

ESTIMATING CHANGES IN FINE-SEDIMENT STORAGE AT EDDY-  
SANDBARS ON THE COLORADO RIVER, GRAND CANYON, AZ  
USING OBLIQUE IMAGERY FROM REMOTE CAMERAS

By Ryan E. Lima

A Dissertation

Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

in Earth Science & Environmental Sustainability

Northern Arizona University

August 2021

Approved:

Temuulen T. Sankey Ph.D., Chair

Paul E. Grams Ph.D.

Erich R. Mueller Ph.D.

Taylor J. Joyal Ph.D.

## ABSTRACT

# ESTIMATING CHANGES IN FINE-SEDIMENT STORAGE AT EDDY-SANDBARS ON THE COLORADO RIVER, GRAND CANYON, AZ USING OBLIQUE IMAGERY FROM REMOTE CAMERAS

RYAN E. LIMA

This research is a fluvial remote sensing study demonstrating methods for sub-annual monitoring via a dataset containing over 1,000,000 ground-based oblique images that capture daily observations at 43 sandbar monitoring sites in the dam-affected Colorado River in Grand Canyon since 1990. Over half of the world's major rivers are affected by dams. In many dam-affected rivers, sediment limited conditions have led to increased erosion of banks and fine-sediment deposits for hundreds of kilometers downstream. Quantifying the short-term rates of erosion and measuring the effect of dam operations and beach-building high flows on sediment storage at sub-annual scales is critical to managing downstream resources effectively. Sandbars in the Grand Canyon provide relatively flat, vegetation-free substrates utilized by nearly 25,000 river runners annually. Sandbars are also essential components of riverine systems creating habitat for native fish and storing sediment which would otherwise be transported downstream. This study is the most comprehensive attempt at quantitative analysis of this dataset. I present methods for estimating sandbar volume and hypsometry from the remote imagery. I demonstrate a deep learning approach to semantic segmentation, which allowed for detailed image-derived sandbar area analysis of over 13,000 images across 10 years at three sites. Significant variability was observed in the sub-annual area change due

to current dam operations. I determined that erosion, deposition, and the resulting mean monthly area at sandbar sites are more closely correlated with antecedent sandbar size than monthly flow metrics. The analysis of time-lapse videos at 41 sites revealed links between daily and seasonal discharge patterns and mass failure rapid erosion events. These insights increase our understanding of the dynamics of fluvial bedforms in dam-affected, canyon-bound rivers and might improve the adaptive management of the Colorado River in the Grand Canyon. These methods could be applied broadly to remote-camera monitoring efforts in many other fluvial and coastal settings for measuring erosion rates and improving, modeling and sediment budgeting efforts.

## ACKNOWLEDGMENTS

Throughout the writing of this dissertation, I have received a great deal of support, guidance, and kindness from individuals whom I would like to acknowledge. I would first like to thank my primary advisor, Temuulen “Teki” Sankey, whose expertise was invaluable during this process. Teki worked tirelessly to keep me supported throughout these five years and provided timely and insightful feedback that pushed me, and my work to a higher level. I would like to thank the other members of my committee, Paul Grams, Erich Mueller, and Taylor Joyal, who generously afforded their time to improve my research questions and methodologies. A special thanks to Paul for allocating resources to keep this project going through all of its ups and downs. Thanks to Dan Buscombe for his time, patience, guidance, and for providing the tools integral to this research.

Thanks to Joe Hazel, and Matt Kaplinski for providing data, filling in gaps in my funding, and for stimulating conversations which inspired and informed my research. Thanks to the entire GCMRC and Namdor communities for their remarkable efforts monitoring these sandbars for over three decades.

In addition, I would like to thank my parents, Fred and Laura for their willingness to read hundreds of pages of my drafts, and offer a sympathetic ear when things got difficult. Thanks to Charles, Hannah, Danny, and countless other friends who helped and supported me in countless ways.

## TABLE OF CONTENTS

TITLE PAGE .....	i
ABSTRACT.....	ii
ACKNOWLEDGMENTS .....	iv
TABLE OF CONTENTS .....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
DEDICATION.....	x
PREFACE .....	xi
CHAPTER 1: INTRODUCTION TO EDDY SANDBAR DYNAMICS.....	1
INTRODUCTION .....	1
BACKGROUND .....	2
1.1.1.  SANDBARS .....	2
1.1.1.1. EDDY-FAN COMPLEX AND GRAND CANYON SANDBARS .....	3
1.1.1.2. SANDBAR EROSION .....	5
1.1.2.  GLEN CANYON DAM.....	6
1.1.2.1. FLOW REGIMES .....	8
1.1.2.2. HIGH FLOW EXPERIMENTS .....	10
1.1.3.  SANDBAR MONITORING .....	12
SUMMARY AND MOTIVATING QUESTIONS.....	14
REFERENCES .....	16
CHAPTER 2: CHAPTER 2: USING OBLIQUE IMAGERY TO MEASURE HYPSONETRIC CHANGES IN SANDBAR VOLUME FOLLOWING CONTROLLED FLOODS IN THE GRAND CANYON.....	23
ABSTRACT.....	23
2.1. INTRODUCTION .....	24
2.2. METHODS .....	27
2.2.1. STUDY AREA.....	27
2.2.2. IMAGE-DERIVED AREA MEASUREMENT.....	28
2.2.2.1. REGISTRATION .....	28
2.2.2.2. RECTIFICATION.....	28
2.2.2.3. SEGMENTATION.....	28
2.2.2.4. AREA MEASUREMENT .....	29
2.2.2.5. ESTIMATING WATER-SURFACE ELEVATION .....	29
2.2.3. VOLUME FROM IMAGE-DERIVED DEMS .....	29
2.2.4. HYPSONETRIC ANALYSIS.....	30
2.3. RESULTS AND DISCUSSION .....	32
2.3.1. SURVEYED HYPSONETRIES .....	32
2.3.2. HYPSONETRICALLY-DERIVED VOLUME .....	33
2.3.3. ACCURACY OF IMAGE WATER-SURFACE ELEVATIONS.....	35
2.3.4. VOLUME FROM IMAGE-DERIVED DEMS .....	36
2.3.5. LIMITATIONS .....	38
2.4. CONCLUSIONS.....	39

2.5. REFERENCES .....	40
CHAPTER 3: SEMANTIC SEGMENTATION OF TIME-SERIES IMAGERY USING CONVOLUTIONAL NEURAL NETWORKS: A CASE STUDY OF SANDBARS IN GRAND CANYON .....	43
ABSTRACT .....	44
3.1. INTRODUCTION .....	45
3.2. DATA .....	49
3.3. METHODS .....	52
3.3.1. IMAGE PRE-PROCESSING .....	52
3.3.2. REMOVING DISTORTION .....	52
3.3.2.1. RECTIFICATION .....	53
3.3.2.2. REGISTRATION .....	54
3.3.2.3. LABELING IMAGERY .....	55
3.3.3. CNN SEMANTIC SEGMENTATION .....	56
3.3.3.1. NEURAL NETWORK ARCHITECTURES .....	56
3.3.3.2. HYPERPARAMETERS .....	58
3.3.3.3. IMAGE AUGMENTATION .....	60
3.3.3.4. TRAINING, TESTING, AND VALIDATION DATA .....	61
3.3.3.5. EXPERIMENTS .....	63
3.3.3.6. COMPUTING HARDWARE AND SOFTWARE .....	63
3.4. RESULTS .....	64
3.4.1. OPTIMIZER AND LOSS FUNCTION .....	64
3.4.2. ARCHITECTURE PERFORMANCE .....	64
3.4.3. IMAGE AUGMENTATION .....	66
3.4.4. EVALUATING TRAINING SAMPLE SIZE .....	66
3.5. DISCUSSION AND CONCLUSIONS .....	67
3.6. ACKNOWLEDGEMENTS, SAMPLES, AND DATA .....	71
3.7. REFERENCES .....	72
CHAPTER 4: MEASURING EDDY SANDBAR SUB-ANNUAL EROSION AND DEPOSITION PATTERNS USING OBLIQUE IMAGE TIME-SERIES OF THE COLORADO RIVER, GRAND CANYON .....	78
4.1. ABSTRACT .....	78
4.2. INTRODUCTION .....	79
4.2.1. REGIONAL SETTING AND SITE DESCRIPTIONS .....	82
4.2.1.1. GRAND CANYON SANDBARS .....	83
4.2.2. FLOW REGIMES AND CONTROLLED FLOODS .....	85
4.2.3. SANDBAR BEHAVIOR .....	87
4.2.3.1. SANDBAR EROSION .....	88
4.2.3.2. OBSERVATIONS FROM PAST SANDBAR STUDIES .....	91
4.3. MATERIALS AND METHODS .....	94
4.3.1. MEASURING CHANGES IN SANDBAR SIZE .....	94
4.3.1.1. ESTIMATING DISCHARGE .....	95
4.3.1.2. CNN IMAGE SEGMENTATION .....	97
4.3.1.3. DAILY AND MONTHLY SANDBAR AREA .....	98
4.3.1.4. ESTIMATING EROSION AND DEPOSITION MAGNITUDE .....	98
4.3.1.5. DISCHARGE PATTERNS .....	99
4.3.1.6. APPLICATION-SPECIFIC CONSIDERATIONS .....	100
4.3.2. MASS FAILURE TIME-LAPSE ANALYSIS .....	101
4.4. RESULTS .....	102

4.4.1. ACCURACY OF IMAGE-DERIVED AREA .....	102
4.4.2. IMAGE-AREA EROSION AND DEPOSITION RATES AT 22-MILE, 30-MILE, AND 145-MILE SANDBARS .....	105
4.4.2.1. TIME-SERIES OF NET AREA CHANGE .....	105
4.4.2.2. RESPONSE TO CONTROLLED FLOODS .....	107
4.4.2.3. MONTHLY NET AREA CHANGES AND MONTHLY FLOW .....	109
4.4.3. MASS FAILURES IN TIME-LAPSE ANALYSIS .....	111
4.4.3.1. MASS FAILURES AT 30-MILE AND 145-MILE SANDBARS .....	116
4.5. DISCUSSION .....	116
4.5.1. IMAGE-AREA AND SPATIAL VARIABILITY AMONG SANDBARS.....	116
4.5.2. MASS FAILURES AT 22-MILE, 30-MILE, AND 145-MILE SANDBARS.....	120
4.5.3. TIME-LAPSE MASS FAILURE ANALYSIS .....	122
4.6. CONCLUSIONS .....	126
4.7. ACKNOWLEDGEMENTS .....	127
4.8. REFERENCES .....	128
CHAPTER 5: CONCLUSIONS AND MANAGEMENT IMPLICATIONS.....	136
5.1. INTRODUCTION .....	136
5.2. RESEARCH QUESTIONS AND KEY FINDINGS .....	136
5.3. MANAGEMENT IMPLICATIONS .....	139
5.4. REFERENCES .....	141
APPENDIX .....	143
UNITS AND ACRONYMS .....	143
DATA AVAILABILITY AND ASSOCIATED CODE.....	146

## LIST OF TABLES

<b>Table 2.1</b> Coefficients of determination and root-mean-squared error (RMSE) for linear models relating surveyed sandbar volume.....	34
<b>Table 3.1</b> Selected Accuracies for CNNs trained on various train and test sets.....	67
<b>Table 4.1</b> Discharge metrics calculated for daily and monthly flow periods.....	99
<b>Table 4.2</b> Variability in 283m3s-1 area at three timescales: Annual, monthly, and daily.....	107
<b>Table 4.4</b> Number of observed mass failures in time-lapse videos and length of the time-lapse record.....	112

## LIST OF FIGURES

<b>Figure 1.1</b> Flow patterns and configuration of the eddy-fan complex.....	5
<b>Figure 1.2</b> Hydrograph of different flow regimes.....	8
<b>Figure 1.3</b> Annual hydrographs of daily discharge from Lees Ferry gage 09380000 for a randomly selected year within each of the four flow regimes discussed.....	9
<b>Figure 2.1</b> The study area.....	27
<b>Figure 2.2</b> Relationships between the area and elevation plotted in three ways.....	32
<b>Figure 2.3</b> Relationship between survey-derived area measurements scaled to the ref. elevation.....	33
<b>Figure 2.4</b> Comparison between estimated water-surface elevation and surveyed water-surface elevations.....	36
<b>Figure 2.5</b> Comparison between survey-derived volumes above the reference elevation (blue dots) and volumes estimated from image-derived DEMs (red line).....	37
<b>Figure 2.6</b> Comparison between survey-derived volumes above the reference elevation and image-derived volumes for four corresponding dates. ....	38
<b>Figure 2.7</b> (a) Image frequency at RM30.7 by discharge. (b) Image times plotted over the estimated hydrograph for RM30.7 .....	39
<b>Figure 3.1</b> Map of the study area in Grand Canyon National Park.....	51
<b>Figure 3.2</b> Image processing steps. ....	54
<b>Figure 3.3</b> A simplified diagram of UNet-1 showing the U-shaped structure,.....	58
<b>Figure 3.4</b> Examples of image augmentation performed on a rectified image from the 30-mile site. ....	60
<b>Figure 3.5</b> <i>F1</i> scores across five folds of the one-site 100 train and validation set. ....	65
<b>Figure 3.6</b> Accuracy of all five folds of Res-UNet CNNs. ....	66
<b>Figure 4.1</b> Study Area and the Grand Canyon National Park.....	83
<b>Figure 4.2</b> Annual (top) and daily (bottom) hydrographs. ....	86
<b>Figure 4.3</b> Diagram of flow zones on a conceptualized Grand Canyon sandbar.....	88
<b>Figure 4.4</b> An example of a mass failure at the 41-mile sandbar.....	89
<b>Figure 4.5</b> Correlations between image-areas and survey-areas at the three selected sites. ....	103
<b>Figure 4.6</b> A comparison of image-derived normalized sandbar area from annual, monthly, and daily time-series. ....	106
<b>Figure 4.7</b> Standardized mean monthly sandbar area. ....	108
<b>Figure 4.8</b> Time-series of image-derived mean monthly area at 22-mile, 30-mile and 145-mile. ....	109
<b>Figure 4.9</b> Correlation matrices for 22-mile, 30-mile, and 145-mile.....	110
<b>Figure 4.10</b> Discharge patterns and mass failure frequency in the Grand Canyon.....	114
<b>Figure 4.11</b> Flow metrics on failure days (F), lag2 days (L2), and randomly selected days (R). ....	115
<b>Figure 4.12</b> Imagery of 22-mile (A-C), 30-mile (D-F), and 145-mile (G-I).....	118
<b>Figure 4.13</b> Hypsometric curves showing 22-mile, 30-mile, and 145-mile surveys .....	121

## DEDICATION

This dissertation is dedicated to my community of loving family and friends, but most of all to my parents Fred and Laura Lima. Their patience, encouragement, and unconditional support got me through the most difficult parts of this process.

*“Nothing of significance was ever achieved by an individual acting alone.*

*Look below the surface and you will find that all seemingly solo acts*

*are really team efforts.” – John C. Maxwell*

## PREFACE

This dissertation is comprised of an introduction chapter (1), three manuscripts targeted for specific journals or proceedings, (Chapters 2-4), and a conclusion chapter (Chapter 5). Chapter 2 is formatted for and was published in the Proceedings of the Federal Interagency Sedimentation and Hydrologic Modeling Conference (2019). Chapter 3 was formatted for publication in the journal Water Resources Research, to which it will be submitted in short order. Chapter 4 is formatted for upcoming submission to the journal Geomorphology. Since these chapters were prepared as stand-alone manuscripts, there is some redundancy, which results from combining them to meet the university formatting requirements.

