced habitat diversity	levee setbacks on glacially fed river results in greater riparian habitat diversity	6	medium	discharge is regulated for power generation	Konrad et al., 2008
Chilliwack River, Canada	enhanced habitat diversity and biodiversity	floodplain reconnection resulted in new habitat for coho salmon	1-4	medium	none known	Ogston et al., 2014

Note. [a]Elapsed time indicates the years between levee alteration and data collection. [b]Magnitude indicates the degree of levee alteration. [c]Other stressors indicate other anthropogenic stressors that continue to operate in that river's basin.

Table 4.2. Selected case studies of efforts to restore floodplains by levee alteration in Europe and Japan.

Location	Floodplain functions analyzed	Summary	Elapsed time[a] (years)	Magnitude[b]	Other stressors[c]	Reference
Rhine, Rhone, Moesa, Hinterrhein, Emme, and Thur Rivers, Switzerland	fluxes, enhanced habitat diversity and biodiversity	comparison of carbon storage and soil organic matter stabilization from levee setbacks and natural floodplains; potential catchment scale effects on habitat and biodiversity relationships	4-11	medium	relocated banks stabilized; discharge is regulated on some rivers by dams or locks	Rohde et al., 2005; Pasquale et al., 2011; Bullinger-Weber et al., 2014
Danube River, Austria	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat diversity and biodiversity, hazard mitigation	reconnecting floodplains on free-flowing part of channel below Vienna impacts microbiota and fish guilds with mixed results	1-20	medium	flow is highly regulated by upstream chain of impoundments	Tockner and Schiemer, 1997; Tockner et al., 1998, 1999; Schiemer et al., 1999; Luer et al., 2007; Reckendorfer et al., 2013; Griselda et al., 2019; Ramler and Keckeis, 2019; Mayr et al., 2020
Danube River, Germany	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat diversity and biodiversity	floodplain reconnection and improved habitat diversity results in improved biodiversity; target riparian vegetation establishment inhibited by floods and limited study time	1-3	medium	flow is regulated for hydropower	Stammel et al., 2011; Pander et al., 2018; Stammel et al., 2021.
Rhine and Meuse Rivers, The Netherlands	enhanced habitat diversity and biodiversity, hazard mitigation	room for the river, which involves levee setbacks, reduces flood levels and flood consequences; other stressors reduced restored habitat and fish biodiversity over time	13-30	high	rivers and basins continue to be highly influenced by humans through canalization, flow regulation for hydropower, and land cover changes	Klijn et al., 2018; Schmitt et al., 2018; Stoffers et al., 2021
Skjern River, Denmark	fluxes, enhanced habitat diversity and biodiversity	restoration design allows limited reconnection of floodplain;habitat improvement rate is low	8-10	high	basin is mostly agricultural and channel lacks LW due to stream management	Pedersen et al., 2007; Kristensen et al., 2014

Location	Floodplain functions analyzed	Summary	Elapsed time[a] (years)	Magnitude[b]	Other stressors[c]	Reference
Órbigo River, Spain	enhanced habitat diversity and biodiversity	levee removal altered riparian vegetation towards natural state	2-4	high	flow regime is regulated by upstream reservoir to allow diversions for agriculture	Martínez-Fernández et al., 2017
Middle Ebro River, Spain	enhanced habitat diversity and biodiversity, hazard mitigation	levee removal and flood flow improved habitat diversity and biodiversity	3-20	medium	limited flow regulation by dams and diversions	Gumiero et al., 2013; González et al., 2017
Long Eau River, England	enhanced habitat diversity and biodiversity, hazard mitigation	levee removal and setback improved flood mitigation and improved floodplain biodiversity	17	low	highly regulated, dredged, banks are mown	Gumiero et al., 2013
Pite and Ume Rivers, Sweden	enhanced habitat diversity and biodiversity	levee removal on streams used for timber harvest increased floodplain connectivity and biodiversity with different results on habitat and biodiversity relationships between vegetation and aquatic organisms	1-20	medium	ongoing restoration efforts of the same type continue; the Ume is regulated	Lepori et al., 2005; Helfield et al., 2007; Helfield et al., 2012
Tummel River, Scotland	enhanced habitat diversity and biodiversity	floodplain landforms and vegetation biodiversity return to natural state after 50 years of levee abandonment	100	medium	flow regulated for hydropower	Parsons and Gilvear, 2002.
Allt Lorgy, Scotland	fluxes	levee alteration increased channel-floodplain interaction and bank erosion	5	low	restored section represents ~ 70% length of impacted length	Williams et al., 2019
Kushiro River, Japan	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat diversity and biodiversity	levee removal induced floodplain reconnection and riparian vegetation biodiversity	1	medium	channelized reaches upstream	Nakamura et al., 2014

Note. [a]Elapsed time indicates the years between levee alteration and data collection. [b]Magnitude indicates the degree of levee alteration. [c]Other stressors indicate other anthropogenic stressors that continue to operate in that river's basin

4.3.1 Limits of selected case studies from English-language journals.

The limits of the case studies are apparent in geographic extent (Figure 4.2), the degree of restoration compared to the degree of alteration by artificial levees and other stressors (Tables 4.1 and 4.2; chapters 2 and 3), and the limited elapsed time between restoration and data collection (Tables 4.1 and 4.2). Case studies represent the northern hemisphere mid-latitudes ranging between 27 °N (Kissimmee River, U.S.) and 66 °N (Pite and Ume River, Sweden). Conspicuously absent are documented examples of floodplain restoration via levee alteration from South America, Africa, and Australia.

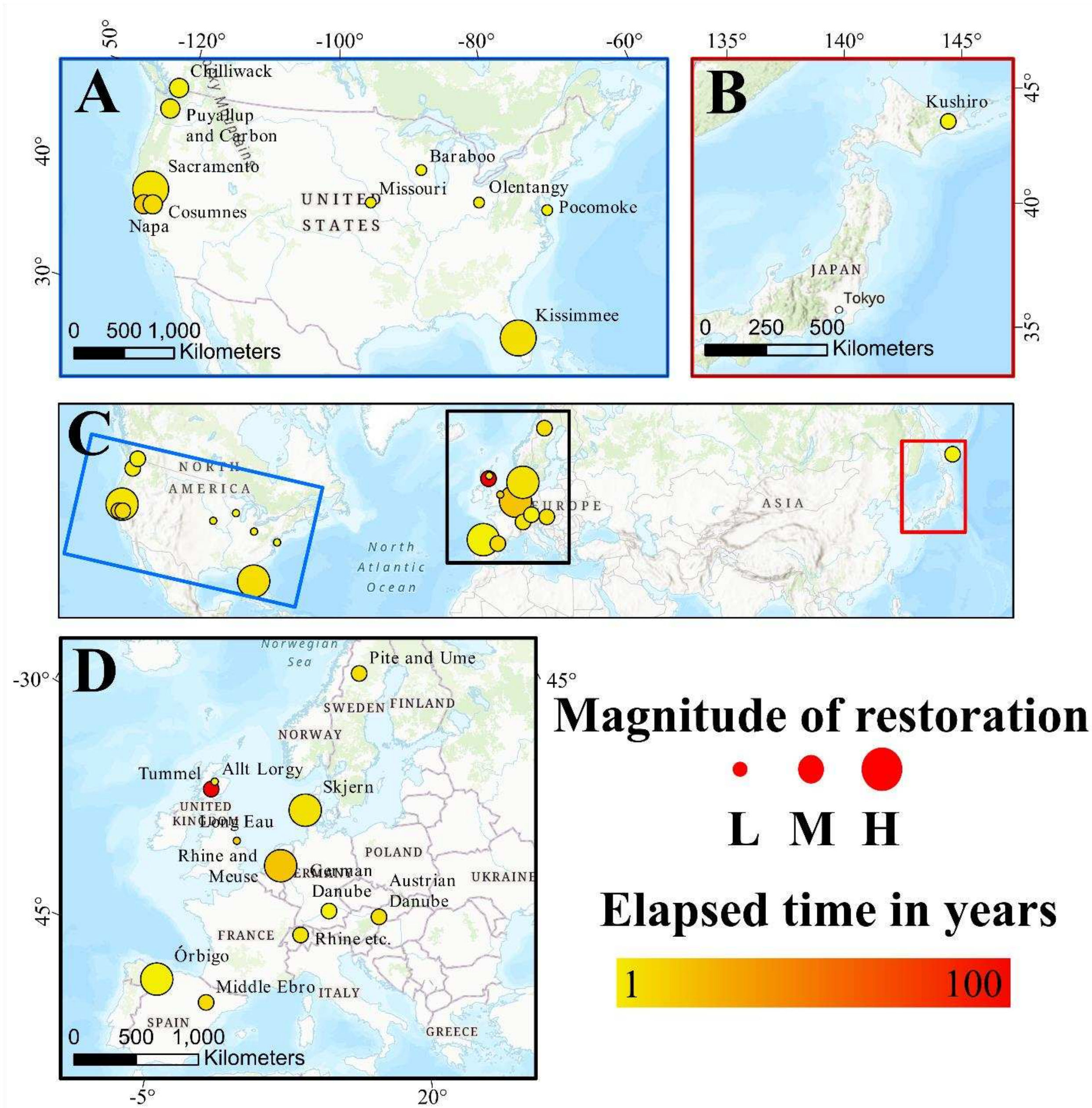


Figure 4.2. Selected case study location, magnitude of alteration, and elapsed time between alteration and data collection. (C) Case studies are concentrated in the northern hemisphere in (A) North America, (B) Japan, and (D) Europe. Low (L) magnitude impacts one or two levees and less than 10 river kilometers. Medium (M) magnitude restoration impacts between 10 and 100 river kilometers. High (H) magnitude restoration impacts hundreds of river kilometers.

The case studies summarized here indicate that the floodplain area impacted by restoration is very small compared to the total area impacted by artificial levees and other stressors. Using the United States as an example, the length of artificial levee alteration from case studies in Table 1 (the largest two examples are 161 km on the Sacramento River (Golet et al., 2008) and 70 km on the Kissimmee River (Koebel and Bousquin, 2014)) and impacts are far less than 1% of total artificial levee length estimates in the continental U.S. (~228,000 km, Knox et al., 2022a). In terms of floodplain reconnection to channels, restoration efforts only reconnect 1-2% of disconnected floodplain area (the two largest examples are ~54 km^2 on the Sacramento River (Golet et al., 2008) and ~80 km^2 on the Kissimmee River (Koebel and Bousquin, 2014)) in the U.S. (8,100 km^2, Knox et al., 2022b). I cannot make comparable quantitative assessments for Europe or Japan because the total extent of artificial levees there is unknown.

Artificial levees are just one (very influential) anthropogenic factor stressing floodplain functions, with almost every floodplain restoration project contending with outside stressors that impact restoration effectiveness (Tables 1 and 2). Effective restoration for target species must occur at the relevant habitat scale (Lepori et al., 2005). Nearby artificial levee (either upstream/downstream or setback) constraints on effectiveness of restoration on the Cosumnes and Pocomoke Rivers and rivers in Switzerland (Rohde et al., 2005) indicate the ability of anthropogenic features to adversely impact low and medium magnitude restoration. Negative impacts of flow regulation include the combination of minimal environmental flows with regulated flows (e.g., the Dutch portions of the Rhine and Meuse Rivers), a complete lack of environmental flows (e.g., the Spanish Órbigo River), and physical barriers to the movement of rare species (e.g., the German portion of the Danube). Many case studies recognized that effectiveness was limited by the degree of elapsed time between restoration and data collection, an issue that may become more important as the scale of restoration increases (Wohl, Lane, & Wilcox, 2015). Case studies with limited time between implementation and evaluation included those on the Olentangy, Napa, Sacramento, and Cosumnes Rivers in the US, as well as rivers in Germany, Denmark, Sweden, and Spain.

I do not necessarily see a consistent difference in the effectiveness of Old versus New World floodplain restoration projects. Site-specific details, such as magnitude of the restoration effort and constraints external to the restoration (e.g., flow regulation), appear to exert a greater influence than length of history of human alteration. The length of time that a site has been altered could certainly influence floodplain response to restoration. Plausible scenarios include such a long period of alteration that aquatic or riparian species have gone extinct or terrestrialisation of the floodplain has been so thorough that simply inundating the floodplain cannot restore lost functions. However, the existing literature on floodplain restoration projects is not yet sufficient to determine whether there is a consistent difference in floodplain response in Old versus New World settings as a result of the differences in history of alteration.

4.3.2 Reconnection and reconfiguration.

Bernhardt and Palmer (2011) make the distinction between restoration that involves reconfiguration (designing artificial channels or connections to floodplains) and reconnection (removing barriers to connection between natural channels and floodplains). They note the scant record of reconfiguration successes. Restoration is an experimental process situated within a wider social context (Gross, 2002) that involves value assignment and careful consideration of potential costs, benefits, and tradeoffs of possible outcomes. Several case studies indicate that reconfiguration strategies are one way in which practitioners select certain components and outcomes over others. By employing hardened intake structures to secondary channels, for example, the designers of the Chilliwack restoration assigned greater value to habitat stability in the near term at the cost of potential habitat decline in the long term due to decreased scouring flows (Ogston et al., 2014). In similar fashion, designers of the Skjern River restoration engineered channel floodplain connections to limit smolt predation at the cost of limited floodplain inundation and habitat development (Kristensen et al., 2014). In contrast, the reconnection of channels replaced by straightened canals seems to be very effective over a range of elapsed times (e.g., Kissimmee, Napa, and Kushiro Rivers). Compared to different restoration methods, floodplain-channel reconnection can be an effective method to improve fish biodiversity (Ramler and Keckeis, 2019). A secondary point is that some of these successful restorations represent studies with comparatively longer elapsed time, indicating that restoration outcomes should be evaluated over a longer time horizon.

4.3.3 Restoration scale dilemma.

One aspect of this dilemma is the inverse relationship of restoration scale and observation resolution. This issue is illustrated by Ahearn et al. (2006), who were able to record primary productivity and other data at high spatial and temporal resolution in the Cosumnes River floodplain because of the small floodplain area (0.36 km^2). This contrasts with a larger-scale restoration project (~10 km river corridor) where finer features such as boulders or large wood and subdivisions of aquatic habitat were undetectable using remote sensing (Konrad et al., 2008). The need to restore floodplains on a larger scale, discussed in section 3.1, poses challenges for the need to study floodplain complexity at finer spatial and temporal resolutions (Ahearn et al., 2006). The second component of the dilemma is that small-magnitude restoration, which allows for finer resolution analyses, is commonly impacted by other stressors outside the study area, as discussed above. The Cosumnes River floodplain experienced artificially low inundation rates because of the small floodplain size and fringing artificial levees. Similarly, the performance of levee setback sites for biodiversity across Switzerland was based mostly on proximity to nearly natural sites, which indicates the important role of nearby stressors (Rohde et al., 2005). The literature reviewed here at least provides a way forward given this dilemma, which will not be solved solely through more data, faster computing, better algorithms, and high-resolution remote sensing. Specifically, lessons learned and improved conceptual models developed at small magnitude restoration studies can facilitate more effective restoration at greater spatial scales.