Citation for Chapter 2:

Lima, R.E., Buscombe, D., Sankey, T., Grams, P.E., Mueller, E.R., 2019. Using Oblique Imagery to Measure Hypsometric Changes in Sandbar Volume Following Controlled Floods in the Grand Canyon, in: Proceedings of the Federal Interagency Sedimentation and Hydrologic Modeling Conference. Reno, NV.

# CHAPTER 1: INTRODUCTION TO EDDY SANDBAR DYNAMICS

## INTRODUCTION

Sandbars are prominent features along the Colorado River within Grand Canyon National Park that provide substrate for riparian vegetation, create backwater habitat for native fish species, serve as a source of aeolian sand, and provide beaches utilized by tens of thousands of recreationalists each year (Kearsley et al., 1994; Valdez et al., 1995, Converse et al., 1998; Stevens et al., 2001; Sankey and Draut, 2004; East et al., 2016). Monitoring the dynamics of these sandbars in response to changes in sediment supply and discharge patterns resulting from the closure and operation of Glen Canyon Dam is of great interest to dam managers charged with balancing societal needs for water storage and power generation with the health of the downstream ecosystem and the provisioning of ecosystem services (Wright et al., 2005). This research utilizes remote sensing, a process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation with sensors, to improve sandbar monitoring (Lillesand et al., 2004). In the following chapters, I explore various ways of utilizing a dataset of oblique images captured by remote cameras to measure changes in sandbar size and reveal linkages between discharge patterns and morphodynamics at a sample of 41 sandbar monitoring sites spanning 362 *km* between Lees Ferry and Diamond Creek in the Grand Canyon. This dataset, which contains over 1,000,000 images, poses significant challenges from a remote sensing methodological standpoint, but also provides a valuable opportunity to quantify sandbar dynamics occurring at finer temporal resolutions than ever before.

In the following chapters, I explore methods for measuring changes in the elevational storage of sand using image-derived hypsometric curves and experiments with volume estimation by combining sandbar area measurements from multiple images. I utilize a deep learning approach to rapidly segment imagery and measure the planar area of the emergent sandbars in thousands of images. I investigate cyclic patterns of area change and the occurrence of mass failure erosional events using time-series of image-derived area estimates, and time-lapse videos. Finally, I attempt to relate observed area change with daily and monthly discharge patterns to provide new insights into the effects of dam operations on sandbars. Such insights increase our understanding of sediment dynamics in dam-affected, canyon bound rivers and may improve the adaptive management of the Colorado River in the Grand Canyon. This research explores the possibilities and limitations of one network of remote cameras but can also inform others of the utility of such datasets in other fluvial and coastal settings. This chapter will introduce the concepts and historical context needed to understand sandbar dynamics in the Grand Canyon. I discuss findings from previous studies and illuminate gaps in our current understanding. Finally, I summarize the research questions explored in subsequent chapters.

## **BACKGROUND**

### **1.1.1. Sandbars**

Fluvial sandbars are elongate bedforms comprised of sediment transported by river or streams; these bars may occur within channels or at channel margins. While some fluvial sandbars may migrate along the bed, within the channel, or as point bars on the inside of meander bends in alluvial rivers, other fluvial bars persist as relatively permanent features

(Bierman and Montgomery, 2014). Persistent fluvial bars are common in fan-affected canyon bound rivers with bedrock channels, like those found in the Grand Canyon (Baker, 1984).

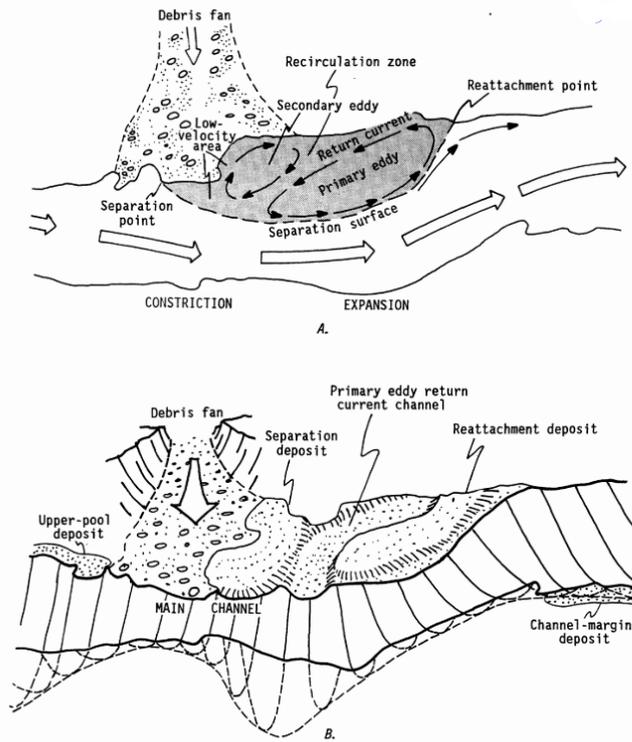
Rivers with bedrock channels generally have steeper gradients, and larger stage increases relative to discharge than in alluvial rivers. These factors lead to high shear stresses resulting in high sediment transport rates and coarser bed materials (Baker, 1984; Tinkler and Wohl, 1998). During high flows, a fraction of the fine-sediment ( $< 2mm$ ) which would otherwise be transported downstream, is deposited in zones of local low-velocity flow in the lee of large boulders or channel constrictions. Because of rapid changes in stage, high flows can deposit fine-sediment at relatively high elevations, where it becomes less vulnerable to erosion by more frequent, smaller magnitude flows (Tinkler and Wohl, 1998). Therefore, sandbars are significant components of the fluvial ecosystem. These persistent fluvial bars or beaches are referred to hereafter as sandbars.

#### ***1.1.1.1. Eddy-fan complex and Grand Canyon sandbars***

The Colorado River drops 670m between Lees Ferry and Lake Mead, most of which occurs in the span of 161 major rapids (Leopold, 1969; Dolan et al., 1974). These major rapids occur where debris flows from tributary side canyons have created debris fans, which act as hydrologic controls and constrict the flow of the river by as much as 50% (Kieffer, 1985; Webb et al., 1989). Constrictions result in shooting flow, or increased velocity of the main channel current. Once past the apex of the debris fan, the main current becomes separated laterally from the bank before reattaching downstream (Webb et al., 1989). This separation creates a zone of low-velocity recirculating water between the primary current and the bank downstream of the fan. Such zones may contain one or more eddies (Schmidt and Graf, 1990). Suspended sediment advected into these zones of recirculating water may be deposited forming sandbars.

Sandbars in Grand Canyon are primarily composed of sand with smaller amounts of gravel, silt and clay (Howard and Dolan, 1981). Sandbar deposits may exceed 13m in vertical thickness, but average about 4m in thickness, and are often mantled by boulders or bedrock (Rubin et al., 1994). Sandbars in the Grand Canyon are often characterized by their depositional environment. Reattachment bars form as spits of sand projecting out from the bank where the river's primary current reattaches to the bank and decelerates (Rubin et al., 1990). A second eddy may occur upstream of the reattachment point leading to deposition on the downstream end of the debris fan near the point, where the current separates from the bank. These deposits are known as separation bars (Schmidt and Graf, 1990; Andrews et al., 1999). At higher flows, the upstream flow may carve out a return-current channel, between the reattachment bar and the bank, which serves as a vital backwater habitat at lower flows (Andrews et al., 1999). In some circumstances, the surface of separation and reattachment deposits may merge into a large eddy deposit, which cannot be distinguished from one or the other (Schmidt and Graf, 1990). Sandbars, that form in the slack water upstream of a debris fan, are known as upper pool bars (Schmidt and Graf, 1990).

Together, these three sandbar types along with the debris fan and flow separation zone comprise what is known as the eddy-fan complex (*sensu* Schmidt and Rubin, 1995)(Figure 1.1). Much of the exposed sand along the Colorado River within the Grand Canyon is found within eddy-fan complexes. In the debris fan-affected sections of the Green River, extending 169 km downstream of Flaming Gorge Dam, 64% of the fine-sediment resides in eddy fan complexes (Grams and Schmidt, 1999). Less common depositional environments within the Grand Canyon include point bars, on the inside of meander bends, and channel margin deposits found in the lee of small boulders or bedrock outcrops along the banks (Schmidt and Graf, 1990).



**Figure 1.1** Flow patterns and configuration of the eddy-fan complex. Figure from Schmidt and Graf (1990). (A) Flow patterns. (B) Configuration of bed deposits.

### 1.1.1.2. Sandbar erosion

Erosion is a major cause of morphodynamic change of sandbars. Sandbars can be thought of as discrete alluvial banks. Bank erosion models focus on the interplay between hydraulic forces and gravitational failures (Simon, 1989; Simon et al., 1999). Hydraulic forces involve the energy imparted by moving water and gravitational failures occur as the force of gravity overcomes the mechanical strength of banks (O'Neill and Kuhns, 1994). Fluvial erosion of sandbars in Grand Canyon results from several interrelated mechanisms including tractive scour, wave-action, seepage-driven erosion, gravitational failures, static liquefaction, and breaching (Schmidt and Graf, 1990; Beus and Avery, 1992; Bauer and Schmidt, 1993; Van den Berg, 2002; Wright and Kaplinski, 2011; Mastbergen et al., 2019). Sandbar characteristics such as grain size, porosity, pore pressure, vegetation cover, and grain

mineralogy affect the mechanical strength of the banks (Fischenich and Allen, 2000; Travis et al., 2010). Bank slope also affects susceptibility to different erosional mechanisms (Alvarez and Schmeckle, 2013). Sandbar erosion typically proceeds gradually, resulting in meters of lateral erosion over weeks to months, but can lead to more rapid erosion through mass failure, where meters of lateral erosion occur in minutes to hours under certain conditions.

### **1.1.2. Glen Canyon Dam**

Glen Canyon Dam began impounding the Colorado River in 1963 and was constructed because of the 1956 Colorado River Storage Project Act (U.S. Department of Interior, 1970). Prior to its closure, sediment from the upper basin, and sediment brought in by tributary side canyons would accumulate on the bed of the Colorado River during the late summer, fall, and winter (Topping et al., 2000a; Wright et al., 2005). Other than short-duration flash floods associated with the North American Monsoon, the river's flow was relatively low during these times, with discharges as low as  $71\text{m}^3\text{s}^{-1}$  (National Resource Council, 1996). In the springtime, the river would swell with annual snowmelt from the Central and Southern Rockies. Flows during the spring sometimes exceeded  $2,831\text{m}^3\text{s}^{-1}$  (National Resource Council, 1996). Such floods would scour the bed and carry most of the sediment downstream, but a significant amount of sediment would be deposited in hundreds of eddies as subaerial open sandbars (Topping et al., 2000a; Wright et al., 2005). Annual high floods prevented riparian vegetation from colonizing large portions of these sand deposits below the average yearly high-water mark (Dolan et al., 1974; Stevens et al., 2001).

Since 1963, Glen Canyon Dam and its operation have led to several significant changes with regards to sediment inputs and flow on the Colorado River downstream. The dam stops sediment from the Colorado's upper basin from entering the Grand Canyon. There was considerable uncertainty and variability in the pre-dam sediment load carried by the

Colorado River in any given year, but Wright et al. (2005) estimated that an average of 25 million tons of sand flowed into the Grand Canyon from the upper basin annually.

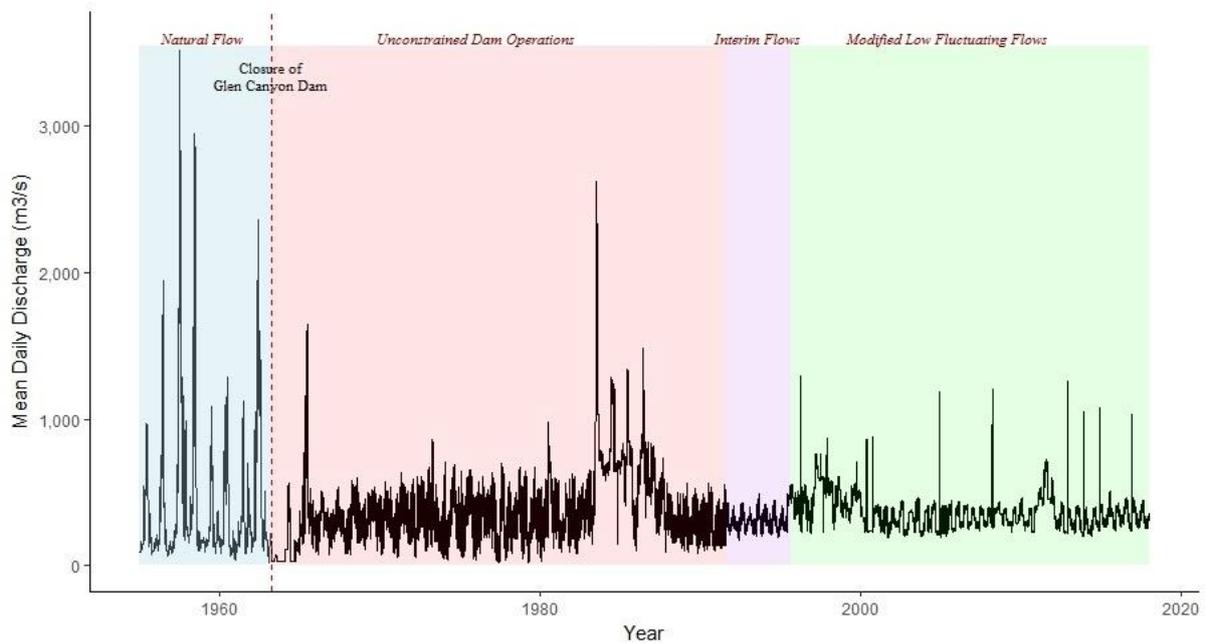
Downstream of the dam, the Paria River and the Little Colorado River are the primary sources of fine-sediment to the canyon. They supply, on average, a combined 3.6 million tons of sand annually or between 15% and 20% of the pre-dam average (Topping et al., 2000b, Wright et al., 2005).

Coupled with a reduced sediment supply, significant changes to the character of the river's flow regime have resulted from dam operations. Annual snowmelt-driven floods have been eliminated as well as the typical late summer, winter, and fall low-flows. High annual flow variability and low daily variability have been replaced by large daily changes in flow and higher median flows. Under this new hydrologic regime, the Colorado River is sediment-limited with respect to sand throughout the year, and no substantial period of sediment accumulation occurs in a typical year (Topping et al., 2000a; Rubin et al., 2002).