Numerical modeling does have an important role to play in floodplain restoration. The increasing computational power of personal computers makes it more feasible to use 2D and 3D models to simulate the effects of floodplain restoration. Application of the models may still be limited by the need to provide spatially and temporally explicit input and validation data, as well as limitations on what the models simulate. Surface-water hydrologic and hydraulic models coupled with sediment transport models have advanced rapidly in recent years (e.g., Van Manh et al., 2015; Gilbert and Wilcox, 2020), for example, but models that effectively couple these physical processes with simulations of biogeochemical cycling or species or biotic community dynamics are limited (e.g., Theng et al., 2022).

Table 3 lists examples of numerical simulations used to evaluate the effects of levee alteration. Studies listed here include 1D hydraulic models (e.g., Remo et al., 2012), combined hydraulic-sedimentologic models (e.g., Jones et al., 2018), and combined hydraulic-plant growth models (e.g., Ahn et al., 2006). For at least the next few years, conceptual models of interactions among water, sediment, biogeochemical processes, and biotic communities are more likely to be used than numerical models. Conceptual models, like numerical models, can be most effective if they are based on knowledge of multiple, interacting variables and if they are informed by monitoring of restoration effects over timespans relevant to the process of interest (e.g., years to decades for vegetation community response; Shafroth et al., 2010; Kui et al., 2017).

Table 4.3. Selected case studies of simulations to restore floodplains by levee alteration

Location	Floodplain functions analyzed	Summary	Model type(s)	Reference
Middle Mississippi River, US	hazard mitigation	simulation scenarios indicate effectiveness of levee setbacks combined with buy backs given floodplain development and decreased flood stages from levee alteration	1D hydraulic (HEC-RAS & Hazus-MH)	Dierauer et al., 2012; Remo et al., 2012
Wisconsin, River, Wisconsin, US	enhanced habitat abundance and diversity, enhanced biomass and biodiversity, hazard mitigation	simulation and field data indicate levee setbacks provide some flood mitigation with little impact to vegetation biodiversity	1D hydraulic (HEC-RAS)	Gergel et al., 2002
Sangamon River, Illinois, US	fluxes, enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat abundance and diversity, enhanced biomass and biodiversity	simulation of different options for levee setbacks or gates on N and P storage and biodiversity involves tradeoffs in N/P storage and biodiversity	2D environmental (CASM)	Bartell et al., 2020
Illinois River, Illinois, US	fluxes, enhanced habitat abundance and diversity, enhanced biomass and biodiversity	different approaches involving levees and levee pumping are evaluated to restore floodplains	1D hydraulic; 2D plant growth; UNET	Sparks et al., 1990; Ahn et al., 2006
Illinois River, Illinois, US	enhanced habitat abundance and diversity, enhanced biomass and biodiversity, hazard mitigation	simulation of analysis and site selection of tradeoffs between economic costs of setbacks, flood risks, and biodiversity	1D hydraulic (HEC-RAS)	Guida et al., 2016; Remo et al., 2017
Iguacu River, Brazil	hazard mitigation	simulation of levee removal to increase flood storage in an urban watershed	pseudo 3D hydrologic- hydraulic (MODCEL)	Miguez et al., 2015
Upper Mid-western, US	enhanced spatial heterogeneity of hydrology and biogeochemistry	simulation of levee removal indicates improved nitrate-nitrogen processing in the floodplain	2D nitrogen biogeochemical numerical; 1D hydraulic (HEC-RAS)	Gergel et al., 2005
Lower White River, Washington, US	storage	simulation indicates improved sediment storage in reconnected floodplain after levee removal or setback	1D hydraulic (HEC-RAS); 2D sedimentologic (AdH)	Jones et al., 2018

Sacramento River, California, US	enhanced spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat abundance and diversity	simulation of levee setback impacts to floodplain reworking and connection to cutoffs indicates site-specific thresholds can be used to maximize habitat with minimal cost	2D river channel migration model	Larsen et al., 2006
American River, California, US	hazard mitigation	simulation of flood risk from future development and climate	2D hydrologic (HadCM2); 1D hydraulic (HEC-RAS); economic model for climate and urban scenarios	Zhu et al., 2007
White River, California, US	enhanced habitat abundance and diversity, enhanced biomass and biodiversity, hazard mitigation	simulation of levee setbacks reduces flood heights and improved salmonid habitat	2D hydraulic (RiverFlo-2D); 3D Stream Tube model; 3D fish foraging and bioenergetics model	Black et al., 2016
San Joaquin River, California, US	enhanced habitat abundance and diversity, enhanced biomass and biodiversity	simulations of impacts of levee setbacks, bypasses, and climate projections indicate that successful restoration must include both floodplain reconnection and augmented reservoir releases	2D climate (B1PCM & A2GFDL); 2D hydro-ecologic (HEC-EFM); 1D hydraulic (HEC-RAS)	Matella and Merenlender, 2015

The use of Decision Support Systems (DSS) that can integrate numerical and conceptual versions of parameter response to changes in floodplain-channel connectivity may prove to be particularly useful in the context of floodplain restoration. An example for environmental flows is the DRIFT DSS software developed for integrated flow assessments by King, Brown, and others (King et al., 2004; Brown et al., 2006, 2013), which has now been applied to multiple watersheds in southern Africa. Another example evaluates fish and vegetation habitat availability given different flow scenarios (Passero, 2020). This approach, when combined with the societal designation of acceptable levels of alteration in a floodplain (King and Brown, 2018), explicitly provides a mechanism for including stakeholder perceptions and values. Given the societal context in which river and floodplain restoration occurs, including socioeconomic considerations such as stakeholder perceptions is likely to be critical to efforts to expand the magnitude and spatial extent of floodplain reconnection via modification of artificial levees.

4.3.4 Unique place-based challenges.

This is the simplest way to explain why certain restoration measures are effective at one location but not at another. The nexus of river and floodplain restoration in altering functions and form experiences the same tension that arises in fluvial geomorphology between the need to identify universal physical processes amidst site-based contingency and characteristics (Wohl, 2014). The restoration of floodplain connection on high-gradient bedrock rivers, which are the exception to the more common restoration on lowland alluvial rivers, illustrates this dilemma. Several restoration projects (Puyallup, Chilliwack, and the Pite/Ume Rivers) are distinct for their location along high-gradient bedrock rivers. Two of these experienced almost immediate positive restoration effects over a period of 6 years at the Puyallup River (Konrad et al., 2008) and 1-4 years at the Chilliwack River (Ogston et al., 2014). Levee removal along the Pite and Ume Rivers (Lepori et al., 2005; Helfield et al. 2007, 2012), which had differing levels of success with respect to desired outcomes, illustrates why expectations for restoration projects along similar types of rivers may be disappointed. Riparian vegetation responded quickly to levee removal along the Ume River but slowly along the Pite River (Helfield et al., 2007, 2012). The different vegetation responses are attributed to (i) different substrate at the restoration sites because of differences in glacial history, (ii) the limited elapsed time, which was up to 20 years, (iii) different pre-existing vegetation patterns on the Pite River secondary channels and the Ume River primary channels, and (iv) incorrect restoration scale compared to the habitat scale of the target organism (Helfield et al., 2012). Cursory similarities (i.e., high-gradient bedrock streams in northern Sweden, similar artificial levee type and restoration method) fail to explain the differences in riparian vegetation response. Notably, none of the case studies on high-gradient bedrock rivers mention other stressors as a reason for restoration ineffectiveness despite flow regulation for hydropower above the Puyallup site and many kilometers of fringing artificial levees near the Pite/Ume Rivers restoration sites. These case studies illustrate the potential for successful restoration on high-gradient bedrock rivers, but also the need to appreciate site-specific characteristics.

4.4 Conclusion

My intent is to explore floodplain functions and how they are impacted by artificial levees. I define five floodplain functions (material fluxes, enhancement of spatial heterogeneity of hydrology and biogeochemistry, enhanced habitat abundance and diversity, enhanced biomass and biodiversity, and hazard mitigation) and selected floodplain restoration case studies that involve alteration to artificial levees. Floodplain functions are highly integrative and based primarily on lateral connectivity between the channel and the floodplain, which is why artificial levees are so harmful to floodplain functions. Case studies are concentrated in North America and Europe on lowland alluvial rivers and generally include data collection within 30 years of restoration. Artificial reconfiguration of floodplain connectivity achieved limited success. Reconnection of channels and floodplains seems more likely than reconfiguration to set floodplains on a trajectory to more fully restore floodplain functions. Case studies highlight the dichotomy between restoration site scale and study resolution, although future case studies will continue to inform conceptual models of restoration and it is critical to continue multi-decadal monitoring of the effects of floodplain restoration. Restoration effectiveness varied by location and highlights the need to apply restoration techniques that are relevant to a specific location. Some of these considerations include the impacts of other stressors (e.g., flow regulation) on connectivity after barrier removal, the site’s geologic history, and the scale of restoration required by the target species.

Chapter 5 Conclusion

Fully functioning floodplains are a vital component of river ecosystems and of great value to humans. It is very important to understand how anthropogenic features influence floodplain functions so that humans can accurately understand the costs and benefits associated with different decisions involving river corridors and floodplains. As such, it is important to investigate how artificial levees, a widely used anthropogenic feature, impact floodplain functions. In this dissertation, I estimated artificial levee locations in the continental United States. Over 182,000 km of potential levees were detected in the CONUS, with a high concentration along smaller streams in the Mississippi and Missouri Basins. These results suggest that the NLD includes about 20% of total artificial levee lengths. Given the long history of artificial levee construction in the CONUS, including construction by landowners and municipal levee boards (Hudson et al., 2008), I contend that the combination of potential levees and those in the NLD is an underestimation of total artificial levee length in the CONUS. Surprisingly, the best performing model used for detecting artificial levees was one that did not use topographic data or the shape of levees, but instead used the land cover, basin, and distance from stream variables, indicating the value of land cover and spatial patterns of anthropogenic features to modeling. This modeling methodology avoids detection of natural levees for two reasons. First, artificial levees occupy a very specific land cover - distance from stream variable space which arises from patterns of artificial levee construction. Second, natural levees develop on rivers with high suspended sediment concentrations and overbank flooding, and many of these locations now have artificial levees installed. I used the locations of potential levees and NLD levees to modify the 1 arc-second (30 m resolution) NED digital elevation model by deleting levees from it. I then used the hydrogeomorphic floodplain algorithm, GFPLAIN, to compare floodplain extent from topography with and without levees. Surprisingly (again), the floodplain extent on the unmodified topography was slightly larger than floodplain extent with levees deleted. This previously unknown dimension of artificial levee impacts to floodplains illustrates that we have massively underestimated the ecological and hydrological damage of levees. The areas of agreement, anthropogenically connected, and anthropogenically disconnected floodplains were 957,111 km^2, 8,911 km^2, and 8,100 km^2, respectively. Over 30% of the CONUS floodplain was either cultivated or developed land cover, indicating one of the reasons for the trend noted by White (2000) in the last century.

I identified five floodplain functions: (1) fluxes of water, solutes, and particulate materials; (2) enhanced spatial heterogeneity of hydrology and biogeochemistry; (3) enhanced habitat abundance and diversity; (4) enhanced biomass and biodiversity; and (5) hazard mitigation. Restoration case studies involving artificial levee alteration indicated the heavy concentration in North America and Europe, with an elapsed time of less than 30 years between restoration and evaluation. In the United States, restoration targeted much less than 1% of river kilometers impacted by artificial levees and 1-2% of disconnected floodplains. Most case studies were impacted by stressors outside the study site such as flow regulation, which emphasizes that the degree of impacts to floodplain functions are likely much larger than impacts just from artificial levees. Researchers struggled to balance the scale of restoration with the study resolution in time and space and discovered that different restoration techniques performed differently in varied locations for diverse reasons.

The combined effects of artificial levees, flow regulation, channelization, and land drainage on floodplain functions are massive. In the CONUS alone, the impacts to floodplains documented in this dissertation are just one ecological impact among many that converted more than 50% of the estimated 900,000 km^2 of wetlands present historically to other uses since 1780 (Vileisis, 1999). The 17,000 km^2 cumulative alteration of CONUS floodplains identified here is comparable to other estimates of floodplain development in the CONUS ranging between 14,000 and 23,000 km^2 (Vileisis, 1999). Placing this dissertation in the context of even more extensive alterations to floodplains and floodplain functions creates an even greater sense of urgency to understand and promote floodplain restoration.

These analyses contain several limitations. I was not able to ground-truth most of the potential levees. This is partially a limitation inherent to artificial levees given their shape (created to mimic a natural feature and two-dimensional at certain scales) and their history (constructed for hundreds of years and maintained with varying levels of rigor). However, data collected at the location of potential levees by researchers or citizen scientists could assist validation. These types of data include: local historical sources from public libraries or historical societies; accurate measurements of potential levee dimensions; and sediment cores. Another limitation was the degree of filtering of the NLD as training data, which resulted in the inclusion of dams that are classified as artificial levees in the NLD. The use of the 100-year hydrogeomorphic floodplain as the study extent mitigated against the inclusion of coastal levees in the training data. The deletion of artificial levees from topography is limited in that this method does not necessarily result in the height that natural levees would have if the location was undisturbed. I was limited by only searching English-language journals and the subsequent focus on North America and Europe for restoration. I avoided defining floodplain restoration beyond the specific situation of artificial levee adjustment. The limited elapsed time between floodplain restoration and analyses is agnostic on whether floodplain functions were actually restored, or if the stage was set for the restoration of functions.

Some exploration of the impacts of data resolution and aggregation on the analyses and results is called for because floodplain functions operate across multiple spatial and temporal scales (e.g., Amoros and Bornette, 2002; Tockner and Stanford, 2002). This dissertation's incorporation of only the spatial component of artificial levees (about half of the entries in the NLD list a construction date) has several implications. This snapshot approach of levee identification, which combines the current location and land cover associations of artificial levees constructed sometime in the past, weakens the land cover – levee patterns on which the CONUS model (i.e., model 12) is based. This approach also limits the ability to estimate potential levee construction date and any subsequent ability to use that date to limit inclusion of levees to, for example, more recent construction dates, for floodplain extent analyses. It is still unclear whether the relatively poor ability of models to detect artificial levees using levee shape is more related to inappropriate DEM resolution or the inherent limits of using levee shape for this purpose given the similarity of levee shape to other natural and anthropogenic landforms. The trend of this study to aggregate results at the HUC2 basin and CONUS level fills an important gap in our understanding of the impacts of artificial levees on floodplains at a continental scale. However, it's my hope that the public availability of the floodplain extent results at a 30 m resolution can be used to improve understanding of levee impacts to fluvial dynamics at the scales of both the floodplain and individual waterbodies (e.g., Amoros and Bornette, 2002).

These analyses also provide other opportunities for future research. Artificial levee identification of this type should be attempted in Europe. Future work on identification could also include testing additional variables from different data types (e.g., spectral data, road data) and distance from streams of order 7-10. Ground-truthing of potential levees could continue digitally, making use of the smaller geographic extent of potential levees by testing variables derived from higher resolution topographic data, as well as on the ground by researchers and citizen scientists.

Floodplain calibration could take into account different methods that weigh stream order contribution differently and include calibration at a finer spatial resolution. The method by which artificial levees are deleted from the topography using focal means could be replaced by a process of modeling natural levee height. I could have compared three floodplain extents (as-is topography with levees, topography with NLD levees deleted, and topography with potential levees deleted) to compare the contribution of NLD and potential levees to disconnection separately. The ability to compare potential levee height with modelled natural levee height provides another data point to validate potential levees.

To improve the global coverage of my restoration review in the future, I could collaborate with researchers fluent in the languages used in regional journals. Future work could also determine how to remotely identify floodplain functions or at least their proxies.

In summary, this is an exciting area of research with much more work to do. This dissertation highlights the importance of understanding artificial levee impacts on lateral connectivity and floodplain functions and identifies some pathways for future research.

References

Ahearn, D. S., Viers, J. H., Mount, J. F., & Dahlgren, R. A. (2006). Priming the productivity pump: flood pulse driven trends in suspended algal biomass distribution across a restored floodplain. Freshwater Biology, 51(8), 1417-1433.

Ahn, C., Johnston, D. M., Sparks, R. E., & White, D. C. (2006). Analysis of naturalization alternatives for the recovery of moist-soil plants in the floodplain of the Illinois River. In Living Rivers: Trends and Challenges in Science and Management (pp. 217-228). Springer, Dordrecht.

Allen, J. R. (1965). A review of the origin and characteristics of recent alluvial sediments. *Sedimentology*, *5*(2), 89-191.

American Society of Civil Engineers (2017). 2017 Infrastructure Report Card: Levees, Reston, Va. [Available at https://www.infrastructurereportcard.org/cat-item/levees/.]

Amoros, C., & Bornette, G. (1999). Antagonistic and cumulative effects of connectivity: a predictive model based on aquatic vegetation in riverine wetlands. Arch. Hydrobiol, 11, 311-327.

Amoros, C., & Bornette, G. (2002). Connectivity and biocomplexity in waterbodies of riverine floodplains. Freshwater biology, 47(4), 761-776.

Annis, A., Nardi, F., Morrison, R. R., & Castelli, F. (2019). Investigating hydrogeomorphic floodplain mapping performance with varying DTM resolution and stream order. Hydrological Sciences Journal, 64(5), 525-538.

Annis, A., Karpack, M., Morrison, R. R., & Nardi, F. (2022). On the influence of River Basin Morphology and Climate on Hydrogeomorphic Floodplain delineations. Advances in Water Resources, 104078.

Ansari, A., & Golabi, M. H. (2019). Prediction of spatial land use changes based on LCM in a GIS environment for Desert Wetlands–A case study: Meighan Wetland, Iran. International soil and water conservation research, 7(1), 64-70.

Appling, A. P., Bernhardt, E. S., & Stanford, J. A. (2014). Floodplain biogeochemical mosaics: a multidimensional view of alluvial soils. Journal of Geophysical Research: Biogeosciences, 119(8), 1538-1553.

Argiroff, W. A., Zak, D. R., Lanser, C. M., & Wiley, M. J. (2017). Microbial community functional potential and composition are shaped by hydrologic connectivity in riverine floodplain soils. Microbial ecology, 73(3), 630-644.

Arscott, D. B., Tockner, K., van der Nat, D., & Ward, J. V. (2002). Aquatic Habitat Dynamics along a Braided Alpine River Ecosystem (Tagliamento River, Northeast Italy). Ecosystems, 5(8), 802–814. http://www.jstor.org/stable/3658883

Atkinson, C. L., Allen, D. C., Davis, L., & Nickerson, Z. L. (2018). Incorporating ecogeomorphic feedbacks to better understand resiliency in streams: a review and directions forward. Geomorphology, 305, 123-140.

Baldwin, D. S., & Mitchell, A. M. (2000). The effects of drying and re‐flooding on the sediment and soil nutrient dynamics of lowland river–floodplain systems: a synthesis. Regulated Rivers: Research & Management: An International Journal Devoted to River Research and Management, 16(5), 457-467.

Barry, J.M. (2007). Rising Tide: The Great Mississippi Flood of 1927 and How it Changed America, Simon and Schuster, 528 p.

Bartell, S. M., Woodard, C. T., Theiling, C. H., & Dahl, T. A. (2020). Development and application of the CASM-SL to support nutrient management in potential Sangamon River levee setbacks.

Bayley, P. B. (1995). Understanding large river: floodplain ecosystems. BioScience, 45(3), 153-158.

Bechtol, V., & Laurian, L. (2005). Restoring straightened rivers for sustainable flood mitigation. Disaster prevention and management: an international journal.

Bellmore, J. R., & Baxter, C. V. (2014). Effects of geomorphic process domains on river ecosystems: a comparison of floodplain and confined valley segments. River Research and Applications, 30(5), 617-630.

Berhane, T. M., Lane, C. R., Wu, Q., Autrey, B. C., Anenkhonov, O. A., Chepinoga, V. V., & Liu, H. (2018). Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory. Remote sensing, *10*(4), 580.

Bernhardt, E. S., Blaszczak, J. R., Ficken, C. D., Fork, M. L., Kaiser, K. E., & Seybold, E. C. (2017). Control points in ecosystems: moving beyond the hot spot hot moment concept. Ecosystems, 20(4), 665-682.

Bernhardt, E. S., & Palmer, M. A. (2011). River restoration: the fuzzy logic of repairing reaches to reverse catchment scale degradation. Ecological applications, 21(6), 1926-1931.

Black, R. W., Czuba, C. R., Magirl, C. S., McCarthy, S., Berge, H., & Comanor, K. (2016). Effect of a levee setback on aquatic resources using two-dimensional flow and bioenergetics models (No. 2016-5025). US Geological Survey.

Blann, K. L., Anderson, J. L., Sands, G. R., & Vondracek, B. (2009). Effects of agricultural drainage on aquatic ecosystems: a review. Critical reviews in environmental science and technology, 39(11), 909-1001.

Blanton, P. & Marcus, W.A. (2009). Railroads, roads and lateral disconnection in the river landscapes of the continental United States, Geomorphology, 112, 212-227.

Blaschak, T. S. (2012). Examining geomorphic effects of flow diversions on low-gradient mountain streams in the Routt National Forest, Colorado (Doctoral dissertation, Colorado State University).

Blessing, R., Sebastian, A., & Brody, S. D. (2017). Flood risk delineation in the United States: how much loss are we capturing? Natural Hazards Review, 18(3), 04017002.

Bouska, K. L., Houser, J. N., De Jager, N. R., Van Appledorn, M., & Rogala, J. T. (2019). Applying concepts of general resilience to large river ecosystems: A case study from the Upper Mississippi and Illinois rivers. Ecological Indicators, 101, 1094-1110.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Brierley, G. J., Brooks, A. P., Fryirs, K., & Taylor, M. P. (2005). Did humid-temperate rivers in the Old and New Worlds respond differently to clearance of riparian vegetation and removal of woody debris?. Progress in Physical Geography, 29(1), 27-49.

Brierley, G. J., Ferguson, R. J., & Woolfe, K. J. (1997). What is a fluvial levee?. Sedimentary Geology, 114(1-4), 1-9.

Briggs, M. A., Lautz, L. K., Hare, D. K., & González-Pinzón, R. (2013). Relating hyporheic fluxes, residence times, and redox-sensitive biogeochemical processes upstream of beaver dams. *Freshwater Science*, *32*(2), 622-641.

Brinson, M. M., Lugo, A. E., & Brown, S. (1981). Primary productivity, decomposition and consumer activity in freshwater wetlands. Annual Review of Ecology and Systematics, 12(1), 123-161.

Brooks, G. R. (2003). Holocene lateral channel migration and incision of the Red River, Manitoba, Canada. Geomorphology, 54(3-4), 197-215.

Brooks, A. P., & Brierley, G. J. (2002). Mediated equilibrium: the influence of riparian vegetation and wood on the long-term evolution and behaviour of a near-pristine river. Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 27(4), 343-367.

Brown, A. G., Tooth, S., Bullard, J. E., Thomas, D. S., Chiverrell, R. C., Plater, A. J., ... & Wainwright, J. (2017). The geomorphology of the Anthropocene: emergence, status and implications. Earth Surface Processes and Landforms, 42(1), 71-90.

Brown, C.A., Pemberton, C., Birkhead, A, Bok, A., Boucher, C., Dollar, E., Harding, W., et al. (2006). In support of water-resource planning – highlighting key management issues using DRIFT: A case study. Water SA, 32, 181-192.

Brown, C.A., Joubert, A.R., Beuster, J., Greyling, A., & King, J.M. (2013). DRIFT: DSS Software Development for Integrated Flow Assessments. Report to the Water Research Commission, WRC Report No. 1873/1/13, South Africa. http://www.wrc.org.za/wp-content/uploads/mdocs/TT%20575-13.pdf

Bullinger-Weber, G., Le Bayon, R. C., Thébault, A., Schlaepfer, R., & Guenat, C. (2014). Carbon storage and soil organic matter stabilisation in near-natural, restored and embanked Swiss floodplains. Geoderma, 228, 122-131.

Bunn, S. E., Abal, E. G., Smith, M. J., Choy, S. C., Fellows, C. S., Harch, B. D., ... & Sheldon, F. (2010). Integration of science and monitoring of river ecosystem health to guide investments in catchment protection and rehabilitation. Freshwater Biology, 55, 223-240.

Burt, T. P. (1996). The hydrological role of floodplains within the drainage basin system. Buffer zones: their processes and potential in water protection, 21-32.

Buto, S. G., & Anderson, R. D. (2020). NHDPlus High Resolution (NHDPlus HR)---A hydrography framework for the Nation (No. 2020-3033). US Geological Survey.

Butzer, K. W. (1976). Early hydraulic civilization in Egypt: a study in cultural ecology. Prehistoric archeology and ecology (USA).

Cannon, C. M. (2015). Landforms along the Lower Columbia River and the influence of humans (Doctoral dissertation, Portland State University).

Caskey, S. T. (2013). Downstream effects of diversion dams on riparian vegetation communities in the Routt National Forest, Colorado (MS thesis, Colorado State University).

Caskey, S. T., Blaschak, T. S., Wohl, E., Schnackenberg, E., Merritt, D. M., & Dwire, K. A. (2015). Downstream effects of stream flow diversion on channel characteristics and riparian vegetation in the Colorado Rocky Mountains, USA. Earth Surface Processes and Landforms, 40(5), 586-598.

Castro, J. M., & Thorne, C. R. (2019). The stream evolution triangle: Integrating geology, hydrology, and biology. River Research and Applications, 35(4), 315-326.

Chaparro, G., O'Farrell, I., & Hein, T. (2019). Multi-scale analysis of functional plankton diversity in floodplain wetlands: Effects of river regulation. Science of the Total Environment, 667, 338-347.

Cho, E., Jacobs, J. M., Jia, X., & Kraatz, S. (2019). Identifying subsurface drainage using satellite Big Data and machine learning via Google Earth Engine. Water Resources Research, 55(10), 8028-8045.

Choi, C., Kim, J., Han, H., Han, D., & Kim, H. S. (2020). Development of water level prediction models using machine learning in wetlands: A case study of Upo wetland in South Korea. Water, 12(1), 93.

Clawson, R. G., Lockaby, B. G., & Rummer, B. (2001). Changes in production and nutrient cycling across a wetness gradient within a floodplain forest. Ecosystems, 4(2), 126-138.

Cole, J. J., Findlay, S., & Pace, M. L. (1988). Bacterial production in fresh and saltwater ecosystems: a cross-system overview. Marine ecology progress series. Oldendorf, 43(1), 1-10.

Covino, T. (2017). Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. Geomorphology, 277, 133-144.

Criss, R. E. and Shock, E. L. (2001). Flood enhancement through flood control. Geology, 29(10), 875-878.

Cuffney, T. F. (1988). Input, movement and exchange of organic matter within a subtropical coastal black water river-flood plain system. Freshwater Biology, 19(3), 305-320.

Czech, W., Radecki-Pawlik, A., Wyżga, B., & Hajdukiewicz, H. (2016). Modelling the flooding capacity of a Polish Carpathian river: a comparison of constrained and free channel conditions. Geomorphology, 272, 32-42.

Czuba, C., Williams, B., Westman, J., & LeClaire, K. (2015). An Assessment of Two Methods for Identifying Undocumented Levees Using Remotely Sensed Data. US Department of the Interior, US Geological Survey.

Darke, A. K., Walbridge, M. R., & Lockaby, B. G. (1996). Changes in Al and Fe crystallinity and P sorption capacity in a flood-plain forest soil subjected to artificially manipulated flooding regimes in field mesocosms. Wetlands Ecology and Management, 4(4), 235-244.

Deines, J. M., Kendall, A. D., & Hyndman, D. W. (2017). Annual irrigation dynamics in the US Northern High Plains derived from Landsat satellite data. Geophysical Research Letters, 44(18), 9350-9360.

Dennis, I. A., Coulthard, T. J., Brewer, P., & Macklin, M. G. (2009). The role of floodplains in attenuating contaminated sediment fluxes in formerly mined drainage basins. Earth Surface Processes and Landforms, 34(3), 453-466.