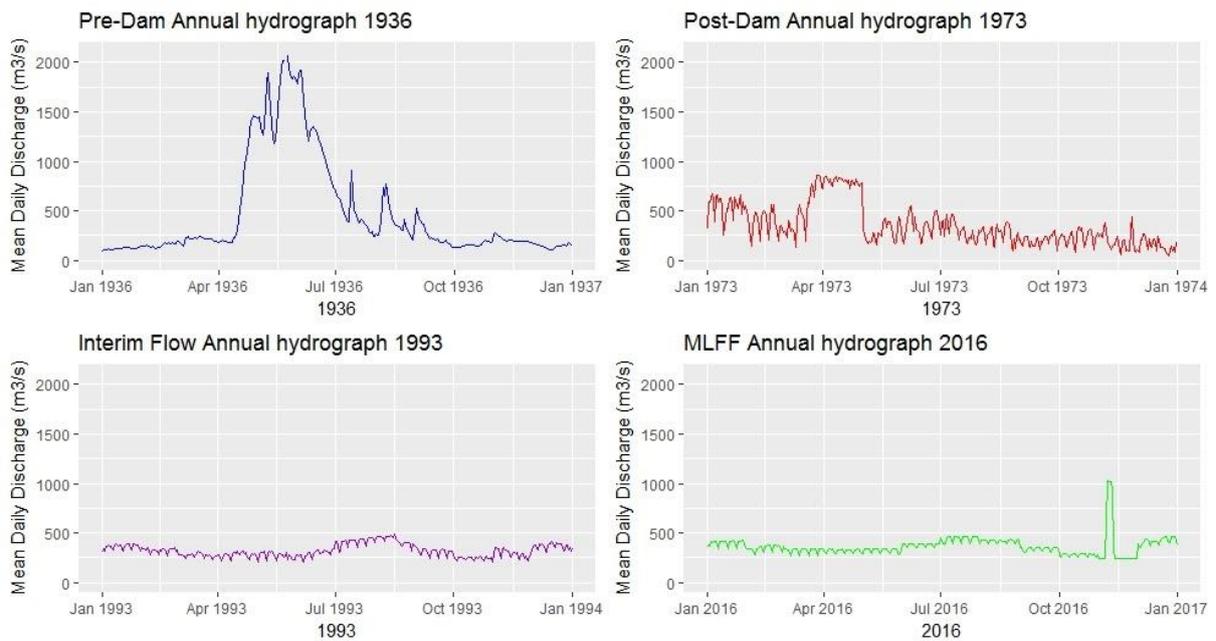
The impacts of Glen Canyon Dam to sediment resources in the canyon were first described in the scientific literature, just over a decade after its closure, by Dolan et al. (1974). Dolan and colleagues describe the widespread erosion of beaches used for camping and the encroachment of vegetation on subaerial, or emergent, sandbars in areas once scoured by large floods. This reduction in areas suitable for camping was occurring just as the use of the river was on the rise. In the early 1950's, approximately 200 people had taken a boat through the Grand Canyon, but by the 1970's more than 10,000 people per year were participating in trips down the canyon and the threat to burgeoning recreation in the canyon led to management interest in the Grand Canyon's sandbars (Dolan et al., 1974).

### 1.1.2.1. Flow regimes

Understanding the effect of dam operations on sandbars requires a closer look at the discharge patterns, which result from the rules guiding the operation of Glen Canyon Dam since its closure. For this study, I delineate four distinct flow regimes: (1) Natural Flow (NF): this is the pre-dam period ending in March 1963; (2) Unconstrained Dam Operations (UDO): the period between 1964 and August 1991; (3) Interim Flows (IF): the flows between August 1991 and October 1996; and (4) Modified Low Fluctuating Flows (MLFF): the flows occurring since the 1996 record of decision went into effect. These delineations are based on the primary dam operations at the time and the characteristics of the resulting flows. It should be noted that there was a period of experimentation from 1990-1991. I combine that period in the UDO because its hydrograph more closely resembles that of the UDO regime. The differences between these periods are illustrated in Figures 1.2 and 1.3.



**Figure 1.2** Average daily discharge in  $m^3 s^{-1}$  from Lees Ferry gage 09380000 between 1955-2018 showing the four flow regimes: Natural Flow (NF), Unconstrained Dam Operations (UDO), Interim Flows (IF) and Modified Low Fluctuating Flows (MLFF).



**Figure 1.3** Annual hydrographs of daily discharge from Lees Ferry gage 09380000 for a randomly selected year within each of the four flow regimes discussed.

During UDO, discharge and the resulting hydrograph was shaped primarily by energy demand and concerns about meeting water delivery obligations (National Resource Council, 1996; Topping et al., 2000a; Wright et al., 2005). Daily maximum flows approached  $994m^3s^{-1}$  and minimum daily flows were as low as  $29m^3s^{-1}$  in the winter and  $85m^3s^{-1}$  in the summer. The median range of daily discharges during UDO was greater than  $227m^3s^{-1}$ . This high daily variability and specifically rapid down-ramping rates were cited as drivers of seepage erosion, creating mass-wasting events, or bank failures (Budhu and Gobin, 1994; Budhu and Gobin, 1995; Alvarez and Schmeckle, 2013).

In 1991 scientists and dam managers began implementing IF. These changes reduced the daily range of discharge to below  $284m^3s^{-1}$ , reduced the maximum daily flow to  $556m^3s^{-1}$ , and increased the minimum flows to  $227m^3s^{-1}$  during the day and  $142m^3s^{-1}$  at night. IF also changed the ramp rates up and down. These temporary changes were intended to reduce erosion of sandbars until an EIS could be completed and new flow guidelines could be agreed upon (LaGory et al., 1993).

The 1995 Environmental Impact Statement, and the subsequent 1996 record of decision led to the adoption of MLFF and the approval of the Glen Canyon Dam Adaptive Management Plan. The shift from IF to MLFF increased the maximum allowable daily flows to  $708m^3s^{-1}$  and increased the allowable up-ramp rate from  $71m^3s^{-1}$  per hour to  $113m^3s^{-1}$  per hour, but the biggest change between IF and MLFF was the addition of beach-building flows, or high flow experiments (HFEs).

#### ***1.1.2.2. High Flow Experiments***

During the 1990s, the scientific consensus was that fine-sediment from tributary canyons was accumulating on the bed during normal dam operations and that high flows were needed to redistribute sand from the bed to high elevation eddy deposits (National Resource Council, 1996). Controlled floods, or High Flow Experiments (HFEs), are short-duration flows more than power plant capacity ( $940m^3s^{-1}$ ). The first HFE occurred from March 26th - April 7th, 1996. Discharges reached  $1274m^3s^{-1}$  for seven consecutive days (Schmidt and Grams, 2011).

When the waters had receded, the HFE had proved successful at redistributing sediment to high elevation sandbars, but, at least in Marble Canyon, the stretch of river between Lees Ferry and the Little Colorado River, most of the sand deposited as high-elevation bars was scoured from lower elevation areas in the eddies, and not from the bed. Scientists realized that they had been working on the incorrect hypothesis-- that sediment was accumulating on the bed of Grand Canyon during normal dam operations, poised to be mobilized by high flows (Schmidt, 1999; Rubin et al., 2002).

Topping et al. (2000a) found that fine-sediment deposits in the bed showed a trend of coarsening upward with regards to grain size, a trend evident throughout the year, suggesting that the Grand Canyon is sediment limited with respect to sand, and no substantial

accumulation was occurring. Sediment transport data collected between 1999 and 2004 shows that when discharges are above about  $255\text{ m}^3\text{ s}^{-1}$  to  $283\text{ m}^3\text{ s}^{-1}$  for an extended period, both fine-sediment stored in the bed and recent inputs from tributaries are rapidly transported downstream (Topping et al., 2000b; Lovich and Melis, 2007). Since the closure of Glen Canyon Dam, flows have exceeded  $255\text{ m}^3\text{ s}^{-1}$  most of the time. This reiterates the importance of sandbars as locations of high elevation storage for fine-sediment, where it is less vulnerable to the erosive flows of the main channel.

Schmidt (1999), in his synthesis of the effects of the 1996 HFE, suggests that future HFEs be timed to follow high sediment inputs from tributaries before normal flows can export that sediment out of the canyon. The 1996 HFE showed that while the mean volume of sediment in high-elevation eddy deposits increased across the canyon, there was a fair amount of variability at individual sites, with some sites experiencing decreases in sediment storage in high-elevation deposits (Schmidt, 1999). Following the 1996 HFE, two more were conducted in November 2004 and March 2008. These were timed to follow shortly after flash flooding of tributary canyons when antecedent storage of sand in upper Marble Canyon was high (Schmidt and Grams, 2011). These HFEs were shorter in duration, just 60 hours each (Schmidt and Grams, 2011). Wiele and Torizzo (2005) found that the highest rates of deposition in eddy sandbars occurred during the first day of an HFE. Schmidt (1999) suggested that exposing eddy sand deposits to high flows for shorter periods would reduce the risk of erosion while providing enough fine sediment deposition to increase sandbar area. A complete summary of the effects of 1996, 2004, and 2008 HFEs can be found in Melis et al. (2011).

In 2011, the Bureau of Reclamation released an Environmental Assessment creating a protocol for HFEs for a ten-year period from 2012 - 2020 (U.S. Department of the Interior, 2011). Remotely operated gages throughout the canyon provide information to create

sediment budgets which inform dam managers about sediment concentrations (Schmidt and Grams, 2011). The HFE protocol requires that antecedent sand enrichment reaches a predetermined level before an HFE can be triggered. HFEs are also limited to  $1,274m^3s^{-1}$ , the total amount of water that can pass through the generators and bypass tubes. Such releases can last for up to 96 hours (Grams et al., 2015; Schmidt and Grams, 2011; U.S Department of the Interior, 2011).

This protocol was used to conduct HFEs in 2012, 2013, 2014, 2016, and 2018. Collectively, annual topographic surveys have shown that HFEs do increase sediment storage at high elevation eddy deposits across the river, but such gains are often eroded away in the months following HFEs (Hazel et al., 2010; Grams et al., 2015; Hazel et al., *in press*). This generalization of overall sandbar size increase obscures the complexity of sandbar responses to HFEs. There is a large amount of variability in the response of specific sandbars to a particular HFE, both across the longitudinal profile of the canyon and to bars within the same reach. How any particular sandbar responds to high flows is dependent on its particular channel geometry, the flow patterns that result, the accommodation space available within the eddy to store sediment (antecedent conditions), and the amount of suspended sediment during a particular high flow (Schmidt and Grams, 2011; Grams et al., 2013).

### **1.1.3. Sandbar monitoring**

The most accurate data we have on the dynamics of Grand Canyon sandbars comes from topographic surveys conducted annually since 1990 at up to 44 sites, and reach-based bathymetric surveys occurring once every three to five years (Hazel et al., 2008; Kaplinski et al., 2009; Hazel et al., *in press*). However, these annual surveys provide only a yearly snapshot of the net change over 12 months. Researchers using remotely operated cameras found that annual surveys likely underestimate lateral erosion rates, and that lateral changes

occurring over a single day may be greater than or equal in magnitude to the changes occurring in an entire year. Cycles of erosion followed by deposition obscure the signal of these frequent high magnitude events (Dexter and Cluer, 1999). Dexter and Cluer (1999) used remote camera imagery to measure lateral erosion, however they only measured maximum lateral erosion rates (see also Cluer 1995; Dexter et al., 1995; Dexter and Cluer, 1996). Grams et al. (2018) utilized imagery to measure changes in sandbar area at a single discharge for the sandbar located 30.7 river miles (RM) downstream from Lees Ferry, RM30.7, discontinuously from 2009 to 2016. However, neither of these approaches enables quantitative and comprehensive monitoring of the nuance of sub-annual sandbar morphodynamics.

An analysis of topographic and bathymetric data of sandbars from RM-0 to RM-225 found that much of the spatial variability of decadal trends in sandbar size and response to HFEs can be explained by a classification scheme that used principal component analysis to classify bars by their geomorphic setting and annual trends in size (Mueller et al., 2018). Mueller (2018) classifies sandbars into 4 groups: (1) subaerial reattachment bars, (2) subaerial separation bars, which tend to have steep reach-averaged channel gradients, (3) subaerial upper-pool bars, typically in low energy environments, and (4) eddy-deposits, which generally contain a small subaerial separation bar and a reattachment bar that is typically submerged or subaqueous. The sandbars in group (1) are further subdivided into three groups (1a, 1b, and 1c), which correlate with changes in the extent of vegetation cover, stage change, and channel width. Relatively open, unvegetated bars in narrow reaches of the canyon are designated 1a, and mostly vegetated bars in wider reaches designated 1c, with 1b bars falling somewhere in the middle (Mueller et al., 2018). Such groups explain some of the spatial variability between sites with regards to long-term trends or HFE response, but significant differences in sub-annual variability and behavior exist, which may not be fully

explained by these groupings (Hazel et al., *in press*). The topographic changes observed at individual sites are influenced by a complex combination of factors including eddy size, antecedent storage condition, grain size, distance downstream from GCD, discharge, suspended sediment concentration, and local hydraulics (Beus and Avery, 1992; Wright and Kaplinski, 2011; Grams et al., 2013)

## **SUMMARY AND MOTIVATING QUESTIONS**

Few places in the world have garnered as much scientific inquiry as the Colorado River in Grand Canyon, yet the remoteness of the canyon makes frequent in-person observations infeasible. Acquisition of satellite data at high-spatial- and temporal-resolutions is currently cost prohibitive, and the use of unmanned aerial vehicles is prohibited in Grand Canyon National Park. Furthermore, an existing dataset of oblique imagery spanning 43 monitoring sites and nearly three decades has not been comprehensively examined until now and offers promising opportunities with the advance of computer vision and new segmentation methods.

Several studies have looked at topographic evolution during a handful of discrete flow events and decadal trends in size resulting from the combined effects of annual discharge patterns (Andrews et al., 1999; Hazel et al., 2008; Kaplinski et al., 2009; Wright and Kaplinski, 2011; Hazel et al., *in press*). However, the daily and monthly discharge patterns which create conditions favoring stability, dynamism, deposition, gradual erosion, and mass failures at sandbars throughout the canyon are not clear. Many efforts have been made to model the dynamics of eddy sandbars in Grand Canyon (Sloff et al., 2009; Travis et al., 2010; Nieuwboer, 2012; Sloff et al., 2012; Alvarez et al., 2017). However, models that did not explicitly account for mass failures tended to over-predict low-elevation subaqueous deposition and produced slopes steeper than the angle of repose (Sloff et al., 2009; Sloff et al., 2012; Alvarez et al., 2017). Analysis of the discharge patterns, which result in mass

failures, could improve this modeling. Furthermore, the parameterization of sub-annual erosion and deposition rates in those models could also be improved with better estimates of daily erosion and deposition from remote camera imagery.

Past remote camera studies have examined mass failure occurrence and quantified lateral erosion rates during the end of the UDO and beginning of the IF, but no study since has used the imagery data to examine daily and monthly behavior over multiple years and multiple sites during the MLFF (Cluer 1995; Dexter and Cluer 1999). Grams et al. (2018) demonstrated that the measurement of area time-series at a single site from remote camera imagery was possible. However, the method of image segmentation used in that study was time-consuming and infeasible for processing a dataset which now exceeds 1,000,000 images. I build on the methods described by Grams et al. (2018) in the following chapters and provide the most comprehensive analysis of sub-annual sandbar behavior during the MLFF. I discuss the limitations and demonstrate the potential of oblique time-series imagery for monitoring erosion and deposition in fluvial and coastal settings.