Dierauer, J., Pinter, N., & Remo, J. W. (2012). Evaluation of levee setbacks for flood-loss reduction, Middle Mississippi River, USA. Journal of Hydrology, 450, 1-8.

Diggory, Z. E., & Parker, V. T. (2011). Seed supply and revegetation dynamics at restored tidal marshes, Napa River, California. Restoration ecology, 19(101), 121-130.

Dingman, S.L., & Platt, R.H. (1977). Floodplain zoning: implications of hydrologic and legal uncertainty. Water Resources Research, 14, 519-523.

Dunne, T., & Aalto, R. E. (2013). Large river floodplains (Vol. 9, pp. 645-678).

Dunne, T., Mertes, L.A.K., Meade, R.H., Richey, J.E., & Forsberg, B.R. (1998). Exchanges of sediment between the flood plain and the channel of the Amazon River in Brazil. Geological Society of America Bulletin, 110, 450-467.

Dwivedi, D., Arora, B., Steefel, C. I., Dafflon, B., & Versteeg, R. (2018). Hot spots and hot moments of nitrogen in a riparian corridor. Water Resources Research, 54(1), 205-222.

Esri Inc. (2021). *ArcGIS Pro* (Version 2.8.3). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview.

Fairfax, E., & Whittle, A. (2020). Smokey the Beaver: beaver-dammed riparian corridors stay green during wildfire throughout the western United States. Ecological Applications, 30(8), e02225.

Fitzgerald, R. W., & Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. Remote sensing of Environment, 47(3), 362-368.

Florsheim, J. L., & Mount, J. F. (2002). Restoration of floodplain topography by sand-splay complex formation in response to intentional levee breaches, Lower Cosumnes River, California. Geomorphology, 44(1-2), 67-94.

Florsheim, J. L., & Mount, J. F. (2003). Changes in lowland floodplain sedimentation processes: pre-disturbance to post-rehabilitation, Cosumnes River, CA. Geomorphology, 56(3-4), 305-323.

Foody, G. M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. Remote Sensing of Environment, 239, 111630.

Frings, R. M., Berbee, B. M., Erkens, G., Kleinhans, M. G., & Gouw, M. J. (2009). Human-induced changes in bed shear stress and bed grain size in the River Waal (The Netherlands) during the past 900 years. Earth Surface Processes and Landforms, 34(4), 503-514.

Galat, D. L., Fredrickson, L. H., Humburg, D. D., Bataille, K. J., Bodie, J. R., Dohrenwend, J., ... & Semlitsch, R. D. (1998). Flooding to restore connectivity of regulated, large-river wetlands: natural and controlled flooding as complementary processes along the lower Missouri River. BioScience, 48(9), 721-733.

Galloway Jr, G. E. (1995). Learning from the Mississippi flood of 1993: Impacts, management issues, and areas for research. In US-Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods.

Gergel, S. E., Carpenter, S. R., & Stanley, E. H. (2005). Do dams and levees impact nitrogen cycling? Simulating the effects of flood alterations on floodplain denitrification. Global Change Biology, 11(8), 1352-1367.

Gergel, S. E., Dixon, M. D., & Turner, M. G. (2002). Consequences of human-altered floods: levees, floods, and floodplain forests along the Wisconsin River. Ecological applications, 12(6), 1755-1770.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., & Tyler, D. (2002). The national elevation dataset. Photogrammetric engineering and remote sensing, 68(1), 5-32.

Gesch, D. B., Oimoen, M. J., & Evans, G. A. (2014). Accuracy assessment of the US Geological Survey National Elevation Dataset, and comparison with other large-area elevation datasets: SRTM and ASTER (Vol. 1008). US Department of the Interior, US Geological Survey.

Gilbert, J.T., and Wilcox, A.C. (2020). Sediment routing and floodplain exchange (SeRFE): a spatially explicit model of sediment balance and connectivity through river networks. Journal of Advances in Modeling Earth Systems, 12, e2020MS002048.

Golden, H. E., Sander, H. A., Lane, C. R., Zhao, C., Price, K., D'Amico, E., & Christensen, J. R. (2016). Relative effects of geographically isolated wetlands on streamflow: a watershed-scale analysis. Ecohydrology, 9(1), 21-38.

Golet, G. H., Gardali, T., Howell, C. A., Hunt, J., Luster, R. A., Rainey, W., ... & Williams, N. (2008). Wildlife response to riparian restoration on the Sacramento River. San Francisco Estuary and Watershed Science, 6(2).

González, E., Masip, A., Tabacchi, E., & Poulin, M. (2017). Strategies to restore floodplain vegetation after abandonment of human activities. Restoration Ecology, 25(1), 82-91.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote sensing of Environment, 202, 18-27.

Gordon, B. A., Dorothy, O., & Lenhart, C. F. (2020). Nutrient retention in ecologically functional floodplains: A review. Water, 12(10), 2762.

Graf, W. L. (1999). Dam nation: A geographic census of American dams and their large-scale hydrologic impacts. Water resources research, 35(4), 1305-1311.doi: 10.1029/1999WR900016.

Graf, W. L. (2001). Dam age control: restoring the physical integrity of America's rivers. Annals of the Association of American Geographers, 91(1), 1-27.

Gren, M., Groth, K. H., & Sylvén, M. (1995). Economic values of Danube floodplains. Journal of Environmental Management, 45(4), 333-345.

Grill, C., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019). Mapping the world's free-flowing rivers. Nature, 569, 215-221.

Gross, M. (2002). New natures and old science: hands-on practice and academic research in ecological restoration. Science & Technology Studies, 15(2), 17-35.

Guida, R. J., Remo, J. W., & Secchi, S. (2016). Tradeoffs of strategically reconnecting rivers to their floodplains: The case of the Lower Illinois River (USA). Science of the total environment, 572, 43-55.

Guilliams, D. (1998). The Big Ditch: The Wichita-Valley Center Flood Control Project. Fairmount Folio: Journal of History, 2.

Gumiero, B., Mant, J., Hein, T., Elso, J., & Boz, B. (2013). Linking the restoration of rivers and riparian zones/wetlands in Europe: sharing knowledge through case studies. Ecological Engineering, 56, 36-50.

Hanberry, B. B., Kabrick, J. M., & He, H. S. (2015). Potential tree and soil carbon storage in a major historical floodplain forest with disrupted ecological function. Perspectives in Plant Ecology, Evolution and Systematics, 17(1), 17-23.

Harvey, J., & Gooseff, M. (2015). River corridor science: Hydrologic exchange and ecological consequences from bedforms to basins. Water Resources Research, 51(9), 6893-6922.

Hauer, F. R., Locke, H., Dreitz, V. J., Hebblewhite, M., Lowe, W. H., Muhlfeld, C. C., ... & Rood, S. B. (2016). Gravel-bed river floodplains are the ecological nexus of glaciated mountain landscapes. Science Advances, 2(6), e1600026.

Hecht, J.S., Lacombe, G., Arias, M.E., Dan, T.D., & Piman, T. (2019). Hydropower dams of the Mekong River basin: a review of their hydrological impacts. Journal of Hydrology, *568*, 285-300.

Hedin, L. O., Von Fischer, J. C., Ostrom, N. E., Kennedy, B. P., Brown, M. G., & Robertson, G. P. (1998). Thermodynamic constraints on nitrogentransformations and other biogeochemicalprocesses at soil–stream interfaces. Ecology, 79(2), 684-703.

Heiler, G., Hein, T., Schiemer, F., & Bornette, G. (1995). Hydrological connectivity and flood pulses as the central aspects for the integrity of a river-floodplain system. Regulated Rivers: Research & Management, 11(3-4), 351-361.

Heine, R. A., & Pinter, N. (2012). Levee effects upon flood levels: an empirical assessment. Hydrological Processes, 26(21), 3225-3240.

Helfield, J. M., Capon, S. J., Nilsson, C., Jansson, R., & Palm, D. (2007). Restoration of rivers used for timber floating: effects on riparian plant diversity. Ecological Applications, 17(3), 840-851.

Helfield, J. M., Engström, J., Michel, J. T., Nilsson, C., & Jansson, R. (2012). Effects of river restoration on riparian biodiversity in secondary channels of the Pite River, Sweden. Environmental management, 49(1), 130-141.

Henry, L., & Wickham, H. (2020). purrr: Functional Programming Tools. R package version 0.3.4.

Hirschboeck, K. K. (1991). Climate and floods. US Geol. Surv. Water-Supply Pap, 2375, 67-88.

Hoagland, B., Schmidt, C., Russo, T. A., Adams, R., & Kaye, J. (2019). Controls on nitrogen transformation rates on restored floodplains along the Cosumnes River, California. Science of The Total Environment, 649, 979-994.

Holling, C. S. (1973). Resilience and stability of ecological systems. Annual review of ecology and systematics, 4(1), 1-23.

Hood, G. A., & Bayley, S. E. (2008). Beaver (Castor canadensis) mitigate the effects of climate on the area of open water in boreal wetlands in western Canada. Biological Conservation, 141(2), 556-567.

Horritt, M. S., & Bates, P. D. (2001). Effects of spatial resolution on a raster based model of flood flow. Journal of Hydrology, 253(1-4), 239-249.

House, W. A. (2003). Geochemical cycling of phosphorus in rivers. Applied Geochemistry, 18(5), 739-748.

Hudson, P. F., & Colditz, R. R. (2003). Flood delineation in a large and complex alluvial valley, lower Panuco basin, Mexico. Journal of Hydrology, 280(1-4), 229-245.

Hudson, P. F., Heitmuller, F. T., & Leitch, M. B. (2012). Hydrologic connectivity of oxbow lakes along the lower Guadalupe River, Texas: the influence of geomorphic and climatic controls on the “flood pulse concept”. Journal of Hydrology, 414, 174-183.

Hudson, P. F., Middelkoop, H., & Stouthamer, E. (2008). Flood management along the Lower Mississippi and Rhine Rivers (The Netherlands) and the continuum of geomorphic adjustment. Geomorphology, 101(1-2), 209-236.

Humphries, P., Keckeis, H., & Finlayson, B. (2014). The river wave concept: integrating river ecosystem models. BioScience, 64(10), 870-882.

Jackson, J. K., Huryn, A. D., Strayer, D. L., Courtemanch, D.L., & Sweeney, B. W. (2005). Atlantic coast rivers of the northeastern United States. In Rivers of North America (pp. 20-71). Academic Press.

Jacobson, R. B., Lindner, G., & Bitner, C. (2015). The role of floodplain restoration in mitigating flood risk, Lower Missouri River, USA. In Geomorphic approaches to integrated floodplain management of lowland fluvial systems in North America and Europe (pp. 203-243). Springer, New York, NY.

Jafarzadegan, K., & Merwade, V. (2017). A DEM-based approach for large-scale floodplain mapping in ungauged watersheds. Journal of Hydrology, *550*, 650-662.

Jenkins, K. M., & Boulton, A. J. (2003). Connectivity in a dryland river: Short-term aquatic microinvertebrate recruitment following floodplain inundation. Ecology, 84(10), 2708-2723.

Jin, S., Homer, C., Yang, L., Danielson, P., Dewitz, J., Li, C., ... & Howard, D. (2019). Overall methodology design for the United States national land cover database 2016 products. Remote Sensing, 11(24), 2971.

Johnson, M. F., Thorne, C. R., Castro, J. M., Kondolf, G. M., Mazzacano, C. S., Rood, S. B., & Westbrook, C. (2020). Biomic river restoration: A new focus for river management. River Research and Applications, 36(1), 3-12.

Jones, J., Börger, L., Tummers, J., Jones, P., Lucas, M., Kerr, J., ... & de Leaniz, C. G. (2019). A comprehensive assessment of stream fragmentation in Great Britain. Science of the Total Environment, 673, 756-762.

Jones, K. E., Dahl, T. A., & Corum, Z. P. (2018). Modeled Sedimentation in the Lower White River Countyline Levee Setback, Washington State: Comparison of 1D (HEC RAS) and 2D (AdH) Results. ERDC-CHL Vicksburg United States.

Jones, T. A. (2017). Impacts on ecosystems, corrective restoration practices, and prospects for recovery: nine case studies in the continental United States. The Rangeland Journal, 39(6), 431-450.

Junk, W. J., Bayley, P. B., & Sparks, R. E. (1989). The flood pulse concept in river-floodplain systems. Canadian special publication of fisheries and aquatic sciences, 106(1), 110-127.

Karasiak, N., Dejoux, J. F., Monteil, C., & Sheeren, D. (2021). Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. Machine Learning, 1-25.

Karpack, M. N., Morrison, R. R., & McManamay, R. A. (2020). Quantitative assessment of floodplain functionality using an index of integrity. Ecological Indicators, 111, 106051.

King, J.M. & Brown, C.A. (2018). Environmental flow assessments are not realizing their potential as an aid to basin planning. Frontiers in Environmental Science, 6, 113. Doi.org/10.3389/fenvs.2018.00113.

King, J.M., Brown, C.A., Paxton, B.R., & February, R.J. (2004). Development of DRIFT, a scenario-based methodology for environmental flow assessments. Report to the Water Research Commission, WRC Report No. 1159/1/04, South Africa. http://www.wrc.org.za/wp-content/uploads/mdocs/1159-1-041.pdf

Klijn, F., Asselman, N., & Wagenaar, D. (2018). Room for rivers: Risk reduction by enhancing the flood conveyance capacity of the netherlands' large rivers. Geosciences, 8(6), 224.

Knox, R. L., & Latrubesse, E. M. (2016). A geomorphic approach to the analysis of bedload and bed morphology of the Lower Mississippi River near the Old River Control Structure. Geomorphology, 268, 35-47.

Knox, R.L., Morrison, R.R., & Wohl, E.E. (2022a). Identification of artificial levees in the contiguous United States. Manuscript submitted for publication.

Knox, R.L., Morrison, R.R., & Wohl, E.E. (2022b). A river ran through it: floodplains as America's newest relict landform. Manuscript submitted for publication.

Knox, R.L., Wohl, E.E., & Morrison, R.R. (2022c). Levees don't protect, they disconnect: how artificial levees impact floodplain functions. Manuscript submitted for publication.

Koebel Jr, J. W., & Bousquin, S. G. (2014). The Kissimmee River restoration project and evaluation program, Florida, USA. Restoration Ecology, 22(3), 345-352.