In Chapter 2 of this dissertation, I explore the following question: Can remote imagery be used to estimate sandbar volume at sub-annual scales? I utilize segmentation methods described by Grams et al. (2018) and demonstrate that estimating sandbar volume using oblique imagery is possible, but this is limited by the availability of imagery over a wide range of discharges. I applied hypsometric analysis, the study of elevational-area distribution, to measure the response of sandbars to HFES and quantify changes in elevational storage of fine-sediment.

In Chapter 3, I determine if rapid segmentation and measurement of sandbar in imagery can be applied to the current imagery dataset. I improve on the methods in Chapter 2 by applying a deep learning approach to image segmentation, drastically reducing the time needed to identify and measure sandbar area in oblique images. I show the process and

practical considerations needed to train a convolutional neural network end-to-end for the purpose of binary semantic segmentation. I examine the effects of hyperparameter choice and training sample size on the accuracy of three different symmetrical encoder-decoder convolutional neural networks.

In the fourth chapter of this dissertation examines the following questions: How much variability is there between sandbars in a similar geomorphic setting? How often do mass failures occur and are they correlated with flow patterns and comparable to monthly or daily change? To address these questions, I apply the new segmentation approach described in Chapter 3 to segment over 13,000 images, creating daily and monthly time-series of sandbar area at three dynamic reattachment bars. I use these high-temporal-resolution time-series to estimate daily and monthly changes in area. More broadly, I compile time-lapse videos to measure the frequency of mass failure events at 41 sites throughout the canyon, demonstrating links between monthly and seasonal shifts in discharge and mass failure occurrence.

In chapter 5, I summarize the findings from the previous three chapters and describe the policy-relevant implications of this work. I discuss the limitations of the remote camera dataset and make recommendations for improving the remote camera sandbar monitoring network in the Grand Canyon. I then conclude by outlining the most fruitful directions for future remote image analysis.

## **REFERENCES**

Alvarez, L. V., Schmeeckle, M.W., 2013. Erosion of river sandbars by diurnal stage fluctuations in the Colorado River in the Marble and Grand Canyons: full-scale laboratory experiments. *River Res. Appl.* 22, 1085–1095. <https://doi.org/10.1002/rra>

- Alvarez, L. V., Schmeeckle, M.W., Grams, P.E., 2017. A detached eddy simulation model for the study of lateral separation zones along a large canyon-bound river. *J. Geophys. Res. Earth Surf.* 122, 25–49. <https://doi.org/10.1002/2016JF003895>
- Andrews, E.D., Johnston, C.E., Schmidt, J.C., Gonzales, M., 1999. Topographic evolution of Sand Bars, in: *The Controlled Flood in Grand Canyon*. American Geophysical Union, p. 15.
- Baker, V.R., 1984. Flood Sedimentation in Bedrock Fluvial Systems, in: Koster, E.H., Steel, R.J. (Eds.), *Sedimentology of Gravel and Conglomerates*. Canadian Society of Petroleum Geologists, pp. 87–96.
- Bauer, B.O., Schmidt, J.C., 1993. Waves and Sandbar Erosion in the Grand Canyon: Applying Coastal Theory to a Fluvial System. *Ann. Assoc. Am. Geogr.* <https://doi.org/10.1111/j.1467-8306.1993.tb01946.x>
- Beus, S.S., Avery, C.C., 1992. *The Influence of Variable Discharge Regimes on Colorado River Sand Bars Below Glen Canyon Dam*. Flagstaff, AZ.
- Bierman, P.R., Montgomery, D.R., 2014. *Key Concepts in Geomorphology*. W.H. Freeman and Company, New York, NY.
- Cluer, B.L., 1995. Cyclic Fluvial Processes and Bias in Environmental Monitoring, Colorado River in Grand Canyon. *J. Geol.* 103, 411–421. <https://doi.org/10.1086/629777>
- Converse, Y.K., Hawkins, C.P., Valdez, R.A., 1998. Habitat relationships of subadult humpback chub in the Colorado River through Grand Canyon: Spatial variability and implications of flow regulation. *Regul. Rivers-Research Manag.* 14, 267–284. [https://doi.org/Doi 10.1002/\(Sici\)1099-1646\(199805/06\)14:3<267::Aid-Rrr501>3.0.Co;2-5](https://doi.org/Doi%2010.1002/(Sici)1099-1646(199805/06)14:3<267::Aid-Rrr501>3.0.Co;2-5)
- Dexter, L.R., Cluer, B.L., 1996. *Extension Study: Daily Dynamics of Grand Canyon Sandbars; Monitoring with Terrestrial Photogrammetry*. Flagstaff, AZ.
- Dexter, L.R., Cluer, B.L., 1999. Cyclic erosional instability of sandbars along the Colorado River, Grand Canyon, Arizona. *Ann. Assoc. Am. Geogr.* 89, 238–266. [https://doi.org/Doi 10.1111/1467-8306.00144](https://doi.org/Doi%2010.1111/1467-8306.00144)
- Dolan, R., Howard, A., Gallenson, A., 1974. Man's impact on the Colorado River in the Grand Canyon. *Am. Sci.* 62, 392–401.

- East, A.E., Collins, B.D., Sankey, J.B., Corbett, S.C., Fairley, H.C., Caster, J., 2016. Conditions and processes affecting sand resources at archeological sites in the Colorado River corridor below Glen Canyon Dam, Arizona. Prof. Pap. <https://doi.org/10.3133/pp1825>
- Fischenich, J.C., Allen, H., 2000. Stream Management. Vicksburg, MS.
- Grams, P.E., Schmidt, J.C., 1999. Geomorphology of the Green River in the Eastern Unita Mountains, Dinosaur National Monument, Colorado and Utah, in: Miller, A.J., Gupta, A. (Eds.), Varieties of Fluvial Form. John Wiley & Sons, Ltd, pp. 81–111.
- Grams, P.E., Topping, D.J., Schmidt, J.C., Hazel, J.E., Kaplinski, M., 2013. Linking morphodynamic response with sediment mass balance on the Colorado River in Marble Canyon: Issues of scale, geomorphic setting, and sampling design. J. Geophys. Res. Earth Surf. 118, 361–381. <https://doi.org/10.1002/jgrf.20050>
- Grams, P.E., Schmidt, J.C., Wright, S.A., Topping, D.J., Melis, T.S., Rubin, D.M., 2015. Building sandbars in the Grand Canyon. Eos (Washington. DC). 96. <https://doi.org/10.1029/2015EO030349>
- Grams, P.E., Tusso, R.B., Buscombe, D., 2018. Automated remote cameras for monitoring alluvial sandbars on the Colorado River in Grand Canyon, Arizona, Open-File Report. Reston, VA. <https://doi.org/10.3133/ofr20181019>
- Hazel Jr., J.E., Kaplinski, M., Parnell, R.A., Kohl, K., Schmidt, J.C., 2008. Monitoring fine-grained sediment in the Colorado River Ecosystem, Arizona-Control network and conventional survey techniques. U.S. Geol. Surv. Open-File Rep. 2008-1276 15.
- Hazel Jr., J.E., Grams, P.E., Schmidt, J.C., Kaplinski, M., 2010. Sandbar response following the 2008 high-flow experiment on the Colorado River in Marble and Grand Canyons. U.S. Geol. Surv. Sci. Investig. Rep. 2010-5015 52.
- Hazel, B.J.E., Kaplinski, M., Grams, P., Hamill, D., Buscombe, D., Mueller, R., Ross, R.P., Jewell, S., *in press*. Multi-Decadal Sandbar Response to Flow Management Downstream from a Large Dam : the Colorado River in Marble and Grand Canyons , Arizona.
- Howard, A., Dolan, R., 1981. Geomorphology of the Colorado River. J. Geol. 89, 269–298.

- Kaplinski, M., Hazel, J.E., Parnell, R., Breedlove, M., Kohl, K., Gonzales, M., 2009. Monitoring Fine-Sediment Volume in the Colorado River Ecosystem, Arizona: Bathymetric Survey Techniques. U.S. Geol. Surv. Open-File Rep. 2009-1207 i, 41.
- Kearsley, L.H., Schmidt, J.C., Warren, K.D., 1994. Effects of Glen Canyon dam on Colorado River sand deposits used as campsites in Grand Canyon National Park. *Regul. Rivers* 9, 137–149.
- Kieffer, S.W., 1985. The 1983 Hydraulic Jump in Crystal Rapid: Implications for River-Running and Geomorphic Evolution in the Grand Canyon. *J. Geol.* 93, 385–406. <https://doi.org/10.1086/628962>
- LaGory, K.E., Hlohowskyi, I., Tomasko, D., Hayse, J.W., Durham, L., 1993. A Review of Proposed Glen Canyon Dam Interim Operating Criteria. Argonne, Illinois.
- Leopold, L.B., 1969. The Rapids and the Pools - Grand Canyon. *United States Geol. Surv. Spec. Publ.* 669D, 131–145.
- Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2004. *Remote Sensing and Image Interpretation*, 5th ed. Wiley and Sons. <https://doi.org/10.1017/CBO9781107415324.004>
- Lovich, J., Melis, T.S., 2005. 10 years of adaptive management, in: *The State of the Colorado River Ecosystem in Grand Canyon*.
- Mastbergen, D.R., Beinssen, K., Nédélec, Y., 2019. Watching the beach steadily disappearing: The evolution of understanding of retrogressive breach failures. *J. Mar. Sci. Eng.* 7. <https://doi.org/10.3390/jmse7100368>
- Melis, T.S.E., 2011. Effects of Three High-Flow Experiments on the Colorado River Ecosystem Downstream from Glen Canyon Dam, Arizona. *U.S. Geol. Surv. Circ.* 1366, 147 p.
- National Research Council, 1996. *River Resource Management in the Grand Canyon*. The National Academies Press, Washington, D.C.
- National Research Council, 1999. *Downstream: Adaptive Management of Glen Canyon Dam and the Colorado River Ecosystem*. National Academy Press, Washington, D.C.

- Nieuwboer, B., 2012. Numerical modelling of Colorado sandbar growth. MSc Thesis. T.U. Delft, Delft, Netherlands.
- O'Neill, M.P., Kuhns, M.R., 1994. Stream Bank Erosion and Flushing Flows, in: Stream Notes. USDA Forest Service, Streams Systems Technology Center, Fort Collins, CO.
- Rubin, D.M., Schmidt, J.C., Moore, J.N., 1990. Origin, Structure, and Evolution of a Reattachment Bar, Colorado River, Grand Canyon, Arizona. *J. Sediment. Petrol.* 60, 982–991.
- Rubin, D.M., Anima, R.A., Sanders, R., 1994. Measurements of Sand Thicknesses in Grand Canyon, Arizona, and a Conceptual Model for Characterizing Changes in Sand-bar Volume Through Time and Space: OFR 94-587. Menlo Park, CA.
- Rubin, D. M., D. J. Topping, J. C. Schmidt, J. E. Hazel, M. A. Kaplinski, and T. S. Melis (2002), Recent sediment studies refute Glen Canyon Dam hypothesis, *EOS Trans. Am. Geophys. Union*, 83(25).
- Sankey, J.B., Draut, A.E., 2014. Gully annealing by aeolian sediment: Field and remote-sensing investigation of aeolian-hillslope-fluvial interactions, Colorado River corridor, Arizona, USA. *Geomorphology* 220, 68–80.  
<https://doi.org/10.1016/j.geomorph.2014.05.028>
- Sloff, K., Wright, S., Kaplinski, M., 2009. High resolution three dimensional modeling of river eddy sandbars, Grand Canyon, USA. *River, Coast. Estuar. Morphodynamics RCEM* 2009 883–890.
- Sloff, K.C.J., Nieuwboer, B., Logan, B., Nelson, J.M., 2012. Effect of sediment entrainment and avalanching on modeling of Grand Canyon eddy bars 791–798.
- Schmidt, J.C., Graf, J.B., 1990. Aggradation and Degradation of Alluvial Sand Deposits, 1965 to 1986, Colorado River, Grand Canyon National Park, Arizona. U.S Geological Survey Professional Paper 1493. Washington, D.C.
- Schmidt, J.C., Rubin, D.M., 1995. Regulated streamflow, fine-grained deposits, and effective discharge in canyons with abundant debris fans. *Nat. Anthropog. Infl. Fluv. Geomorphol.* 177–195.
- Schmidt, J.C., 1999. Summary and Synthesis of Geomorphic Studies Conducted During the 1996 Controlled Flood in Grand Canyon, in: Webb, R.H., Schmidt, J. C., Marzolf,

- R.G., Valdez, R.A. (Eds.), *The Controlled Flood in Grand Canyon*. American Geophysical Union, Washington DC., pp. 329–342.
- Schmidt, J.C., Grams, P.E., 2011. The High Flows, in: Melis, T.S. (Ed.), *Effects of Three High-Flow Experiments on the Colorado River Ecosystem Downstream of Glen Canyon Dam, Arizona: Circular 1366*. U.S. Geological Survey, Reston, VA, pp. 25–50.
- Simon, A., 1989. A model of channel response in disturbed alluvial channels. *Earth Surf. Process. Landforms*. <https://doi.org/10.1002/esp.3290140103>
- Simon, A., Curini, A., Darby, S., Langendoen, E.J., 1999. Streambank Mechanics and the Role of Bank and Near-Bank Processes in Incised Channels, in: Darby, S.E., Simon, A. (Eds.), *Incised River Channels: Processes, Forms, Engineering and Management*. John Wiley & Sons, West Sussex, UK, pp. 123–152.
- Stevens, L., Ayers, T., Bennett, J., Christensen, K., 2001. Planned Flooding and Colorado River Riparian Trade-Offs Downstream from Glen Canyon Dam, Arizona. *Ecol. Appl.* 11, 701–710.
- Topping, D.J., Rubin, D.M., Vierra Jr., L.E., 2000a. Colorado River sediment transport 1. Natural sediment supply limitation and the influence of Glen Canyon Dam. *Water Resour. Res.* 36, 515–542.
- Topping, D.J., Rubin, D.M., Nelson, J.M., Iii, P.J.K., Corson, I.C., 2000b. Colorado River sediment transport 2. Systematic bed-elevation and grain-size effects of sand supply limitation. *Water Resour. Res.* 36, 543–570.
- Travis, Q.B., Houston, S.L., Marinho, F. a. M., Schmeeckle, M., 2010. Unsaturated Infinite Slope Stability Considering Surface Flux Conditions. *J. Geotech. Geoenvironmental Eng.* 136, 963–974. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0000301](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000301)
- U.S. Department of the Interior (1970), *Glen Canyon Dam and Powerplant: Technical Record of Design and Construction*. Bureau of Reclamation, Denver, Colo., 658
- U.S. Department of the Interior, 2011. *Environmental Assessment: Development and Implementation of a Protocol for High-Flow Experimental Releases from Glen Canyon Dam, Arizona, 2011 through 2020*, U.S. Bureau of Reclamation. Salt Lake City, UT. <https://doi.org/10.1007/BF02244392>