Koebel, J.W., Bousquin S. G., Anderson D.H., Cheek, M.D., Carroll, C, Chen, H., Hanlon, C., Spencer, L., Anderson, B.C., Beck, T., Brunell, A., Sawyers, E., Marois, D., Dyson, A. (2021) Graham-Moore, A., & Sculley, S. Kissimmee River Restoration and Other Kissimmee Basin Initiatives. 2021 South Florida Environmental Report – Volume 1.

Kondolf, G. M., Boulton, A. J., O'Daniel, S., Poole, G. C., Rahel, F. J., Stanley, E. H., ... & Nakamura, K. (2006). Process-based ecological river restoration: visualizing three-dimensional connectivity and dynamic vectors to recover lost linkages. Ecology and society, 11(2).

Konrad, C. P., Black, R. W., Voss, F., & Neale, C. M. (2008). Integrating remotely acquired and field data to assess effects of setback levees on riparian and aquatic habitats in glacial-melt water rivers. River Research and Applications, 24(4), 355-372.

Kristensen, E. A., Kronvang, B., Wiberg-Larsen, P., Thodsen, H., Nielsen, C., Amor, E., ... & Baattrup-Pedersen, A. (2014). 10 years after the largest river restoration project in Northern Europe: Hydromorphological changes on multiple scales in River Skjern. Ecological engineering, 66, 141-149.

Kui, L., Stella, J.C., Shafroth, P.B., House, P.K., and Wilcox, A.C. (2017). The long-term legacy of geomorphic and riparian vegetation feedbacks on the dammed Bill Williams River, Arizona, USA. Ecohydrology, 10, e1839.

Kupfer, J. A., Meitzen, K. M., & Gao, P. (2015). Flooding and surface connectivity of Taxodium-Nyssa stands in a southern floodplain forest ecosystem. River Research and Applications, 31(10), 1299-1310.

Larsen, E. W., Girvetz, E. H., & Fremier, A. K. (2006). Assessing the effects of alternative setback channel constraint scenarios employing a river meander migration model. Environmental Management, 37(6), 880-897.

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. Geoscience Frontiers, 7(1), 3-10.

Lehner, B., Liermann, C.R., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., et al. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment, 9, 494-502.

Leopold, L. B., & Maddock, T. (1953). The hydraulic geometry of stream channels and some physiographic implications (Vol. 252). US Government Printing Office.

Lepori, F., Palm, D., Brännäs, E., & Malmqvist, B. (2005). Does restoration of structural heterogeneity in streams enhance fish and macroinvertebrate diversity?. Ecological applications, 15(6), 2060-2071.

Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. R News 2(3), 18--22.

Lindersson, S., Brandimarte, L., Mård, J., & Di Baldassarre, G. (2021). Global riverine flood risk–how do hydrogeomorphic floodplain maps compare to flood hazard maps? Natural Hazards and Earth System Sciences Discussions, 1-38.

Lininger, K. B., & Latrubesse, E. M. (2016). Flooding hydrology and peak discharge attenuation along the middle Araguaia River in central Brazil. Catena, 143, 90-101.

Luef, B., Aspetsberger, F., Hein, T., Huber, F., & Peduzzi, P. (2007). Impact of hydrology on free-living and particle-associated microorganisms in a river floodplain system (Danube, Austria). Freshwater Biology, 52(6), 1043-1057.

Magoulick, D. D., & Kobza, R. M. (2003). The role of refugia for fishes during drought: a review and synthesis. Freshwater biology, 48(7), 1186-1198.

Marron, D. C. (1992). Floodplain storage of mine tailings in the Belle Fourche River system: a sediment budget approach. Earth Surface Processes and Landforms, 17(7), 675-685.

Martínez-Fernández, V., González, E., López-Almansa, J. C., González, S. M., & de Jalón, D. G. (2017). Dismantling artificial levees and channel revetments promotes channel widening and regeneration of riparian vegetation over long river segments. Ecological Engineering, 108, 132-142.

Matella, M. K., & Merenlender, A. M. (2015). Scenarios for restoring floodplain ecology given changes to river flows under climate change: case from the San Joaquin River, California. River Research and Applications, 31(3), 280-290.

Matthews, W. J., Vaughn, C. C., Gido, K. B., & Marsh-Matthews, E. D. I. E. (2005). Southern plains rivers. Rivers of North America, 282-325.

Mayr, M. J., Besemer, K., Sieczko, A., Demeter, K., & Peduzzi, P. (2020). Bacterial community composition and function along spatiotemporal connectivity gradients in the Danube floodplain (Vienna, Austria). Aquatic sciences, 82(2), 1-12.

McCluney, K. E., Poff, N. L., Palmer, M. A., Thorp, J. H., Poole, G. C., Williams, B. S., ... & Baron, J. S. (2014). Riverine macrosystems ecology: sensitivity, resistance, and resilience of whole river basins with human alterations. Frontiers in Ecology and the Environment, 12(1), 48-58.

Meade, R. H., & Moody, J. A. (2010). Causes for the decline of suspended-sediment discharge in the Mississippi River system, 1940–2007. Hydrological Processes: An International Journal, 24(1), 35-49.

Mertes, L. A. (1997). Documentation and significance of the perirheic zone on inundated floodplains. Water Resources Research, 33(7), 1749-1762.

Mertes, L.A. (2000). Inundation hydrology. In E. Wohl (Ed.), Inland flood hazards: Human, riparian, and acquatic communities (pp. 145-166). Cambridge, UK: Cambridge University Press.

Miguez, M. G., Veról, A. P., De Sousa, M. M., & Rezende, O. M. (2015). Urban floods in lowlands—levee systems, unplanned urban growth and river restoration alternative: a case study in Brazil. Sustainability, 7(8), 11068-11097.

Montgomery, D. R. (1999). Process domains and the river continuum 1. JAWRA Journal of the American Water Resources Association, 35(2), 397-410.

Moyle, P. B., & Mount, J. F. (2007). Homogenous rivers, homogenous faunas. Proceedings of the National Academy of Sciences, 104(14), 5711-5712.

Murray, A., & Wunner, R. (1980). A Study of the Jacoby Creek Watershed, Humboldt County, California. Jacoby Creek Canyon Community, Incorporated.

Naiman, R. J., Decamps, H., & Pollock, M. (1993). The role of riparian corridors in maintaining regional biodiversity. Ecological applications, 3(2), 209-212.

Nakamura, F., Ishiyama, N., Sueyoshi, M., Negishi, J. N., & Akasaka, T. (2014). The significance of meander restoration for the hydrogeomorphology and recovery of wetland organisms in the Kushiro River, a lowland river in Japan. Restoration ecology, 22(4), 544-554.

Nanson, G. C., & Croke, J. C. (1992). A genetic classification of floodplains. Geomorphology, 4(6), 459-486.

Nardi, F., Annis, A., Di Baldassarre, G., Vivoni, E. R., & Grimaldi, S. (2019). GFPLAIN250m, a global high-resolution dataset of Earth's floodplains. Scientific data, 6(1), 1-6.

Nardi, F., Biscarini, C., Di Francesco, S., Manciola, P., Ubertini, L. (2013). Comparing a large-scale DEM-based 355 floodplain delineation algorithm with standard flood maps: The Tiber River Basin case study. Irrig. Drain. 62, 356 11-19.

Nardi, F., Morrison, R. R., Annis, A., & Grantham, T. E. (2018). Hydrologic scaling for hydrogeomorphic floodplain mapping: Insights into human-induced floodplain disconnectivity. River Research and Applications, 34(7), 675-685.

Nardi, F., Vivoni, E. R., & Grimaldi, S. (2006). Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. Water Resources Research, 42(9).

Nilsson, C., Reidy, C.A., Dynesius, M., & Revenga, C. (2005). Fragmentation and flow regulation of the world's large river systems. Science, 308, 405-408.

Noe, G. B., Boomer, K., Gillespie, J. L., Hupp, C. R., Martin-Alciati, M., Floro, K., ... & Strano, S. (2019). The effects of restored hydrologic connectivity on floodplain trapping vs. release of phosphorus, nitrogen, and sediment along the Pocomoke River, Maryland USA. Ecological Engineering, 138, 334-352.

Noe, G. B., & Hupp, C. R. (2005). Carbon, nitrogen, and phosphorus accumulation in floodplains of Atlantic Coastal Plain rivers, USA. Ecological Applications, 15(4), 1178-1190.

Noe, G. B. (2013). Interactions among hydrogeomorphology, vegetation, and nutrient biogeochemistry in floodplain ecosystems.

Noe, G. B., Hupp, C. R., & Rybicki, N. B. (2013). Hydrogeomorphology influences soil nitrogen and phosphorus mineralization in floodplain wetlands. Ecosystems, 16(1), 75-94.

Norwood, Joy (2020, March 8). Riverwalk project dies in Brookville. Tri-County Weekend. https://m.thecourierexpress.com/tri_county_sunday/news/local/riverwalk-project-dies-in-brookville/article_34a0f4ec-2da1-5b78-863e-26468c50b3a9.html

Ogston, L., S. Gidora, M. Foy, and J. Rosenfeld (2015), Watershed-scale effectiveness of floodplain habitat restoration for juvenile coho salmon in the Chilliwack River, British Columbia, Can. J. Fish. Aquat. Sci., 72, 479–490.

Opperman, J. J., Galloway, G. E., Fargione, J., Mount, J. F., Richter, B. D., & Secchi, S. (2009). Sustainable floodplains through large-scale reconnection to rivers. Science, 326(5959), 1487-1488.

Opperman, J. J., Moyle, P. B., Larsen, E. W., Florsheim, J. L., & Manfree, A. D. (2017). Floodplains: Processes and management for ecosystem services. Univ of California Press.

Orr, C. H., Stanley, E. H., Wilson, K. A., & Finlay, J. C. (2007). Effects of restoration and reflooding on soil denitrification in a leveed Midwestern floodplain. Ecological Applications, 17(8), 2365-2376.

Palmer, M. A., Hondula, K. L., & Koch, B. J. (2014). Ecological restoration of streams and rivers: shifting strategies and shifting goals. Annual Review of Ecology, Evolution, and Systematics, 45, 247-269.

Pander, J., Mueller, M., & Geist, J. (2018). Habitat diversity and connectivity govern the conservation value of restored aquatic floodplain habitats. Biological Conservation, 217, 1-10.

Park, E., & Latrubesse, E. M. (2017). The hydro-geomorphologic complexity of the lower Amazon River floodplain and hydrological connectivity assessed by remote sensing and field control. Remote Sensing of Environment, 198, 321-332.

Parsons, H., & Gilvear, D. (2002). Valley floor landscape change following almost 100 years of flood embankment abandonment on a wandering gravel-bed river. River Research and Applications, 18(5), 461-479.

Pasquale, N., Perona, P., Schneider, P., Shrestha, J., Wombacher, A., & Burlando, P. (2011). Modern comprehensive approach to monitor the morphodynamic evolution of a restored river corridor. Hydrology and Earth System Sciences, 15(4), 1197-1212.

Passalacqua, P., Belmont, P., Staley, D. M., Simley, J. D., Arrowsmith, J. R., Bode, C. A., ... & Wheaton, J. M. (2015). Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. Earth-Science Reviews, 148, 174-193.

Pedersen, M. L., Andersen, J. M., Nielsen, K., & Linnemann, M. (2007). Restoration of Skjern River and its valley: project description and general ecological changes in the project area. Ecological engineering, 30(2), 131-144.

Peipoch, M., Brauns, M., Hauer, F. R., Weitere, M., & Valett, H. M. (2015). Ecological simplification: human influences on riverscape complexity. BioScience, 65(11), 1057-1065.

Pennington, D. N., Hansel, J. R., & Gorchov, D. L. (2010). Urbanization and riparian forest woody communities: diversity, composition, and structure within a metropolitan landscape. Biological Conservation, 143(1), 182-194.

Pinter, N. (2005). One step forward, two steps back on US floodplains. Science, 308, 207.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., ... & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nature communications, 11(1), 1-11.

Poeppl, R. E., Keesstra, S. D., & Maroulis, J. (2017). A conceptual connectivity framework for understanding geomorphic change in human-impacted fluvial systems. Geomorphology, 277, 237-250.

Poff, N.L., Allen, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., & Stromberg, J.C. (1997). The natural flow regime. BioScience 47, 769-784.

Poff, N. L., Olden, J. D., Merritt, D. M., & Pepin, D. M. (2007). Homogenization of regional river dynamics by dams and global biodiversity implications. Proceedings of the National Academy of Sciences, 104(14), 5732-5737.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rajib, A., Zheng, Q., Golden, H. E., Wu, Q., Lane, C. R., Christensen, J. R., ... & Nardi, F. (2021). The changing face of floodplains in the Mississippi River Basin detected by a 60-year land use change dataset. Scientific data, 8(1), 1-11.

Ramler, D., & Keckeis, H. (2019). Effects of large-river restoration measures on ecological fish guilds and focal species of conservation in a large European river (Danube, Austria). Science of the total environment, 686, 1076-1089.

Reckendorfer, W., Funk, A., Gschöpf, C., Hein, T., & Schiemer, F. (2013). Aquatic ecosystem functions of an isolated floodplain and their implications for flood retention and management. Journal of Applied Ecology, 50(1), 119-128.

Records, R. M., Wohl, E., & Arabi, M. (2016). Phosphorus in the river corridor. Earth-Science Reviews, 158, 65-88.

Remo, J. W., Carlson, M., & Pinter, N. (2012). Hydraulic and flood-loss modeling of levee, floodplain, and river management strategies, Middle Mississippi River, USA. Natural hazards, 61(2), 551-575.

Remo, J. W. F., Guida, R. J., & Secchi, S. (2017). Screening the suitability of levee protected areas for strategic floodplain reconnection along the LaGrange segment of the Illinois River, USA. River Research and Applications, 33(6), 863-878.

Rieger, I., Lang, F., Kowarik, I., & Cierjacks, A. (2014). The interplay of sedimentation and carbon accretion in riparian forests. Geomorphology, 214, 157-167.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913-929.

Roering, J. J., Mackey, B. H., Marshall, J. A., Sweeney, K. E., Deligne, N. I., Booth, A. M., ... & Cerovski-Darriau, C. (2013). 'You are HERE': Connecting the dots with airborne lidar for geomorphic fieldwork. Geomorphology, 200, 172-183.

Rohde, S., Schütz, M., Kienast, F., & Englmaier, P. (2005). River widening: an approach to restoring riparian habitats and plant species. River Research and Applications, 21(10), 1075-1094.

Roni, P., Hall, J. E., Drenner, S. M., & Arterburn, D. (2019). Monitoring the effectiveness of floodplain habitat restoration: A review of methods and recommendations for future monitoring. Wiley Interdisciplinary Reviews: Water, 6(4), e1355.