- Valdez, R.A., Ryel, R.J., 1995. Life History and Ecology of the Humpback Chub in the Colorado Tinkler, K., Wohl, E., 1998. A primer on bedrock channels, in: Tinkler, K., Wohl, E. (Eds.), *Rivers Over Rock: Fluvial Processes in Bedrock Channels*, Geophysical Monograph 107. American Geophysical Union, Washington, D.C., pp. 1–18.
- Van Den Berg, J.H., Van Gelder, A., Mastbergen, D.R., 2002. The importance of breaching as a mechanism of subaqueous slope failure in fine sand. *Sedimentology* 49, 81–95. <https://doi.org/10.1111/j.1525-139X.2006.00168.x-i1>
- Webb, R.H., Pringle, P.T., Rink, G.R., 1989. Debris flows from tributaries of the Colorado River, Grand Canyon National Park, Arizona. U.S. Geol. Surv. Prof. Pap. 1492 39.
- Wiele, S., and M. Torizzo ( 2005), Modelling of sand deposition in archaeologically significant reaches of the Colorado River in Grand Canyon, USA, in Bates, P.D., Lane, S.N., and Ferguson, R.I., eds., *Computational fluid dynamics—Applications in environmental hydraulics*: Hoboken, New Jersey, John Wiley & Sons, chap. 14, p. 357–394, accessed November 29, 2010, at <http://onlinelibrary.wiley.com/doi/10.1002/0470015195.ch14/summary>.
- Wright, S.A., Melis, T.S., Topping, D.J., Rubin, D.M., 2005. Influence of Glen Canyon Dam Operations on Downstream Sand Resources of the Colorado River in Grand Canyon, in: Gloss, S., Lovich, J., Melis, T. (Eds.), *The State of the Colorado River Ecosystem in Grand Canyon*. United States Geological Survey, Reston, VA.
- Wright, S.A., Kaplinski, M., 2011. Flow structures and sandbar dynamics in a canyon river during a controlled flood, Colorado River, Arizona. *J. Geophys. Res. Earth Surf.* 116, 1–15. <https://doi.org/10.1029/2009JF001442>

## **CHAPTER 2: CHAPTER 2: USING OBLIQUE IMAGERY TO MEASURE HYPSONETRIC CHANGES IN SANDBAR VOLUME FOLLOWING CONTROLLED FLOODS IN THE GRAND CANYON**

**Ryan Lima**, Ph.D. Candidate, Northern Arizona University, Flagstaff, AZ, [rl587@nau.edu](mailto:rl587@nau.edu)

**Daniel Buscombe**, Assistant Research Professor, Northern Arizona University, Flagstaff, AZ, [Daniel.Buscombe@nau.edu](mailto:Daniel.Buscombe@nau.edu)

**Temuulen Sankey**, Associate Professor, Northern Arizona University, Flagstaff, AZ, [Temuulen.Sankey@nau.edu](mailto:Temuulen.Sankey@nau.edu)

**Paul Grams**, Research Hydrologist, USGS, Flagstaff, AZ, [pgrams@usgs.gov](mailto:pgrams@usgs.gov)

**Erich Mueller**, Visiting Assistant Professor, University of Wyoming, Laramie, WY, [erich.mueller@uwo.edu](mailto:erich.mueller@uwo.edu)

### **ABSTRACT**

Measuring changes in the elevation distribution of subaerial fine ( $< 2mm$ ) sediment and estimating sandbar volume multiple times per year can improve sediment budget calculations in fluvial systems. In the Grand Canyon of the Colorado River, effects of dam operations on sandbar size and distribution is of long-term management interest. Bar-building controlled floods have been implemented in 1996, 2004, 2008, 2012, 2013, 2014, 2016, and 2018 to mitigate sandbar erosion. Annual topographic surveys provide a single measurement of sandbar change caused by the integrated effects of all flows in one year (both controlled floods and normal dam releases), but do not measure erosion and deposition caused by specific operations or individual floods. At one sandbar monitoring site in Grand Canyon, we

demonstrate that imagery from autonomous digital cameras can be used to provide quantitative measures of sandbar hypsometry multiple times per year without costly and labor-intensive surveys. We describe methods for measuring changes in the storage of fine sediment at monthly or seasonal timescales by constructing hypsometric (area-elevation relation) curves. These curves are created and updated with sandbar area measurements from georectified images taken multiple times each day. As the water surface elevation fluctuates with daily, seasonal, and monthly discharge patterns, sandbar area and volume can be estimated using known stage-discharge relationships. We present parameters extracted from image-derived hypsometries to estimate sandbar volume and elevation relief ratio, which provides a new way to quantitatively measure monthly or seasonal changes in fine sediment storage.

## **2.1. INTRODUCTION**

The closure of Glen Canyon Dam in 1963 significantly altered the discharge patterns and sediment supply of the Colorado River in Grand Canyon. Increases in minimum and mean discharges, increases in daily discharge fluctuation, decreases in maximum discharge, and a reduction of sediment supply have created a deficit of fine sediment (Topping et al. 2000) and reductions in the size and extent of sandbars throughout Grand Canyon (Dolan et al. 1974; Schmidt et al. 2004). The preservation of fine sediment resources downstream of Glen Canyon Dam has been a long-standing management focus. Sandbars are used for camping by nearly 25,000 people each year who float the Grand Canyon. Sandbars also form habitat used by native fish and provide a source of sand for aeolian transport to upslope archaeological resources, protecting those areas from erosion (Sankey et al. 2018).

Controlled floods have been carried out periodically since 1996 to rebuild and maintain sandbars. They are conducted following significant tributary inputs of sediment.

The high flows are intended to redistribute sand from low elevations within the channel or in eddies to high elevations where sand is less vulnerable to erosion from fluctuating flows and can provide the benefits previously listed. Following the high flow events, topographic surveys are the primary method for monitoring the storage of subaerial fine sediment within Grand Canyon. These surveys are typically conducted once per year and provide a detailed snapshot of fine sediment storage at 44 sandbar monitoring sites (Hazel et al. 2006).

To supplement the detailed topographic surveys, daily photographic monitoring of sandbars in Grand Canyon began in the early 1990s with film cameras, which collected one image per day at just a few sandbar monitoring sites. These images revealed that sandbars undergo cycles of rapid short-term erosion followed by gradual deposition throughout the year and that short-term rates of erosion equalled or exceeded the rates of erosion and deposition measured at annual timesteps (Dexter and Cluer, 1999). These findings reiterated the need for more frequent, short-term monitoring in addition to detailed annual monitoring.

The network of remote cameras was expanded throughout the 1990s and early 2000s to monitor 43 sandbars sites, and between 2008 and 2014, film cameras were replaced with digital cameras capturing five or more oblique images each day. These remote camera systems include a solar panel, a data logger, a 12V battery, a camera, and lens sealed in weatherproof boxes. The camera systems and their specifications, referred to hereafter as *remote cameras* are described by Bogle et al., (2013). They operate autonomously with semi-annual maintenance schedules. The sandbar imagery dataset contains over 1,000,000 images and is one of the most comprehensive and longest records of fluvial sandbar monitoring currently available. Recently, imagery from remote cameras has been used to qualitatively assess size changes following controlled floods, and to quantify sandbar area at a single elevation through time (Grams et al. 2018, Tusso et al. 2015).

We describe methods for quantifying changes in fine sediment storage at sandbars in Grand Canyon using oblique imagery captured from the remote cameras. We utilize rectified oblique imagery to: (1) isolate sandbar pixels from images, taken at multiple discharges, to compute sandbar area at different elevations; (2) examine the area-elevation relationship and (3) estimate sandbar volume at monthly or seasonal timescales. Our analysis focuses on the subaerial portion of sandbars, which are emergent during most flows. Since daily discharges are rarely less than  $8,000 \text{ ft}^3 \text{ s}^{-1}$  ( $227 \text{ m}^3 \text{ s}^{-1}$ ), our analyses are limited to the portions of the sandbar which remain subaerial at a discharge of  $8,000 \text{ ft}^3 \text{ s}^{-1}$ . The elevation of the water-surface at  $8,000 \text{ ft}^3 \text{ s}^{-1}$  is hereafter referred to as the reference elevation. At river mile 30.7 the reference elevation is  $855.47 \text{ m}$ .

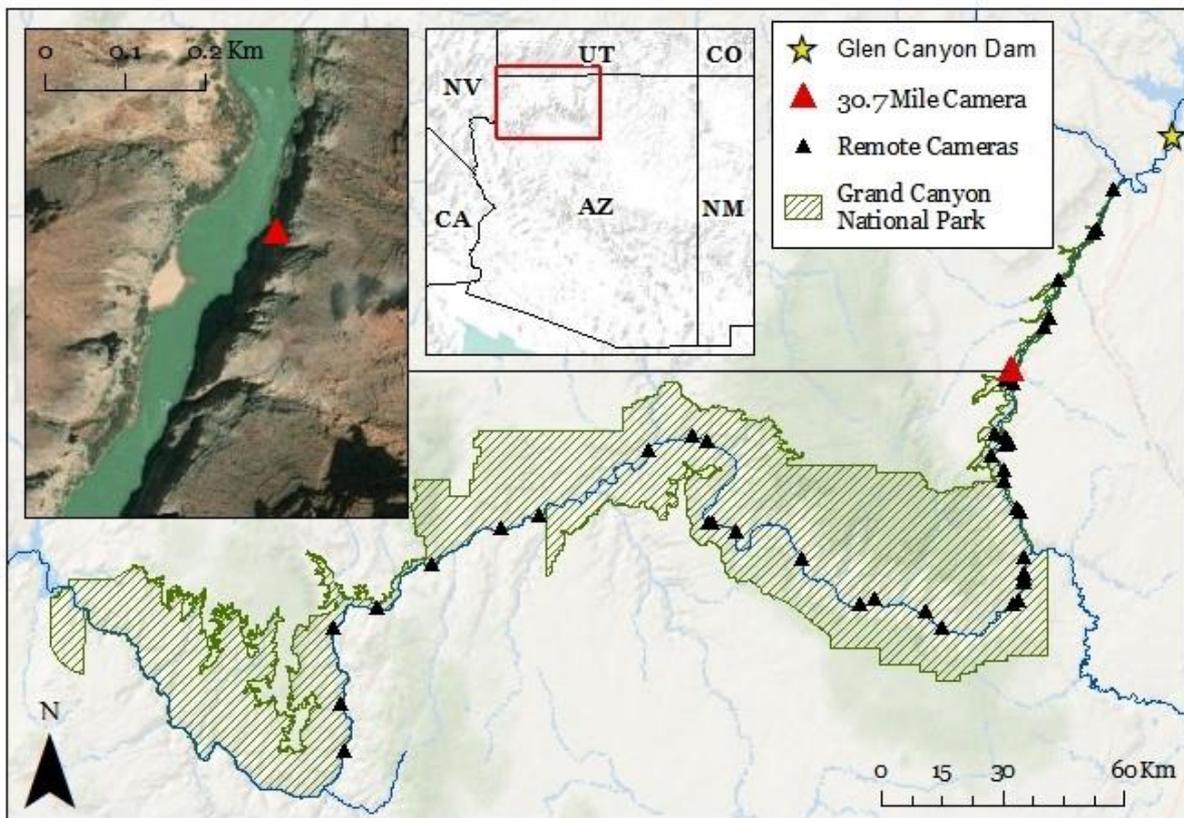
We refer to elevations in two different ways in this study; the first is in meters above sea level. This absolute elevation is useful for understanding changes in storage at a particular site. Secondly, we use discharge as a proxy for elevation based on an observed stage-discharge relationship to show the proportion of the bar inundated at different flow levels. This relative elevation-proxy provides a context for understanding changes in storage in relation to particular flow patterns and allows for comparison across sites throughout the canyon which occur at different absolute elevations.

In this study, we examine hypsometry because it provides a quantitative estimate of overall sandbar steepness and allows quick visualization of where fine sediment is stored in relation to stage-elevations. We estimate sandbar volume by creating digital elevation models (DEMs) from the segmented sandbar outline in the oblique images collected in discrete (one to four day) intervals, where three or more images showing water-surfaces across a broad range of discharges.

## 2.2. METHODS

### 2.2.1. Study Area

We chose to focus on the sandbar site located at 30.7 river miles downstream of Lees Ferry, referred to, hereafter as RM30.7 (Figure 2.1). This site was chosen principally because it is dynamic, but also because it is almost vegetation-free. The lack of vegetation makes automated segmentation of sandbar area simpler, which was important in the development of our new method. Further, RM30.7 is located just 0.7 river miles downstream of USGS gage # 09383050, which provides 15-minute measurements of discharge. In 2009, the location of the camera at RM30.7 was moved, and we primarily use images (up to five times per day) captured between 2009 and 2017. Images were pre-processed and processed via a sequence of analysis and methods. We describe each of the important steps below



**Figure 2.1** The study area, depicting the Grand Canyon in hatched green, the Colorado River, remote camera monitoring sites (black triangles) and the RM30.7 site (red triangle).

## **2.2.2. Image-derived area measurement**

Sandbar area measurements from the images involved registering and rectifying the images to known references and then segmenting the sandbars to calculate their areas

### **2.2.2.1. Registration**

Registration is the process of spatially matching multiple images of the same scene. This is needed because small shifts in the camera viewing angle occur during camera maintenance or as a result of environmental conditions. Batch registration was carried out to register all images taken at RM30.7 to a single reference image using a 2-D Fast Fourier Transform (Grams et al., 2018).

### **2.2.2.2. Rectification**

Once registered, images were rectified using a homography or transformation between two images in the same planar surface. A homography was developed by using the known location of particular rocks (i.e., hardpoints) and panels which were surveyed for their precise location and imaged with the remote cameras; collectively these are known as ground control points (GCPs). This transformation assumes all ground control points are on the same 2-D plane or have the same elevation. The homography could then be applied to images so that distances in pixels could be translated into distances on the Earth's surface. The homography was generated using the OpenCV package in Python 2.7. The workflow is described by Grams et al. (2018).

### **2.2.2.3. Segmentation**

To perform segmentation, or the process of delineating pixels containing sand from non-sand pixels, we used a program called *RCSandseg*, implemented in Python 2.7 and described by Grams et al. (2018). *RCSandseg* allows the user to define a bounding box around the sandbar and uses the GrabCut algorithm (Rother and Kolmogorov, 2004) to

delineate a subject from the background. This supervised method also allows the user to correct misclassified pixels. Each segmentation required minimal user input and takes between one and five minutes.

#### **2.2.2.4. Area measurement**

Once the sandbars were segmented, we imported the coordinates of the sandbar shoreline from all photos, or the boundary between pixels classified as sandbar and those classified as background, into ArcMap 10.6.1 and created a polygon from the points. The area of the polygon was then calculated using the Geometry tool within ArcMap 10.6.1.

#### **2.2.2.5. Estimating water-surface elevation**

To associate a measured water-surface elevation to an image, we first estimated the discharge pictured in each image using the time recorded in the image EXIF metadata and subtracted a lag-time to the upstream gage (USGS-09393050) from the image time. This provides an estimated time when the discharge wave in the image passed the upstream gage. Since the gage records discharge every 15-minutes, discharge was interpolated from the two discharge measurements closest (temporally) to the discharge wave. That discharge was then converted to water-surface elevation using the stage-discharge relationship developed from several decades of field observations at the RM30.7 site (Hazel et al., 2007)

### **2.2.3. Volume from image-derived DEMs**

Our first approach in estimating sandbar volume was to find periods of one to four days where images captured discharges at or below the reference elevation and images at higher discharges  $> 12,000 \text{ft}^3 \text{s}^{-1}$ . A total of twenty-two such periods were identified between 2012 and 2015. An image from at or below the reference elevation was selected for each period along with at least two other images at higher elevations. A total of 68 images from those twenty-two intervals were selected. These images were segmented to identify the

sandbar. The homography (described above) was applied to points along the boundary between the sandbar and the water providing geographic coordinates for those points. An elevation value estimated using the process described above was paired with the water's edge points based on the image time. Points were converted into polylines in ArcMap 10.6.1. Polylines were used to create a triangular irregular network, then the triangular irregular network was converted into a digital elevation model which was used to calculate volume.

#### **2.2.4. Hypsometric analysis**

Hypsometry is a measure of the relationship between area and elevation, usually applied to a basin or watershed. Hypsometric curves are created by plotting the cumulative area of a basin against its relative elevation (Strahler, 1952). For a drainage basin, the shape of this curve is characteristic of the erosional processes occurring within the basin and the geologic setting. Convex hypsometries may indicate a tectonically young basin where much of the mass is stored at higher elevations. In such basins, fluvial processes play a greater role (Strahler, 1952). Elevation relief ratio, shown in Equation 2.1,

$$\text{(Eq. 2.1) } \frac{\text{Mean Elevation} - \text{Min Elevation}}{\text{Max Elevation} - \text{Min Elevation}} = \text{Elevation Relief Ratio}$$

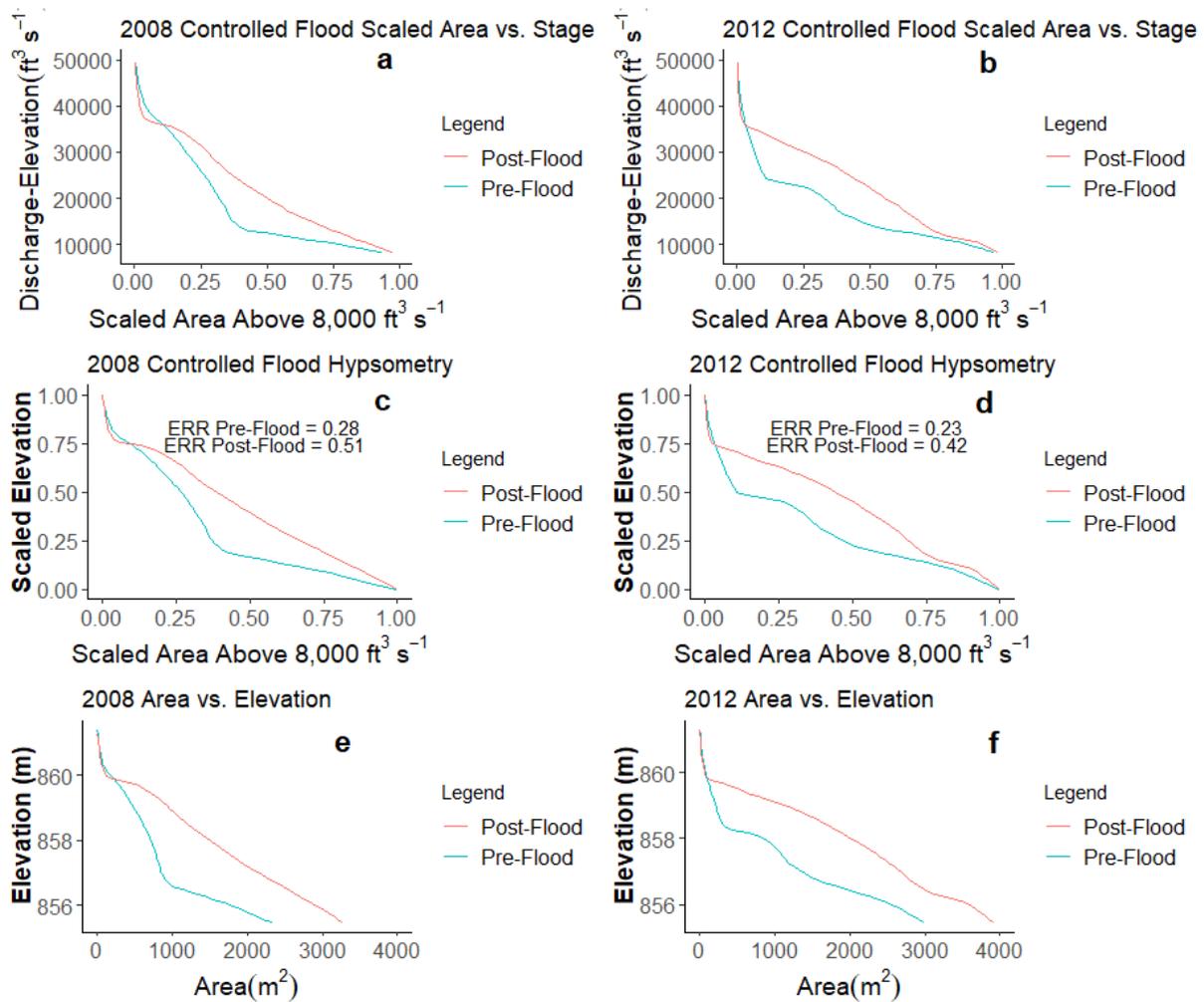
(Pike and Wilson, 1971)

is the normalized elevation of the mean basin area. Described by Pike and Wilson (1971), it provides a single number that can be used to compare basins irrespective of scale. Concave hypsometries result in lower elevation relief ratios (~0.2 – 0.5), whereas convex hypsometries result in higher elevation relief ratios (~0.5-0.8) (Pike and Wilson, 1971). This metric is useful because it distills the complexity of a curve, which may have several inflection points and different slopes, into a single number which can be compared over time at the same site and across several sandbar sites in the canyon.

These concepts are useful to the analysis and monitoring of sandbars in several ways. Like drainage basins, sandbars can have complex morphologies which are not easily quantified. Hypsometric curves, and the elevation relief ratio can provide a simple means of quantifying changes occurring to these sandbars through time. Determining the elevation relief ratio for a sandbar at a particular time might also provide insight into the dominant types of erosion, which can be expected to occur at a given site, and how that might be related to overall sandbar slope or degree of convexity. Alvarez and Schmeeckle (2013), for example, found in the laboratory that the slope angle of sandbars determines the dominant mode of erosion in response to diurnal stage fluctuations. A field test of this finding is one eventual goal of the present work.

Controlled floods in Grand Canyon generally mobilize sediment from low elevations on the channel bed and store it at higher elevations within eddies and on the riverbanks (Grams et al., 2015). The hypsometric curves in figure 2.2 were created from surveys before and after the 2008 and 2012 controlled floods, which supports the conclusions that controlled floods are depositing sediment at higher elevations. Although controlled floods were also conducted in 2014, 2014, 2015, and 2017, sandbar topography was not measured in the field before and after these events.

Maintaining camp-able sandbar area that is above the water-surface most of the time is a primary management concern along the Colorado River in Grand Canyon. In comparison to volume alone, a hypsometric curve provides more information about the usable area since the steeper sandbars may be less desirable for camping and hypsometry provides insight into the overall sandbar slope angle.



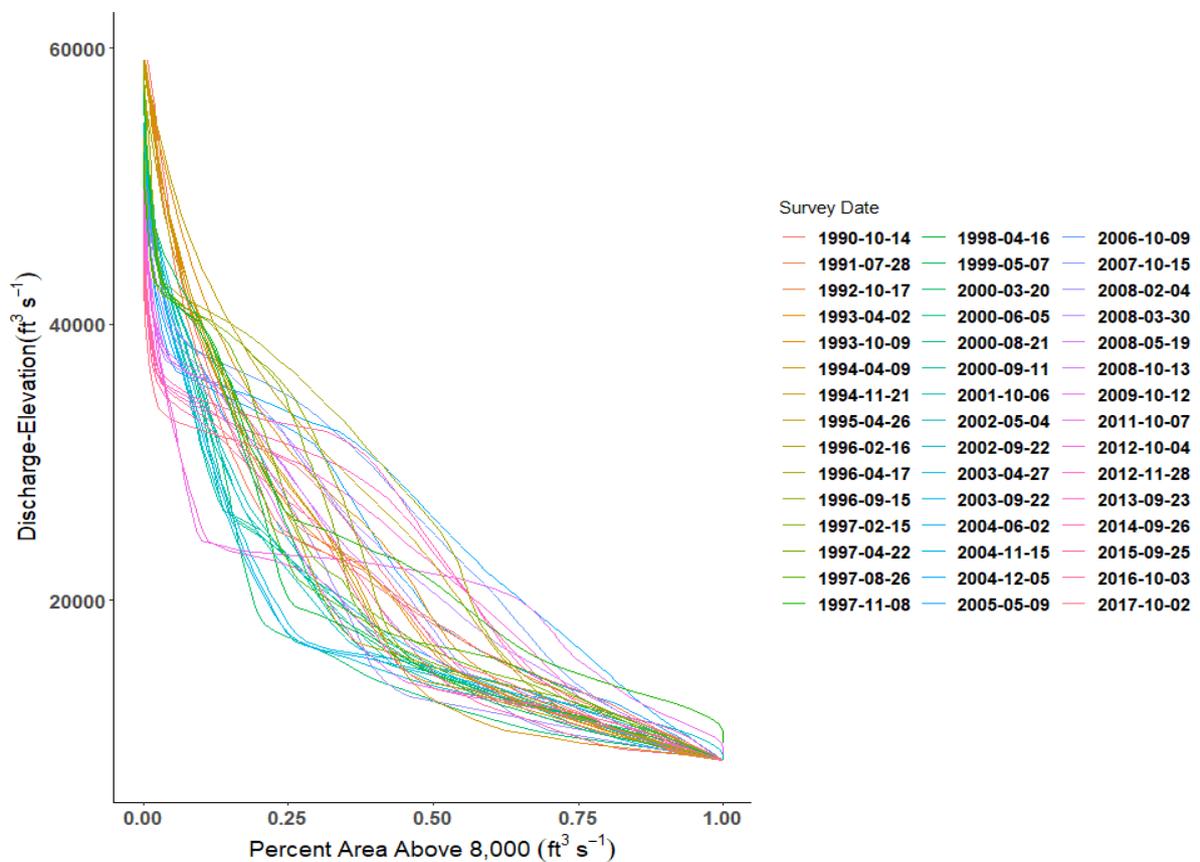
**Figure 2.2** Relationships between the area and elevation for the sandbar at RM30.7 plotted in three ways, before and after the 2008 and 2012 flood using data from topographic surveys. Panels (a) and (b) show sandbar area scaled to the subaerial sandbar area at a discharge of  $8,000 \text{ ft}^3 \text{ s}^{-1}$ , plotted against the water surface elevation  $\text{ft}^3 \text{ s}^{-1}$  for each year. Panels (c) and (d) show normalized area and elevation showing the hypsometry for each year, where ERR is the elevation relief ratio. Panels (e) and (f) show area ( $\text{m}^2$ ) by elevation (m) relationship for each year.

## 2.3. RESULTS AND DISCUSSION

### 2.3.1. Surveyed hypsometries

We built hypsometric curves using data from 42 topographic surveys of RM30.7 collected between 1990 and 2019. The measured area was scaled as a percentage of the sandbar area at the reference elevation and plotted against elevation. Here elevation is expressed as the water-surface elevation at a particular discharge in  $\text{ft}^3 \text{ s}^{-1}$  (Figure 2.3). We

built these hypsometric curves for two reasons. First, we sought to understand the variability in sandbar morphology which has been observed over the multi-decadal topographic survey period. Second, river management would benefit from an improved understanding of the relationship between sandbar area at the reference stage, and area scaled to the reference stage at various other water surface elevations. This could be used to develop a relationship through which sandbar volume can be predicted using two oblique images, one at the reference stage and another at a higher elevation.



**Figure 2.3** Relationship between survey-derived area measurements scaled to the reference elevation and the stage elevations for 42 sandbars surveys at RM30.7.

### 2.3.2. Hypsometrically-derived volume

To estimate sandbar volume using the rectified oblique imagery, we first needed to determine the relationship between area at multiple elevations and volume at the reference elevation. Linear models were fit to the hypsometric relation (figure 2.3) using survey-

derived area measured at the reference elevation and scaled area for each elevation with 1000  $ft^3s^{-1}$  intervals between 8,000 $ft^3s^{-1}$  and 45,000 $ft^3s^{-1}$  as a predictor, and volume at the reference elevation as the response (Equation 2.2):

$$(Eq. 2.2) V_{ref} = \beta_0 + \beta_1 A_{ref} + \beta_2 A_2 + \varepsilon$$

Where  $V_{ref}$  is the volume ( $m^3$ ) at the reference elevation, the predictor variables are  $A_{ref}$ , the area ( $m^2$ ) at the reference elevation, and  $A_2$ , the area ( $m^2$ ) at another elevation scaled to the area at the reference elevation.  $\beta_0$  is the intercept,  $\beta_1$  and  $\beta_2$  are the coefficients for the predictor variables, and  $\varepsilon$  is an error term. The  $lm()$  function in R was used to fit models using ordinary least squares. We examined model fit using a wide range of elevations for the second variable ( $A_2$ ). Coefficients of determination in Table 1 are  $> 0.9$  for linear models utilizing an area at the reference stage and a scaled subaerial sandbar area at discharges between 15,000 $ft^3s^{-1}$  and 26,000 $ft^3s^{-1}$ . Therefore, in a predictive capacity, if area can be measured accurately from oblique imagery, we can estimate sandbar volume within an accuracy of 90% or more, using just two images: one at the reference stage and a second at a water-surface elevation associated with flows between 15,000 $ft^3s^{-1}$  and 26,000 $ft^3s^{-1}$ .