Rykiel JR, E. J. (1985). Towards a definition of ecological disturbance. Australian Journal of Ecology, 10(3), 361-365.

Safanelli, J. L., Poppiel, R. R., Ruiz, L. F. C., Bonfatti, B. R., Mello, F. A. D. O., Rizzo, R., & Demattê, J. A. (2020). Terrain Analysis in Google Earth Engine: A Method Adapted for High-Performance Global-Scale Analysis. ISPRS International Journal of Geo-Information, 9(6), 400.

Sanchez-Pérez, J. M., & Trémolières, M. (2003). Change in groundwater chemistry as a consequence of suppression of floods: the case of the Rhine floodplain. Journal of Hydrology, 270(1-2), 89-104.

Scheel, K., Morrison, R. R., Annis, A., & Nardi, F. (2019). Understanding the Large-Scale Influence of Levees on Floodplain Connectivity Using a Hydrogeomorphic Approach. JAWRA Journal of the American Water Resources Association, 55(2), 413-429.

Schiemer, F. (2000). Fish as indicators for the assessment of the ecological integrity of large rivers. Hydrobiologia, 422, 271-278.

Schiemer, F., Baumgartner, C., & Tockner, K. (1999). Restoration of floodplain rivers: The 'Danube restoration project'. River Research and Applications, 15(1-3), 231-244.

Schlesinger, W. H., Raikes, J. A., Hartley, A. E., & Cross, A. F. (1996). On the spatial pattern of soil nutrients in desert ecosystems: ecological archives E077-002. Ecology, 77(2), 364-374.

Schmitt, L., Morris, D., & Kondolf, G. M. (2018). Managing Floods in large river basins in Europe: The Rhine River. In Managing Flood Risk (pp. 75-89). Palgrave Macmillan, Cham.

Schumm, S. A. (1956). Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. Geological society of America bulletin, 67(5), 597-646.

Seaber, P. R., Kapinos, F. P., & Knapp, G. L. (1987). Hydrologic unit maps: US Geological Survey water supply paper 2294. US Geological Survey.

Sedell, J.R., Reeves, G.H., Hauer, F.R., Stanford, J.A., & Hawkins, C.P. (1990). Role of refugia in recovery from disturbances: modern fragmented and disconnected river systems. Environmental Management, 14(5), 711-724.

Shafroth, P.B., Wilcox, A.C., Lytle, D.A., Hickey, J.T., Andersen, D.C., Beauchamp, V.B., Hautzinger, A., McMullen, L.E., and Warner, A. (2010). Ecosystem effects of environmental flows: modelling and experimental floods in a dryland river. Freshwater Biology, 55, 68-85.

Sharpley, A., Jarvie, H. P., Buda, A., May, L., Spears, B., & Kleinman, P. (2013). Phosphorus legacy: Overcoming the effects of past management practices to mitigate future water quality impairment. Journal of environmental quality, 42(5), 1308-1326.

Sheaffer, J. R., Mullan, J. D., & Hinch, N. B. (2002). Encouraging wise use of floodplains with market-based incentives. Environment: Science and Policy for Sustainable Development, 44(1), 32-43.

Sheibley, R. W., Ahearn, D. S., & Dahlgren, R. A. (2006). Nitrate loss from a restored floodplain in the lower Cosumnes River, California. Hydrobiologia, 571(1), 261-272.

Shi, L., Wang, Y., Jia, Y., Lu, C., Lei, G., & Wen, L. (2017). Vegetation cover dynamics and resilience to climatic and hydrological disturbances in seasonal floodplain: The effects of hydrological connectivity. Frontiers in plant science, 8, 2196.

Simenstad, C. A., Burke, J. L., O'Connor, J. E., Cannon, C., Heatwole, D. W., Ramirez, M. F., ... & Jones, K. L. (2011). Columbia River estuary ecosystem classification: concept and application.

Škute, A., Gruberts, D., Soms, J., & Paidere, J. (2008). Ecological and hydrological functions of the biggest natural floodplain in Latvia. Ecohydrology & Hydrobiology, 8(2-4), 291-306.

Smith, D. L., Miner, S. P., Theiling, C. H., Behm, R. L., & Nestler, J. M. (2017). Levee setbacks: an innovative, cost-effective, and sustainable solution for improved flood risk management. Environmental Laboratory (US).

Sofia, G. (2020). Combining geomorphometry, feature extraction techniques and Earth-surface processes research: The way forward. Geomorphology, 355, 107055.

Sofia, G., Fontana, G. D., & Tarolli, P. (2014). High-resolution topography and anthropogenic feature extraction: testing geomorphometric parameters in floodplains. Hydrological Processes, 28(4), 2046-2061.

Sparks, R. E. (1995). Need for ecosystem management of large rivers and their floodplains. BioScience, 45(3), 168-182.

Sparks, R. E., Bayley, P. B., Kohler, S. L., & Osborne, L. L. (1990). Disturbance and recovery of large floodplain rivers. Environmental management, 14(5), 699-709.

Sparks, R.E., Blodgett, K.D., Casper, A.F., Hagy, H.M., Lemke, M.J., Velho, L.F.M., & Rodrigues, L.C. (2017). Why experiment with success? Opportunities and risks in applying assessment and adaptive management to the Emiquon floodplain restoration project. Hydrobiologia, 804, 177-200.

Stammel, B., Cyffka, B., Geist, J., Müller, M., Pander, J., Blasch, G., ... & Weißbrod, M. (2012). Floodplain restoration on the Upper Danube (Germany) by re-establishing back water and sediment dynamics: a scientific monitoring as part of the implementation. River systems, 20(1-2), 55-70.

Stammel, B., Stäps, J., Schwab, A., & Kiehl, K. (2021). Are natural floods accelerators for streambank vegetation development in floodplain restoration?. International Review of Hydrobiology.

Stanford, J. A., Ward, J. V., Liss, W. J., Frissell, C. A., Williams, R. N., Lichatowich, J. A., & Coutant, C. C. (1996). A general protocol for restoration of regulated rivers. Regulated Rivers: Research & Management, 12(4-5), 391-413.

Steinfeld, C. M., & Kingsford, R. T. (2013). Disconnecting the floodplain: earthworks and their ecological effect on a dryland floodplain in the Murray–Darling Basin, Australia. River Research and Applications, 29(2), 206-218.

Steinfeld, C. M., Kingsford, R. T., & Laffan, S. W. (2013). Semi-automated GIS techniques for detecting floodplain earthworks. Hydrological Processes, 27(4), 579-591.

Stella, J.C., Hayden, M.K., Battles, J.J., Piegay, H., Dufour, S., & Fremier, A.K. (2011). The role of abandoned channels as refugia for sustaining pioneer riparian forest ecosystems. Ecosystems, 14, 776-790.

Stoffers, T., Collas, F. P. L., Buijse, A. D., Geerling, G. W., Jans, L. H., Van Kessel, N., ... & Nagelkerke, L. A. J. (2021). 30 years of large river restoration: How long do restored floodplain channels remain suitable for targeted rheophilic fishes in the lower river Rhine?. Science of the Total Environment, 755, 142931.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society: Series B (Methodological), 36(2), 111-133.

Stout, J.C., & Belmont, P. (2014). TerEx Toolbox for semi-automated selection of fluvial terrace and floodplain features from lidar. Earth Surface Processes and Landforms, 39, 569-580.

Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. Eos, Transactions American Geophysical Union, 38(6), 913-920.

Strahler, A. N. (1952). Hypsometric (area-altitude) analysis of erosional topography. Geological society of America bulletin, 63(11), 1117-1142.

Sutfin, N. A., Wohl, E. E., & Dwire, K. A. (2016). Banking carbon: a review of organic carbon storage and physical factors influencing retention in floodplains and riparian ecosystems. Earth Surface Processes and Landforms, 41(1), 38-60.

Sutfin, N. A., Wohl, E., Fegel, T., Day, N., & Lynch, L. (2021). Logjams and channel morphology influence sediment storage, transformation of organic matter, and carbon storage within mountain stream corridors. Water Resources Research, 57(5), e2020WR028046.

Swab, R. M., Zhang, L., & Mitsch, W. J. (2008). Effect of hydrologic restoration and Lonicera maackii removal on herbaceous understory vegetation in a bottomland hardwood forest. Restoration Ecology, 16(3), 453-463.

Swenson, R. O., Whitener, K., & Eaton, M. (2003). Restoring floods on floodplains: riparian and floodplain restoration at the Cosumnes River Preserve. In California Riparian Systems: Processes and Floodplain Management, Ecology, Restoration, 2001 Riparian Habitat and Floodplains Conference Proceedings, Faber PM (ed.). Riparian Habitat Joint Venture: Sacramento, CA (pp. 224-229).

Syvitski, J.P.M., Vörösmarty, C.J., Kettner, A.J., & Green, P. (2005). Impact of humans on the flux of terrestrial sediment to the global coastal ocean. Science*, 308*, 376-380.

Tank, J. L., Rosi-Marshall, E. J., Griffiths, N. A., Entrekin, S. A., & Stephen, M. L. (2010). A review of allochthonous organic matter dynamics and metabolism in streams. Journal of the North American Benthological Society, 29(1), 118-146.

Tena, A., Piégay, H., Seignemartin, G., Barra, A., Berger, J. F., Mourier, B., & Winiarski, T. (2020). Cumulative effects of channel correction and regulation on floodplain terrestrialisation patterns and connectivity. Geomorphology, 354, 107034.

Theng, V., Hashimoto, K., Uk, S., Tanaka, T., Yoshioka, H., and Yoshimura, C. (2022). Hydrodynamics-based modeling of phosphorus balance and dynamics in a large tropical floodplain. Environmental Advances, 7, 100176.

Thorp, J. H., & Delong, M. D. (1994). The riverine productivity model: an heuristic view of carbon sources and organic processing in large river ecosystems. Oikos, 305-308.

Tobin, G. A. (1995). The levee love affair: a stormy relationship? 1. JAWRA Journal of the American Water Resources Association, 31(3), 359-367.

Tockner, K., Lorang, M. S., & Stanford, J. A. (2010). River flood plains are model ecosystems to test general hydrogeomorphic and ecological concepts. River research and applications, 26(1), 76-86.

Tockner, K., Malard, F., & Ward, J. V. (2000). An extension of the flood pulse concept. Hydrological processes, 14(16-17), 2861-2883.

Tockner, K., & Schiemer, F. (1997). Ecological aspects of the restoration strategy for a river-floodplain system on the Danube River in Austria. Global Ecology and Biogeography Letters, 321-329.

Tockner, K., Schiemer, F., Baumgartner, C., Kum, G., Weigand, E., Zweimüller, I., & Ward, J. V. (1999). The Danube restoration project: species diversity patterns across connectivity gradients in the floodplain system. River Research and Applications, 15(1-3), 245-258.

Tockner, K., Schiemer, F., & Ward, J. V. (1998). Conservation by restoration: the management concept for a river-floodplain system on the Danube River in Austria. Aquatic Conservation: Marine and Freshwater Ecosystems, 8(1), 71-86.

Tockner, K., & Stanford, J. A. (2002). Riverine flood plains: present state and future trends. Environmental conservation, 29(3), 308-330.

Tockner, K., & Ward, J. V. (1999). Biodiversity along riparian corridors. Large rivers, 11(3), 293-310.

Tomasek, A. A., Hondzo, M., Kozarek, J. L., Staley, C., Wang, P., Lurndahl, N., & Sadowsky, M. J. (2019). Intermittent flooding of organic-rich soil promotes the formation of denitrification hot moments and hot spots. Ecosphere, 10(1), e02549.

Toth, L. A. (2010). Unrealized expectations for restoration of a floodplain plant community. Restoration Ecology, 18(6), 810-819.

Toth, L. A., Melvin, S. L., Arrington, D. A., & Chamberlain, J. (1998). Hydrologic manipulations of the channelized Kissimmee River: implications for restoration. BioScience, 48(9), 757-764.

Toth, L. A., & van der Valk, A. (2012). Predictability of flood pulse driven assembly rules for restoration of a floodplain plant community. Wetlands Ecology and Management, 20(1), 59-75.

Trowbridge, W. B. (2007). The role of stochasticity and priority effects in floodplain restoration. Ecological applications, 17(5), 1312-1324.

Valentine, A., & Kalnins, L. (2016). An introduction to learning algorithms and potential applications in geomorphometry and earth surface dynamics. Earth surface dynamics, 4(2), 445-460.

Van Manh, N., Viet Dung, N., Hung, N.N., Kummu, M., Merz, B., and Apel, H. (2015). Future sediment dynamics in the Mekong Delta floodplains: impacts of hydropower development, climate change and sea level rise. Global and Planetary Change, 127, 22-33.

Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. Canadian journal of fisheries and aquatic sciences, 37(1), 130-137.

Vileisis, A. (1999). Discovering the unknown landscape: a history of America's wetlands. Island Press.

Vörösmarty, C.J., Meybeck, M., Fekete, B., Sharma, K., Green, P., and Syvitski, J.P.M. (2003). Anthropogenic sediment retention: major global impacts from registered river impoundments. Global and Planetary Change, 39, 169-190.

Wang, L., Butcher, A. S., Stuart, M. E., Gooddy, D. C., & Bloomfield, J. P. (2013). The nitrate time bomb: a numerical way to investigate nitrate storage and lag time in the unsaturated zone. Environmental Geochemistry and Health, 35(5), 667-681.

Ward, J. V. (1989). The four-dimensional nature of lotic ecosystems. Journal of the North American Benthological Society, 8(1), 2-8.

Ward, J. V., & Stanford, J. A. (1995). Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. Regulated rivers: research & management, 11(1), 105-119.

Ward, J.V., Tockner, K., & Schiemer, F. (1999). Biodiversity of floodplain river ecosystems: ecotones and connectivity. Regulated Rivers: Research and Management, 15, 125-139.

Westbrook, C. J., Cooper, D. J., & Baker, B. W. (2011). Beaver assisted river valley formation. River Research and Applications, 27(2), 247-256.

White, G. F. (2000). Water science and technology: some lessons from the 20th century. Environment: Science and Policy for Sustainable Development, 42(1), 30-38.

White, G. F., Kates, R. W., & Burton, I. (2001). Knowing better and losing even more: the use of knowledge in hazards management. Global Environmental Change Part B: Environmental Hazards, 3(3), 81-92.

Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation. https://dplyr.tidyverse.org/

Williams, R. D., Bangen, S., Gillies, E., Kramer, N., Moir, H., & Wheaton, J. (2020). Let the river erode! Enabling lateral migration increases geomorphic unit diversity. Science of The Total Environment, 715, 136817.

Wing, O. E., Bates, P. D., Neal, J. C., Sampson, C. C., Smith, A. M., Quinn, N., ... & Krajewski, W. F. (2019). A new automated method for improved flood defense representation in large-scale hydraulic models. Water Resources Research, 55(12), 11007-11034. DOI: 10.1029/2019WR02597.

Wing, O. E. J., P. D. Bates, C. C. Sampson, A. M. Smith, K. A. Johnson, and T. A. Erickson (2017), Validation of a 30 m resolution flood hazard model of the conterminous United States, Water Resour. Res., 53, 7968–7986, doi:10.1002/2017WR020917.

Wohl, E. (2005). Disconnected rivers: Human impacts to rivers in the United States. Reviews in Engineering Geology, 16, 19-34.

Wohl, E. (2014). Rivers in the landscape. John Wiley & Sons Inc.

Wohl, E. (2017). Connectivity in rivers. Progress in Physical Geography, 41(3), 345-362., DOI: 10.1177/0309133317714972.

Wohl, E. (2018), Human Alterations of Rivers, in Sustaining River Ecosystems and Water Resources. Springer Briefs in Environmental Science. Springer, Cham, 59-104.

Wohl, E. (2020). Wood process domains and wood loads on floodplains. Earth Surface Processes and Landforms, 45, 144-156.

Wohl, E. (2021). An Integrative Conceptualization of Floodplain Storage. Reviews of Geophysics, 59(2), e2020RG000724.

Wohl, E., Bledsoe, B. P., Jacobson, R. B., Poff, N. L., Rathburn, S. L., Walters, D. M., & Wilcox, A. C. (2015). The natural sediment regime in rivers: Broadening the foundation for ecosystem management. BioScience, 65(4), 358-371.

Wohl, E., Castro, J., Cluer, B., Merritts, D., Powers, P., Staab, B., & Thorne, C. (2021). Rediscovering, reevaluating, and restoring lost river-wetland corridors. Frontiers in Earth Science, 9, 511.

Wohl, E., Hall Jr, R. O., Lininger, K. B., Sutfin, N. A., & Walters, D. M. (2017). Carbon dynamics of river corridors and the effects of human alterations. Ecological Monographs, 87(3), 379-409.

Wohl, E., & Iskin, E. P. (2021). Damming the wood falls. Science advances, 7(50), eabj0988.

Wohl, E., Kramer, N., Ruiz-Villanueva, V., Scott, D. N., Comiti, F., Gurnell, A. M., ... & Fausch, K. D. (2019). The natural wood regime in rivers. BioScience, 69(4), 259-273.

Wohl, E., Lane, S. N., & Wilcox, A. C. (2015). The science and practice of river restoration. Water Resources Research, 51(8), 5974-5997.

Wohl, E., Lininger, K. B., & Baron, J. (2017). Land before water: The relative temporal sequence of human alteration of freshwater ecosystems in the conterminous United States. Anthropocene, 18, 27-46.

Wohl, E., Lininger, K. B., & Scott, D. N. (2018). River beads as a conceptual framework for building carbon storage and resilience to extreme climate events into river management. Biogeochemistry, 141(3), 365-383.

Wohl, E., Marshall, A. E., Scamardo, J., White, D., & Morrison, R. R. (2022). Biogeomorphic influences on river corridor resilience to wildfire disturbances in a mountain stream of the Southern Rockies, USA. Science of The Total Environment, 153321.

Wolman, M.G., & Leopold, L.B. (1957). River flood plains: some observations on their formation. U.S. Geological Survey Professional Paper 282-C, Washington, DC.

Wolman, M.G., & Miller, J.P. (1960). Magnitude and frequency of forces in geomorphic processes. J. Geol., 68, 54–74.

Woltemade, C. J., & Potter, K. W. (1994). A watershed modeling analysis of fluvial geomorphologic influences on flood peak attenuation. Water Resources Research, 30(6), 1933-1942.

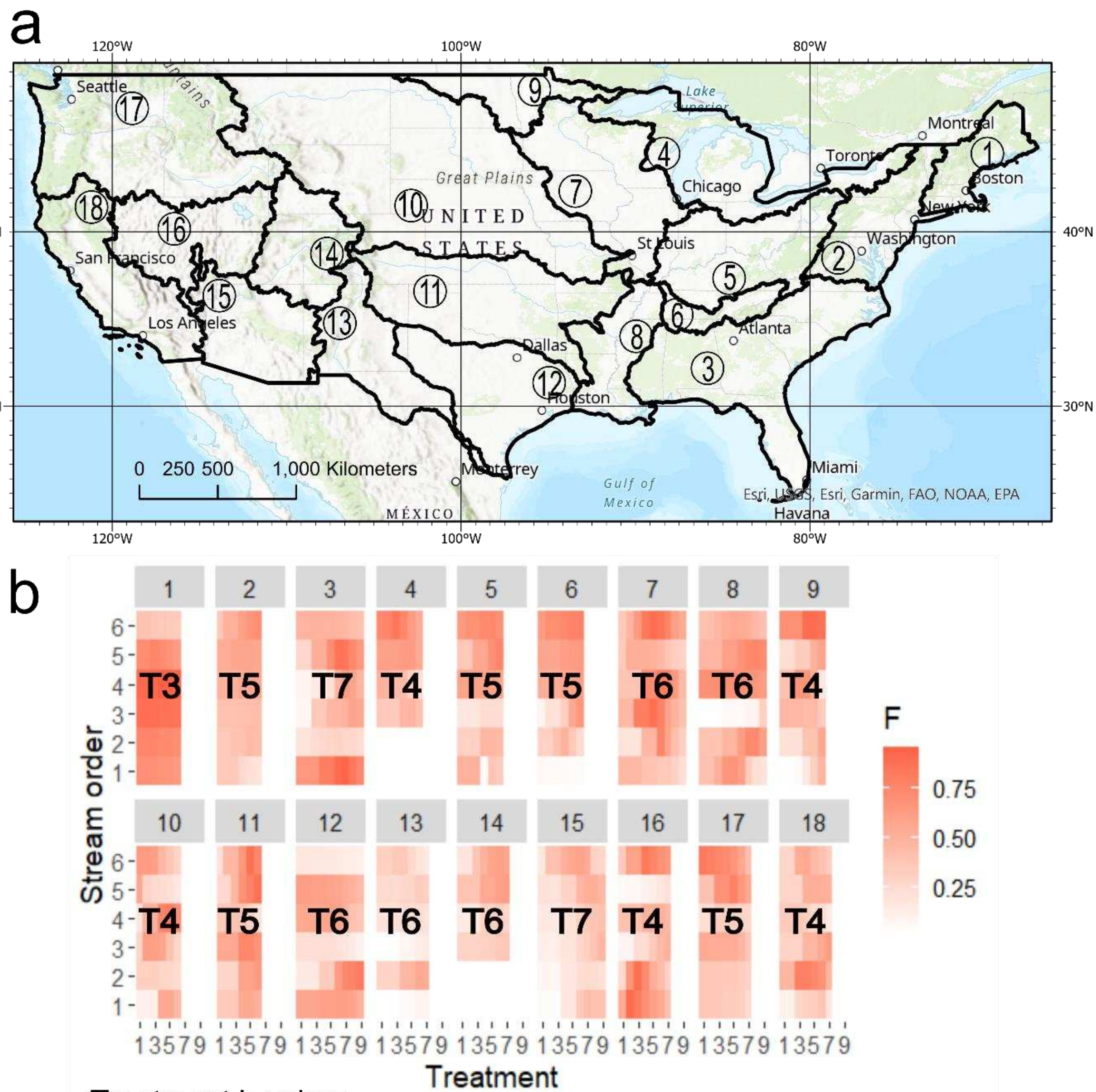
Zhang, L., & Mitsch, W. J. (2007). Sediment chemistry and nutrient influx in a hydrologically restored bottomland hardwood forest in Midwestern USA. River Research and Applications, 23(9), 1026-1037.

Zhu, T., Lund, J. R., Jenkins, M. W., Marques, G. F., & Ritzema, R. S. (2007). Climate change, urbanization, and optimal long-term floodplain protection. Water Resources Research, 43(6).

Appendix A: Supplemental information for Chapter 2: Identification of Artificial Levees in the Contiguous United States

Our initial modeling efforts focused on discovering an accurate and parsimonious model that could be applied to the contiguous United States in a reasonable amount of time using a personal computer and Google Earth Engine. I did not establish an accuracy threshold but knew that an effective model needed to at least detect known artificial levees well. Initial efforts employed SVM and GLM models with five geomorphic variables, the 2016 NLCD variable, a "distance from any stream" variable, and the ~2,100 HUC8 basins as a factor variable with ~2,100 levels. I created a national probability map using a partial GLM model and discovered that the presence of levees in the training data strongly influenced the model output so that it was not possible to identify areas with potential levees in the ~1,200 basins without NLD coverage. In the second iteration, I added Random Forest models and generated training and validation data from the GFPLAIN flood model area. Covering about 10% of the land area of the contiguous United States, the floodplain area appears to provide an effective way to reduce the studies' geographical and variable domain and computational load. The assumption that potential levees fall within the floodplain area is supported by the presence of most NLD levees in the floodplain. However, this assumption could be revisited later. I attempted to generate all of the variables I used here to include distance from stream order 1-10 in GEE. Google Earth Engine required ~130 hours to collect five geomorphic variables, the NLCD, and stream order 4-10 (with NA values beyond 5 km due to memory constraints) at 30,000 locations. This method was too slow to apply nationally and with model performance stabilized at 0.7 kappa, the method also was not sufficiently accurate. These modelling efforts, and those in the case study, led us to the present method and model.

Based on the case study, I expected the full model (model 1) to outperform other models. However, after it took ~24 hours to collect the geomorphic variables from GEE for the larger (n ~ 3,060,000) sample, I began looking for a more parsimonious model. I discovered that the land use and spatial model (model 12) outperformed the full model which included geomorphic variables with the larger sample (0.94 kappa versus 0.84) and similar but slightly lower performance with the smaller sample (0.65 kappa versus 0.69). Model performance in the case study using both sets of variables was similar as well.



Treatment b values

T1	T2	T3	T4	T5	T6	T7	T8	T9
0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.4	0.42

Figure A.1. Calibration results from the 100-year hydrogeomorphic floodplain delineation. (a) CONUS map of the 18 2-digit HUC basins with the number denoting the basin. (b) The F measure of fit is plotted for each stream order and treatment b value in each 2-digit HUC basin. The selected treatment (corresponding to b value), with the highest average F measure of fit in each basin, is listed on top of the tile plot. The b value corresponding to each treatment is listed at the bottom of the plot.

Random Forest, 1,700 samples per basin, Non levee to levee ratio ~1, Model 1

	Reference	
Prediction	Levee	Not levee
Levee	3611	534
Not levee	818	4013

GLM, 1,700 samples per basin, Non levee to levee ratio ~1, Model 1

	Reference	
Prediction	Levee	Not levee
Levee	2996	927
Not levee	1433	3620

SVM, 1,700 samples per basin, Non levee to levee ratio ~1, Model 1

	Reference	
Prediction	Levee	Not levee
Levee	3050	1379
Not levee	637	3910

RF, 1,700 samples per basin, Non levee to levee ratio ~ 0.7, Model 1

	Reference	
Prediction	Levee	Not levee
Levee	4696	672
Not levee	703	3108

RF, 1,700 samples per basin, Non levee to levee ratio ~ 0.7, Model 12

	Reference	
Prediction	Levee	Not levee
Levee	4700	853
Not levee	699	2927

RF, 170,000 samples per basin, Non levee to levee ratio ~ 0.7, Model 1

	Reference	
Prediction	Levee	Not levee
Levee	496,700	37,524
Not levee	25,720	338,013

RF, 170,000 samples per basin, Non levee to levee ratio ~ 0.7, Model 12

	Reference	
Prediction	Levee	Not levee
Levee	515,581	18,185
Not levee	6,839	357,352

RF, 170,000 samples per basin, Non levee to levee ratio ~ 0.7, Model 12 + relative elevation

	Reference	
Prediction	Levee	Not levee
Levee	507,553	31,177
Not levee	14,867	344,360

RF, 170,000 samples per basin, Non levee to levee ratio ~ 0.7, Model 12 + profile curvature

	Reference	
Prediction	Levee	Not levee
Levee	509,562	27,525
Not levee	12,858	348,012

RF, 170,000 samples per basin, Non levee to levee ratio ~ 0.7, Model 12 + aspect difference

	Reference	
Prediction	Levee	Not levee
Levee	504,405	34,704
Not levee	18,015	340,833

RF, 170,000 samples per basin, Non levee to levee ratio ~ 0.7, Model 12 + slope

	Reference	
Prediction	Levee	Not levee
Levee	507,670	31,682
Not levee	14,750	343,855

Figure A.2. Confusion matrices of CONUS models from table 2.4.

Figure A.3. Two artificial levees in the NLD. (a) Over 7 m high, a massive levee west of the overbank structure at the Old River Complex, Louisiana, USA. (b) Almost invisible, Fort Collins North- Cache La Poudre River, ~1 m high, Colorado, USA, indicated by two arrows.

RF, Lower Mississippi Basin, 170,000 samples, Model 12

	Reference	
Prediction	Levee	Not levee
Levee	27,315	1,948
Not levee	1,812	18,964

RF, Lower Mississippi Basin, 1,700,000 samples, Model 12

	Reference	
Prediction	Levee	Not levee
Levee	200,651	3,255
Not levee	7,056	292,352

RF, Lower Mississippi Basin, 1,700,000 samples, Model 12 + Distance from stream of order 10

	Reference	
Prediction	Levee	Not levee
Levee	203,107	1,860
Not levee	4,600	293,747

Figure A.4. Confusion matrices of LMR models from table 2.6.