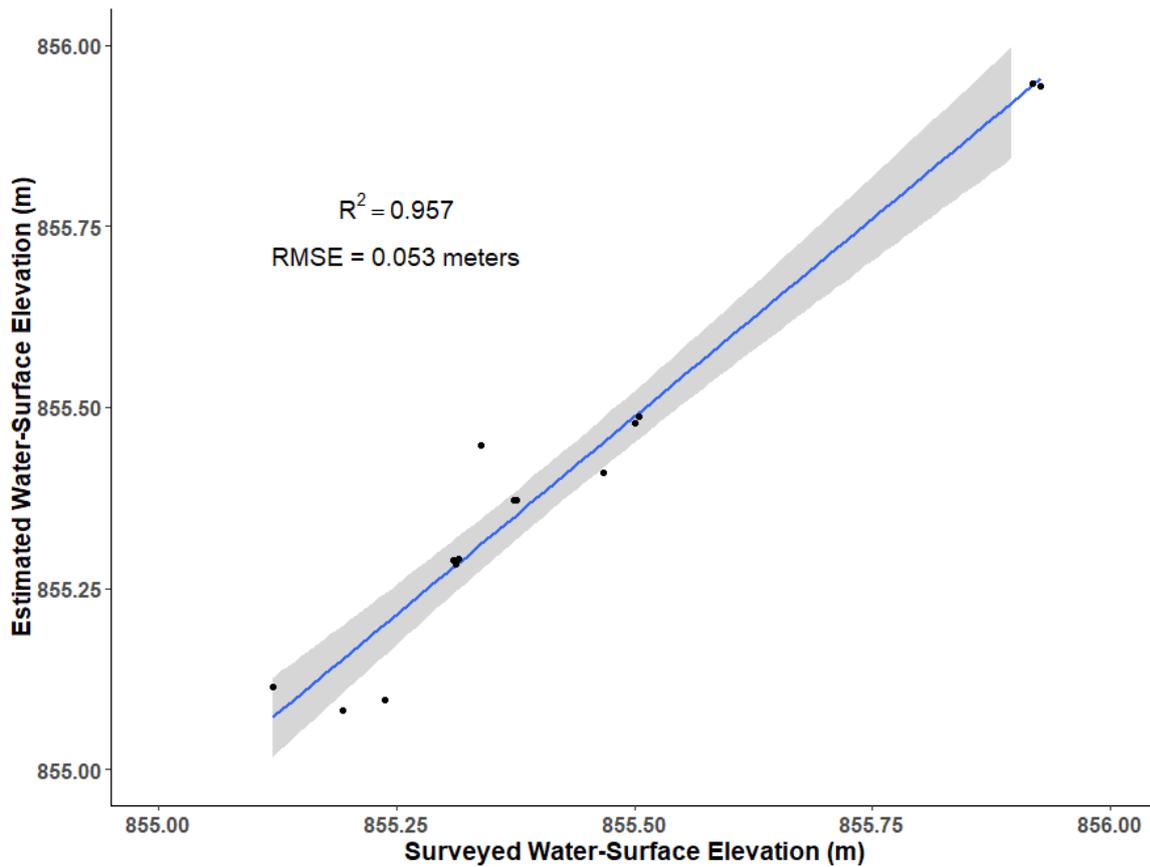
**Table 2.1** Survey-derived area volume relationships. Coefficients of determination and root-mean-squared error (RMSE) for linear models relating surveyed sandbar volume at the reference elevation with (1) the surveyed area at the reference elevation and (2) the scaled area at discharges between 14,000 $ft^3s^{-1}$  and 26,000 $ft^3s^{-1}$  at RM30.7.  $R_{adj}^2 = 1 - (1 - R^2) \frac{n-1}{n-p-1}$ , where  $p = \#$  of predictors.

<b><math>A^2</math> Discharge (<math>ft^3s^{-1}</math>) Elevation</b>	<b><math>R^2</math></b>	<b><math>R_{adj}^2</math></b>	<b>RMSE (<math>m^3</math>)</b>
<b>14000</b>	<b>0.874857</b>	<b>0.86844</b>	<b>601.3845</b>
<b>15000</b>	<b>0.911225</b>	<b>0.906673</b>	<b>506.5172</b>
<b>16000</b>	<b>0.921782</b>	<b>0.91777</b>	<b>475.4492</b>
<b>17000</b>	<b>0.926187</b>	<b>0.922402</b>	<b>461.8654</b>
<b>18000</b>	<b>0.928511</b>	<b>0.924845</b>	<b>454.5362</b>

<b>19000</b>	<b>0.929401</b>	<b>0.92578</b>	<b>451.6995</b>
<b>20000</b>	<b>0.93369</b>	<b>0.930289</b>	<b>437.7633</b>
<b>21000</b>	<b>0.940267</b>	<b>0.937204</b>	<b>415.486</b>
<b>22000</b>	<b>0.958439</b>	<b>0.956307</b>	<b>346.5727</b>
<b>23000</b>	<b>0.963504</b>	<b>0.961632</b>	<b>324.7686</b>
<b>24000</b>	<b>0.919457</b>	<b>0.915326</b>	<b>482.4634</b>
<b>25000</b>	<b>0.912903</b>	<b>0.908437</b>	<b>501.7069</b>
<b>26000</b>	<b>0.906134</b>	<b>0.901321</b>	<b>520.8383</b>

### **2.3.3. Accuracy of image water-surface elevations**

To determine how accurate the predicted water-surface elevation, based on known discharge values and the lag time to any site, was at any given day and time, we compared our results to the water-surface elevation in fourteen surveys where the water's edge was surveyed at RM30.7. The surveyed elevations were compared to estimated water-surface elevations using the same dates and times (Figure 2.4). Our predicted water surface elevations correlated with the surveyed elevations with an  $R^2 = 0.9573$  and produced a root-mean-squared error or (RMSE) of  $0.053m$ , indicating strong agreement between our estimates and survey results.

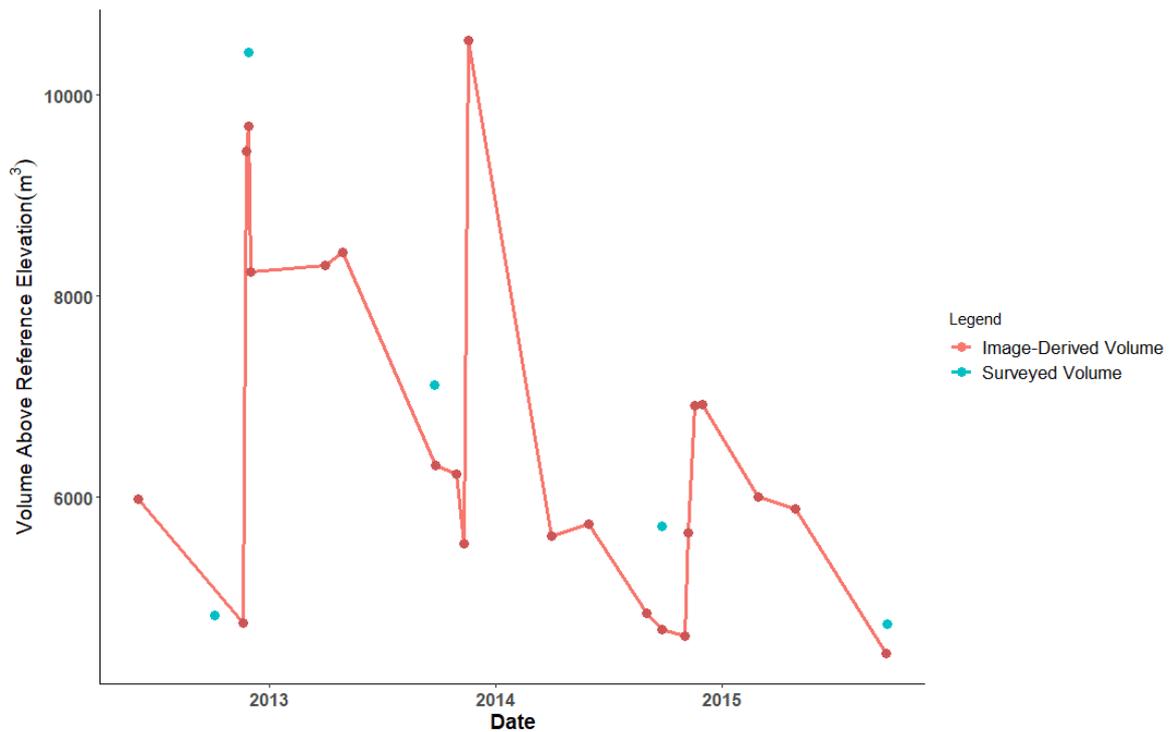


**Figure 2.4** Comparison between estimated water-surface elevation and surveyed water-surface elevations. The reported R-squared value of 0.957 suggests high agreement between estimated and surveyed water surface elevations.

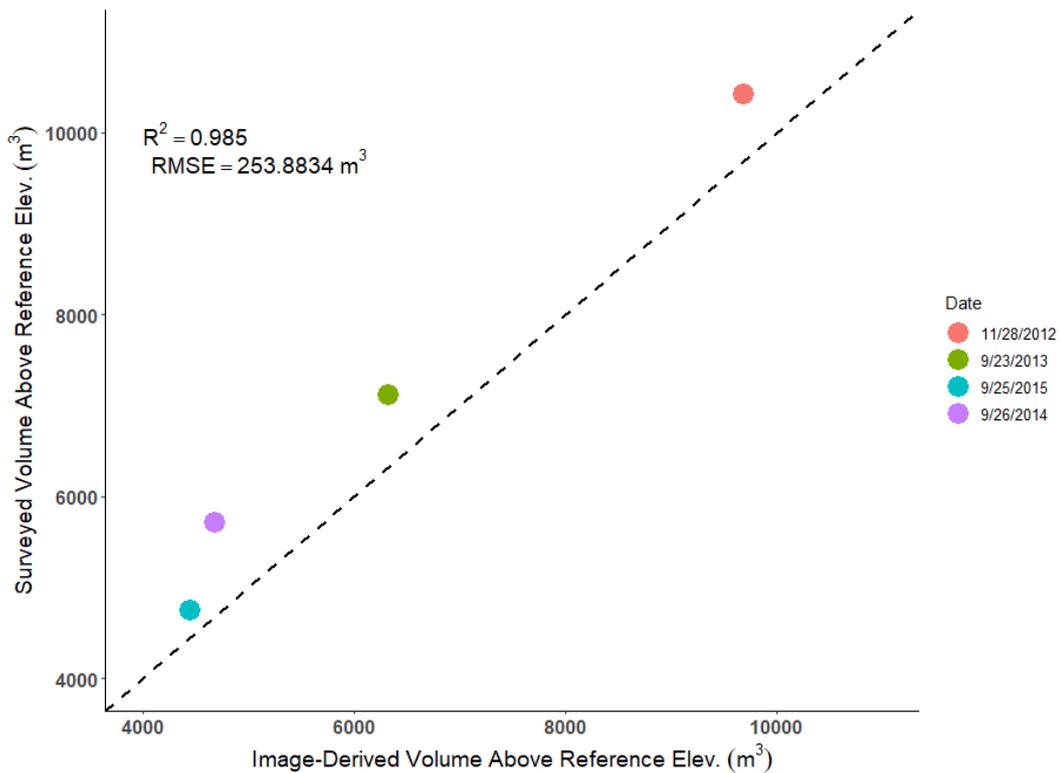
### 2.3.4. Volume from image-derived DEMs

When comparing image-derived sediment volume estimates above the reference elevation to the survey-derived volumes above the reference elevation (Figure 2.5), we observe that the trends in surveyed and estimated volume over time are similar. However, four out of five of the image-derived estimates underestimated sandbar volume (Figure 2.5). This is likely because the surveyed area is not exactly the same as the extent of sub-aerial sand picked up by our segmentation, the surveyed site boundary contains small areas that contain rocks or sandbar covered with vegetation. Therefore, in future studies, an offset or a correction factor, once properly defined, could be applied to image-derived volumes if the errors are found to be within the same range in every case (Figure 2.6). Unfortunately, direct

comparisons between the 22 image-derived volume estimates could only be made on four occasions between 2012 and 2016, when surveys occurred within a day or two of each estimate (Figure 2.6). More data are needed to quantify the accuracy of image-derived volume estimates. New imagery from 2017 – 2019 should allow for more accurate comparison in future work. Evaluating uncertainty is a future research goal and involved estimating the combined errors in the segmentation process, image rectifications, and area estimation, and how these errors propagate into volume measurements of resulting raster surfaces.



**Figure 2.5** Comparison between survey-derived volumes above the reference elevation (blue dots) and volumes estimated from image-derived DEMs (red line).

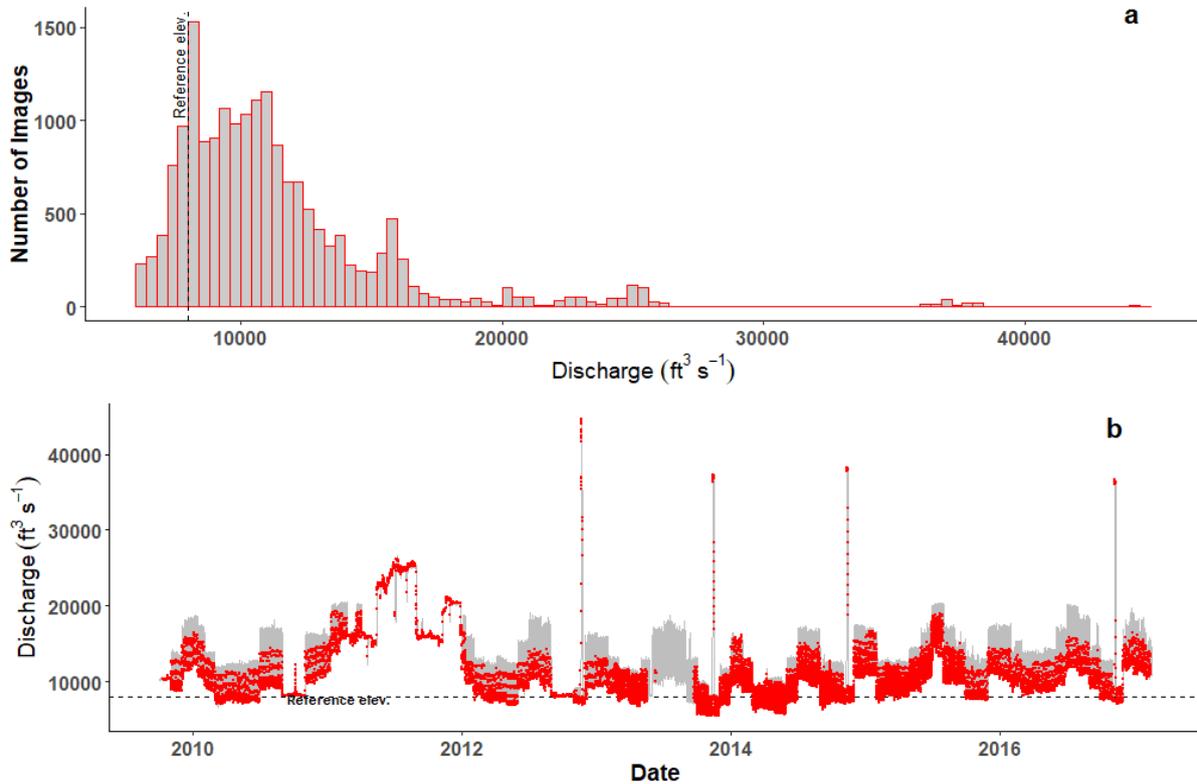


**Figure 2.6** Comparison between survey-derived volumes above the reference elevation and image-derived volumes for four corresponding dates.

### 2.3.5. Limitations

The creation of DEMs and hypsometric curves, and their accuracy, is limited by the frequency with which an image at or very near the reference elevation is taken, along with images at discharges greater than  $15,000 \text{ ft}^3 \text{ s}^{-1}$  within a few days of each other or a period when erosion and deposition can be assumed negligible. The frequency of images captured at discharges greater than  $20,000 \text{ ft}^3 \text{ s}^{-1}$  is low, and the period, where a broad range of discharges can be captured, are limited to infrequent controlled floods. Figure 2.7 shows the frequency of images at various discharges at RM30.7. Since portions of the sandbar that remain subaerial above  $20,000 \text{ ft}^3 \text{ s}^{-1}$  are rarely inundated, we assume that fluvial erosion of those portions of the bar is limited to controlled floods, and episodic mass failures likely resulting from erosion at lower elevations and back wasting. Such failure events are evident in imagery. Overall, we assume these high-elevation areas remain static in the absence of

obvious episodic mass failures detected by manual inspection of imagery. To aid in future analysis, the camera systems could be programmed to take 10 images per day instead of five which would provide additional area measurements and increase the accuracy of image-derived hypsometries, and potentially the accuracy of volume estimates.