Appendix B: Supplemental information for Chapter 3: A river ran through it: floodplains as America's newest relict landform

Floodplain model calibration

I delineated floodplains using the hydrogeomorphic floodplain delineation tool, GFPLAIN (Nardi et al., 2006; Nardi et al., 2013), which runs as three python scripts in ArcGIS Pro. There are two main steps: (1) DEM pit filling, determination of flow direction and cell accumulation, and identification of the river network based on contributing area threshold, and (2) flow height estimation along the network based on the upstream contributing area and the following scaling relationship based on Leopold and Maddock (1953):

$$FH = aA^b \tag{1}$$

where FH is the flow depth for the *A* contributing area, with *a* and *b* dimensionless scaling parameters (Scheel et al., 2019). Regional flow depth – contributing area scaling laws are especially valid (Annis et al., 2022), so I calibrated the scaling parameters in each 2-digit HUC basin with FEMA flood map areas A and AE (Table B.1) along streams of orders one through six, as described in more detail in chapter 2. I selected 50 km^2 for the contributing area threshold, based on previous research (Annis et al., 2019; Scheel et al., 2019) and applied GFPLAIN separately in each HUC2 basin.

Stream order contribution normalization

I compared stream order contribution to disagreement areas to that in the agreement areas using the following equation, where "X" indicates the stream order and "disagreement area" is either anthropogenically connected (AC) or anthropogenically disconnected (AD) floodplain:

(2)

$$Area\ contribution\ normalized\ by\ agreement\ area =$$

$$\frac{stream\ order\ X\ contribution\ to\ disagreement\ area}{stream\ order\ X\ contribution\ to\ agreement\ area}$$

Normalized difference metric

I developed a metric, normalized difference, to indicate the type of alteration most prevalent in a HUC8 or HUC2 basin.

$$Normalized\ difference = \frac{(Sum\ of\ AD\ area) - (Sum\ of\ \mathrm{AC}\ area)}{(Sum\ of\ AD\ area) + (Sum\ of\ \mathrm{AC}\ area)} \tag{3}$$

This metric ranges from -1, indicating completely artificially flooded, to +1, indicating completely disconnected floodplain. I calculated this metric for each HUC2 basin by determining the mean value of HUC8 basins there.

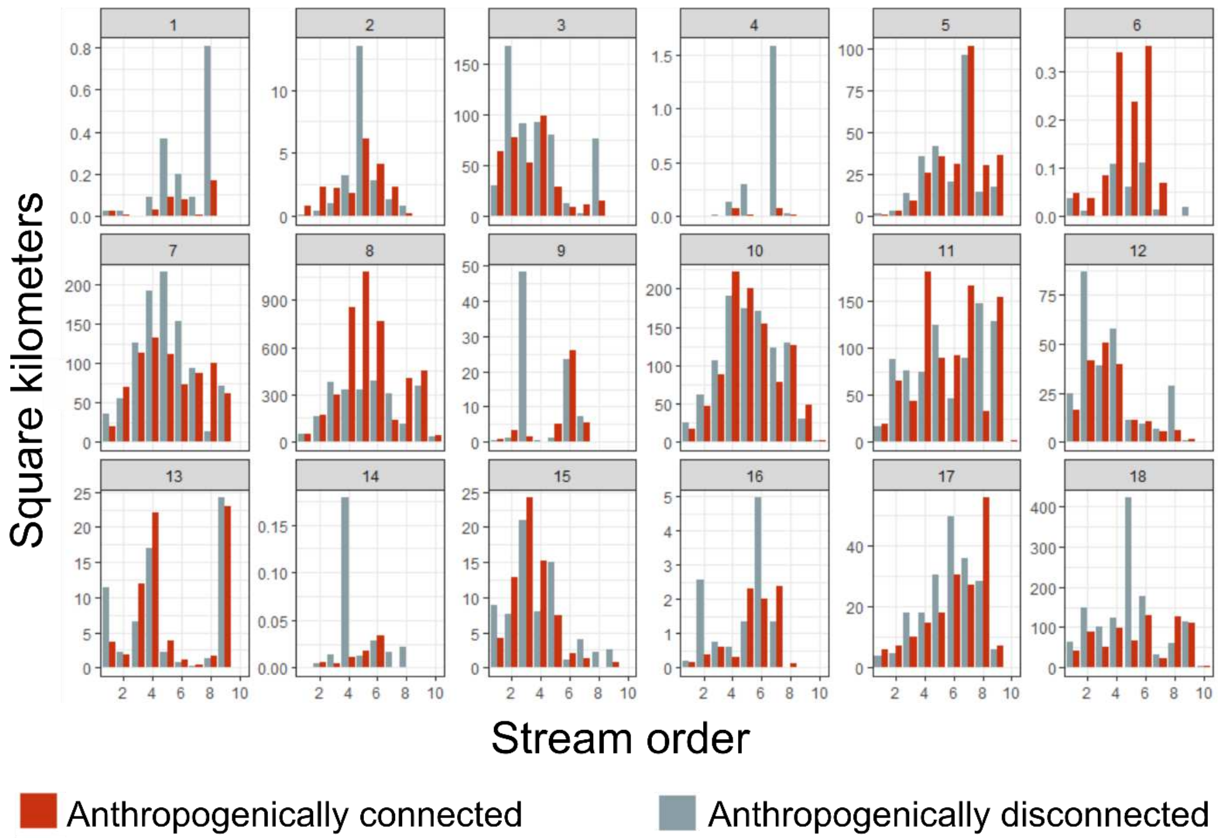


Figure B.1. Area of stream order contribution to anthropogenically connected and disconnected floodplain areas by HUC2 basin.

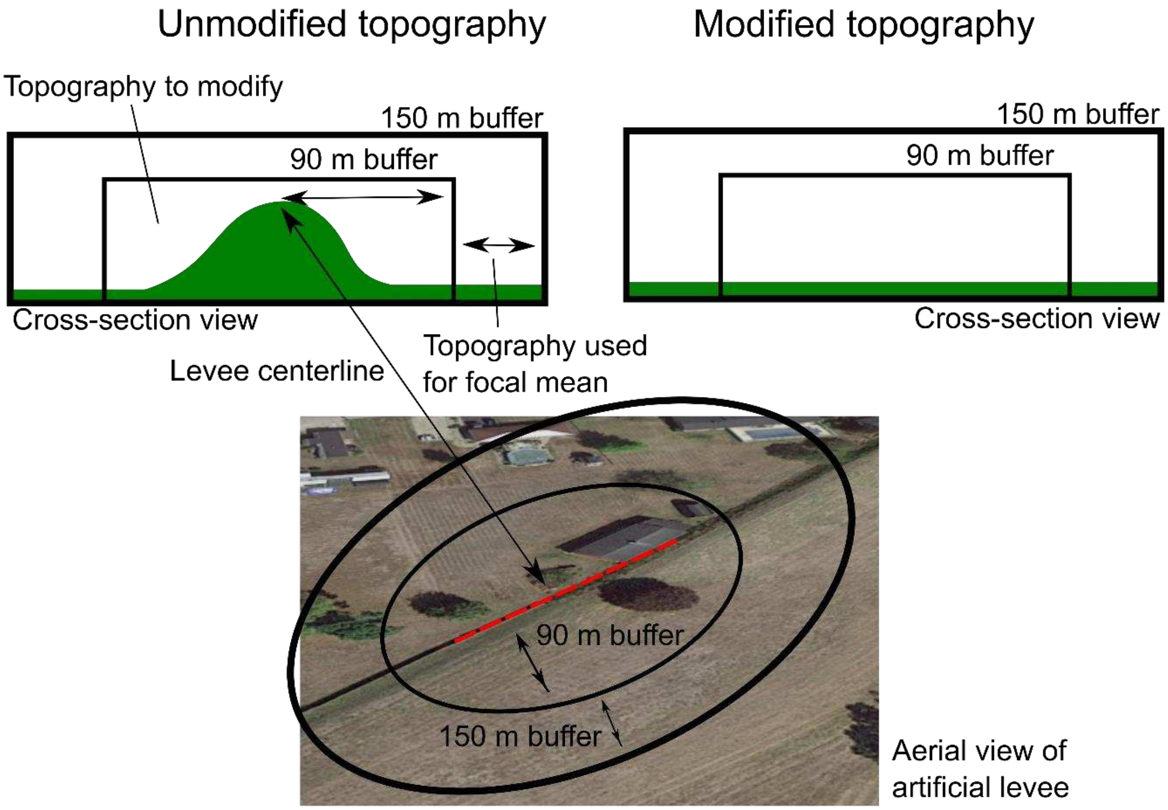


Figure B.2. Topography modification for NLD and potential levees. The topography within 90 m of the levee centerline is modified by applying a focal mean with a 120 m radius using only the topography between the 90 m and 150 m buffers.

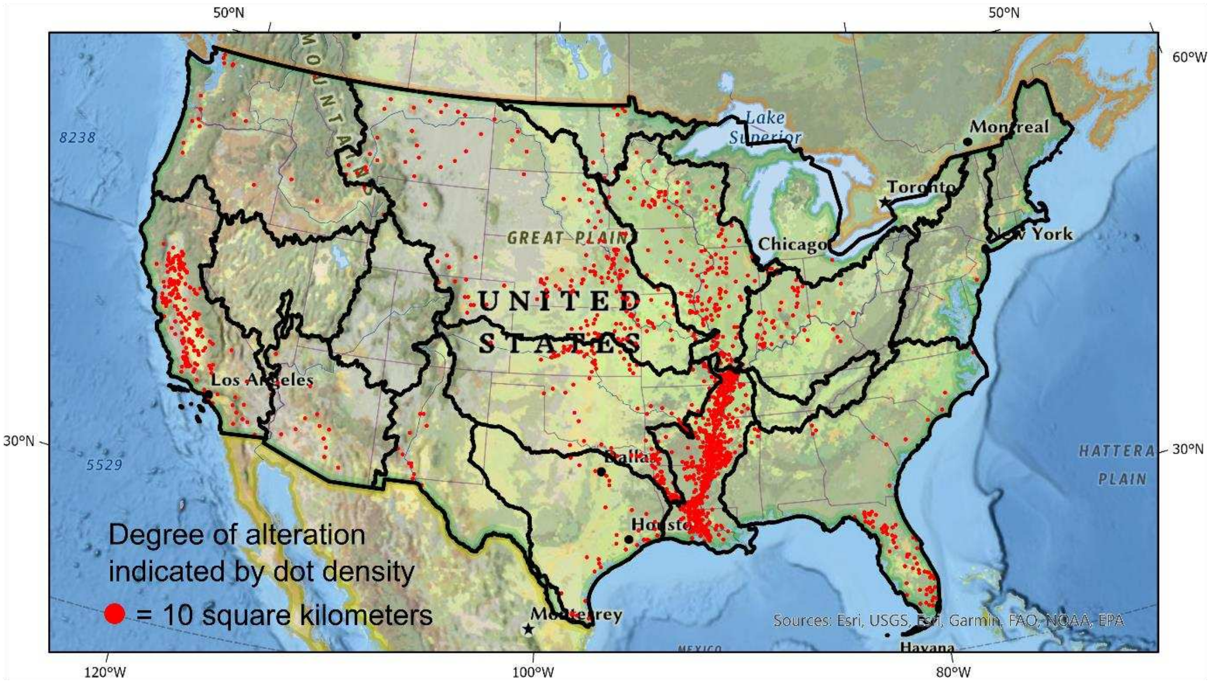


Figure B.3. Cumulative alteration by anthropogenic floodplain disconnection and connection where the degree of alteration is indicated by dot density, with one dot representing 10 square kilometers.

Table B.1. Sources and descriptions of data used in study.

Data	Description	Source
Elevation Derivatives for National Applications (EDNA)-DEM	30-m resolution DEM	Gesch et al. (2002)
National hydrography dataset (NHD) High Resolution	National stream location and orders	Buto et al. (2020)
FEMA Flood maps	"A" and "AE" flood zones indicating the 100 year recurrence interval	https://msc.fema.gov/portal/advanceSearch
1:250,000-scale Hydrologic Units of the United States	2-digit Hydrologic Unit Code (HUC2), Watershed Boundary Dataset units	Seaber et al. (1987)
National levee dataset (NLD)	known locations of artificial levees	https://levees.sec.usace.army.mil/
Potential levees	potential locations of artificial levees not in the NLD	https://doi.org/10.4211/hs.729c0aea00bb48d6b6814c147e4318c4
Land Cover (NLCD)	National Land Cover Database, 2016, 30-m resolution	Jin et al., (2019)

Table B.2. Area in square kilometers of each type of area and the normalized difference by HUC2 basin.

HUC2 basin	**Agreement (km^2)**	**Anthropogenically connected (km^2)**	**Anthropogenically disconnected (km^2)**	**Normalized difference**
New England (1)	13,463	1	2	0.65
Mid-Atlantic (2)	33,647	20	23	0.23
South Atlantic-Gulf (3)	223,063	378	575	0.16
Great Lakes (4)	25,378	10	11	0.18
Ohio (5)	32,939	273	244	0.08
Tennessee (6)	8,850	1	0	0.30
Upper Mississippi (7)	73,917	786	971	0.06
Lower Mississippi (8)	117,658	4,252	2,462	0.00
Souris-Red-Rainy (9)	22,767	44	83	0.07
Missouri (10)	81,588	992	1,024	0.11
Arkansas-White-Red (11)	59,114	850	803	0.08
Texas-Gulf (12)	84,760	190	273	0.09
Rio Grande (13)	29,765	70	68	0.15
Upper Colorado (14)	12,168	0	0	0.36
Lower Colorado (15)	28,366	82	86	0.10
Great Basin (16)	37,210	9	12	0.33
Pacific Northwest (17)	40,731	179	197	0.12
California (18)	31,730	776	1,267	0.16
Total (km^2)	957,113	8,911	8,100	

Table B.3. Max and median slope (in degrees) of anthropogenically connected and disconnected floodplain areas in the Lower Mississippi basin (HUC2 no. 8).

	Anthropogenically connected	**Anthropogenically disconnected**
Max slope per area (mean)	2.3	2.3
Max slope per area (median)	1.02	1.18
Median slope per area (mean)	1.25	1.58
Median slope per area (median)	0.47	0.84

Table B.4. Percent of land use using the 2016 NLCD in anthropogenically connected and disconnected, and agreement floodplains for the CONUS.

Land use	**Anthropogenically connected**	**Anthropogenically disconnected**	**Agreement**
Barren land	0	0	3
Cultivated crops	47	36	18
Hay pasture	8	11	6
Deciduous forest	7	7	4
Evergreen forest	2	6	4
Mixed forest	2	3	1
Developed high intensity	1	1	0
Developed low intensity	3	4	2
Developed medium intensity	2	2	1
Developed open space	5	5	3
Emergent herbaceous wetlands	3	3	7
Woody wetlands	12	8	17
Herbaceous	4	7	7
Open water	2	2	17
Perennial snow ice	0	0	0
Shrub and scrub	2	5	10
Unclassified	0	0	0