**Figure 2.7** (a) Image frequency at RM30.7 by discharge; (b) Image times plotted over the estimated hydrograph for RM30.7. There are 18,699 total images between 2009 and 2017. Relatively few images of discharges above  $15,000 \text{ ft}^3 \text{ s}^{-1}$  are available.

## 2.4. CONCLUSIONS

Our results indicate that sandbars can be successfully delineated in oblique imagery.

Segmented sandbar imagery can be used to accurately estimate sandbar area and water surface elevation. Image-extracted area measurements were successfully combined to create contours and subsequent DEMs at the RM30.7 sandbar three to five times per year from 2012 – 2016, providing a mechanism for monthly or seasonal monitoring of sandbar size and

morphology. Volume measurements from image-derived DEMs show a similar trend of surveyed sandbar volumes. Currently, there are only four occasions where surveys correspond temporally to image-derived DEMs. Therefore, more data are needed to evaluate the accuracy of volume estimates. Analysis of images in 2017 – 2019 will likely provide more corresponding surveys to test image-derived volume estimate accuracy.

Future analysis will attempt to produce hypsometric curves from oblique imagery at more sites. This technique offers an important tool for monitoring the effects of controlled floods between surveys and may provide a single value, the elevation relief ratio, which can be used to compare changes in elevational storage at individual sites and across sites within Grand Canyon.

**Disclaimer:** Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government

## 2.5. REFERENCES

- Alvarez, L. V., and M. W. Schmeeckle (2013), Erosion of river sandbars by diurnal stage fluctuations in the Colorado River in the Marble and Grand Canyons: full-scale laboratory experiments., *River Res. Appl.*, 22(July 2011), 1085–1095, doi:10.1002/rra.
- Bogle, R., Velasco, M., & Vogel, J. (2013). An Automated Digital Imaging System for Environmental Monitoring Applications.
- Dexter, L. R., & Cluer, B. L. (1999). Cyclic erosional instability of sandbars along the Colorado River, Grand Canyon, Arizona. *Annals of the Association of American Geographers*, 89(2), 238–266. <http://doi.org/Doi.10.1111/1467-8306.00144>
- Dolan, R., Howard, A., & Gallenson, A. (1974). Man's impact on the Colorado River in the Grand Canyon. *American Scientist*, 62(4), 392–401. Retrieved from [http://erode.evsc.virginia.edu/papers/dolan\\_GrCyn\\_impact\\_74.pdf](http://erode.evsc.virginia.edu/papers/dolan_GrCyn_impact_74.pdf)

- Grams, P. E., J. C. Schmidt, S. A. Wright, D. J. Topping, T. S. Melis, and D. M. Rubin (2015), Building sandbars in the Grand Canyon, *Eos (Washington, DC)*, 96, doi:10.1029/2015EO030349.
- Grams, P. E., Tusso, R. B., & Buscombe, D. (2018). *Automated remote cameras for monitoring alluvial sandbars on the Colorado River in Grand Canyon, Arizona. Open-File Report*. Reston, VA. Retrieved from <http://pubs.er.usgs.gov/publication/ofr20181019>
- Hazel, J. E., Topping, D. J., Schmidt, J. C., & Kaplinski, M. (2006). Influence of a dam on fine-sediment storage in a canyon river. *Journal of Geophysical Research: Earth Surface*, 111(1), 1–16. <https://doi.org/10.1029/2004JF000193>
- Hazel, J. E., M. Kaplinski, R. Parnell, K. Kohl, and D. J. Topping (2007), Stage-Discharge Relations for the Colorado River in Glen, Marble, and Grand Canyons, Arizona, 1990-2005, 7.
- Pike, R. J., and S. E. Wilson (1971), Elevation-relief ratio, hypsometric integral, and geomorphic area-altitude analysis, *Bull. Geol. Soc. Am.*, 82(4), 1079–1084, doi:10.1130/0016-7606(1971)82[1079:ERHIAG]2.0.CO;2.
- Rother, C., V. Kolmogorov, and A. Blake (2004), “GrabCut”: interactive foreground extraction using iterated graph cuts, *ACM Trans. Graph.*, 23(3), 309, doi:10.1145/1015706.1015720.
- Sankey, J. B., Caster, J., Kasprak, A., & East, A. E. (2018). The response of source-bordering aeolian dunefields to sediment-supply changes 2: Controlled floods of the Colorado River in Grand Canyon, Arizona, USA. *Aeolian Research*, 32(October 2017), 154–169. <https://doi.org/10.1016/j.aeolia.2018.02.004>
- Schmidt, J. C., Topping, D. J., Grams, P. E., & Hazel Jr., J. E. (2004). System-wide changes in the distribution of fine sediment in the Colorado River corridor between Glen Canyon Dam and Bright Angel Creek, Arizona, (Final Report), 107. Retrieved from [http://www.gcmrc.gov/library/reports/Physical/Fine\\_Sed/Schmidt2004.pdf](http://www.gcmrc.gov/library/reports/Physical/Fine_Sed/Schmidt2004.pdf)

Strahler, A. N. (1952), Hypsometric analysis of erosional topography, *Bull. Geol. Soc. Am.*, 63(November), 1117–1142, doi:10.1130/0016-7606(1952)63[1117:HAAOET]2.0.CO;2.

Topping, D. J., Rubin, D. M., & Vierra Jr., L. E. (2000). Colorado River sediment transport 1. Natural sediment supply limitation and the influence of Glen Canyon Dam. *Water Resources Research*, 36(2), 515–542.

Tusso, R. B., Buscombe, D., & Grams, P. E. (2015). Using oblique digital photography for alluvial sandbar monitoring and low-cost change detection. In *Proceedings of the 3rd Joint Federal Interagency Conference on Sedimentation and Hydrologic Modeling, April 19-23* (pp. 79–85).

# **CHAPTER 3: SEMANTIC SEGMENTATION OF TIME-SERIES IMAGERY USING CONVOLUTIONAL NEURAL NETWORKS: A CASE STUDY OF SANDBARS IN GRAND CANYON**

**R. E. Lima<sup>1</sup>, D. Buscombe<sup>2</sup>, T. T. Sankey<sup>3</sup>, P. E. Grams<sup>4</sup>, E. R. Mueller<sup>5</sup>**

<sup>1</sup> School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ

<sup>2</sup> MARDA Science LLC

<sup>3</sup> School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ

<sup>4</sup> U.S. Geological Survey, Southwest Biological Science Center, Grand Canyon Monitoring and Research Center, Flagstaff, AZ

<sup>5</sup> Department of Geosciences, Southern Utah University, Cedar City, UT

## **KEY POINTS**

- We describe a workflow to train a Convolutional Neural Network end-to-end for semantic segmentation of oblique time series imagery.
- Res-UNet outperformed UNet with and without dropout layers in binary semantic segmentation of oblique sandbar imagery from Grand Canyon.
- 25 labeled images with augmentation were sufficient to train and validate a Convolutional Neural Network and achieve an accuracy of  $F1 = 0.93$  for a single site.

## **KEYWORDS**

Deep Learning, Semantic Segmentation, Time-series, Fluvial Monitoring, CNN

## ABSTRACT

The proliferation of remotely sensed images in hydrologic monitoring has increased the need for automated methods of image segmentation. Deep Learning (DL) and Convolutional Neural Networks (CNNs) are becoming increasingly popular for segmentation tasks within the water sciences. However, detailed and accessible discussions about steps required for data pre-processing, decisions about the size of labeled data sets, and hyperparameter selection are often neglected, making it difficult for non-experts to apply DL to their specific needs. We present a case study describing our workflow of image-processing, labeling data, training, validating, and testing three different CNNs to perform binary semantic segmentation on oblique images of Grand Canyon sandbars. Sandbars are used by thousands of river runners each year, provide substrate for riparian vegetation, and create backwater habitat for native fish species. We found that neural networks with residual connections and UNet architecture outperformed two UNet models without residual connections. We also found that a set of 25 labeled images with augmentation were sufficient to train and validate a CNN capable of accurately ( $F1 = 0.93$ ) segmenting imagery from a test set at a single site. However, 40 images from five different sites with augmentation did not provide the variability needed to train a generalized model capable of segmenting imagery from three or more sites with accuracies  $F1 \geq 0.9$ . We made use of cyclic learning rates and trained networks for 100 epochs. Our results demonstrate that optimizer 'Adam' outperformed 'RMSprop' and 'dice-loss' outperformed 'binary cross-entropy' as a loss function with models without residual connections. We provide a workflow that can be used to apply DL to binary classification problems within the growing body of image data sets in earth and water sciences.

## PLAIN LANGUAGE SUMMARY

We describe a step-by-step workflow for pre-processing data, labeling imagery and training a Convolutional Neural Network (CNN) to perform binary semantic segmentation. We use this UNet model with Residual Connections (Res-UNet) to delineate subaerial sandbars in time series imagery from Grand Canyon. We provide discussion and insight into network architecture, training sample size, and hyperparameter choice. Such methods can be adapted for other time series of coastal or fluvial remotely sensed images.

### 3.1. INTRODUCTION

Sandbars are an important component of the physical and biological systems in riparian areas around the world and specifically in the Grand Canyon. Sandbars in Grand Canyon serve as high-elevation storage locations for fine ( $<2mm$ ) sediment, create backwater habitat for native fish (Valdez & Ryel, 1995; Converse et al., 1998; Valdez et al., 2012), and provide a source of aeolian sand that protects archaeological sites and mitigates gully erosion (Draut & Rubin 2007; Sankey et al., 2018). Sandbars are the preferred camping location for nearly 25,000 annual river users (Kearsley et al., 1994; NPS, 2006). Sandbars in Grand Canyon have decreased in size and extent beginning with the 1963 closure of Glen Canyon Dam (Dolan et al., 1974). To mitigate sandbar erosion, controlled floods have been used to re-distribute sand deposited in the channel by tributary flooding to high elevations within eddies (Grams et al., 2015). Images from a network of remote cameras are one of the data sources used to monitor the effects of controlled floods (Grams et al., 2018). These cameras capture five oblique images per day at 43 different sandbar monitoring sites throughout Grand Canyon, resulting in a large data set that currently contains over 1,000,000 images.

The promise of this collection of daily images is the measurement of short-term changes in fine sediment storage and correlation of those changes to particular flow patterns, which is of interest to river managers. However, quantification of sandbar change from a dataset of nearly a million images requires an automated method of accurate segmentation of pixels containing subaerial sand in oblique imagery, which often contains significant shadows, solar glare, and changing water color.

Deep Learning (DL) has increasingly become an important tool in the arsenal of water resource scientists (Shen, 2018). DL is a subset of the larger field of artificial intelligence, which performs automatic feature extraction. Contrast DL with traditional Machine Learning systems, which required features or patterns of interest to be define beforehand (Arif et al., 2020). In DL, the practitioner provides input data and ground-truth data and DL networks “learn” which features or patterns to utilize. This learning process automatically tunes internal adjustable parameters, often called “weights” that define the input-output function. Such parameters would have to be specified beforehand in traditional machine learning methods (Lecun et al., 2015). Convolutional Neural Networks (CNNs) are a class of artificial neural networks (ANNs) within the field of DL. CNNs are designed to perform classification tasks on image-like arrays through a filtering process known as kernel convolution (Lecun et al., 2015). During kernel convolution, image arrays are successively down-sampled creating progressively smaller feature maps that allow for the abstraction of high-level features, object detection, and classification (Lecun et al., 2015). The word ‘deep’ in this context refers to the inclusion of one or more hidden layers within the neural network. Advances such as the Fully Convolutional Network (FCN) (Long et al., 2015), and encoder-decoder architectures (Ronneberger et al., 2015; Badrinarayanan et al., 2017) have improved CNN performance on pixel-wise classification tasks, known as semantic segmentation. CNNs have recently been used to segment water bodies (Li et al., 2018; Chu et al., 2019; Ling et al., 2019; Nath et al.,

2019), clouds (Guo et al., 2020), glaciers (Baumhoer et al., 2019), ocean eddies (Lguensat et al., 2018), underwater vegetation (Burguera, 2020), and a host of other natural textures in remotely sensed imagery (Buscombe et al., 2018; Buscombe & Carini, 2019).

Fluvial landscapes are currently undergoing rapid change due to human activities (Piégay et al., 2020). CNNs offer tremendous potential in the burgeoning field of fluvial remote sensing to monitor and assess those changes (Marcus & Fonstad, 2010; Carbonneau et al., 2012; Shen, 2018; Piégay et al., 2020). The availability of low-cost remote camera systems and UAVs has led to the proliferation of large fluvial and coastal monitoring imagery datasets (Bertoldi et al., 2012). In this paper, we apply DL to one of the first and longest-running fluvial monitoring networks, which has been collecting oblique images of sandbars in Grand Canyon since 1990 (Dexter et al., 1995).

While attempting to apply DL and CNNs to our dataset, we found that many of the practical questions for the application of DL in this context remained unanswered. Most of the CNN studies focus on evaluating network performance on very large, pre-labeled competition datasets containing street scenes, animals, or anthropogenic objects (Cheng et al., 2017; Garcia-Garcia et al., 2017). CNN used in the natural sciences most often involves segmenting tiles from satellite imagery (Chu et al., 2019; Ling et al., 2019; Carbonneau et al., 2020; Guo, Y., et al., 2020; Jiao et al., 2020) with a few notable exceptions (Buscombe et al., 2018; Buscombe & Carini, 2019; Nath et al., 2019; Burguera, 2020). Little work has been done on oblique land- or UAV-based time-series imagery (Carbonneau et al., 2020).

Labeling imagery for training, validation, and testing neural networks is a tedious and time-consuming process. Therefore, we wanted to determine the minimum number of labeled images required to train a CNN that can accurately segment time-series imagery from a single site of interest (a sandbar), and from multiple sandbars with a range of different scales and imaging angles. We compare the performance of three different neural networks, all

variations of the symmetrical encoder-decoder type UNet architecture, which was originally developed for segmenting medical imagery, but has since been used to segmented sea ice flows, water, clouds, and to delineate coastlines (Ronneberger et al., 2015; Yan and Dong, 2019; Guo, H., et al., 2020; Jiao et al., 2020; Heidler et al., 2021; Nagi et al., 2021). Given these recent applications, we use it here to delineate sandbars. If this application is successful along the Colorado River, it could be widely applied to other dam-impacted systems around the world.

In a UNet architecture, down-sampled image arrays from the contracting (encoder) path are concatenated with up-sampled layers from the expanding (decoder) path acting as long-skip connections recovering spatial information lost during down-sampling (Drozdal et al., 2016). One key advantage of UNet models is that they can be trained end-to-end using relatively few training images, if image augmentation is applied to artificially increase the size of the training dataset (Ronneberger et al., 2015).

The addition of residual connections to a CNN was first described by He et al., (2016). In a UNet with Residual connections (Res-UNet), each block of layers is replaced with a residual block containing batch normalization and activation layers creating shortcuts or short-skip connections. Such short-skip connections effectively simplify the loss landscape (Li et al., 2018), allowing for much deeper networks to be trained, and addressing the problem of vanishing gradients when backpropagating across many layers (Drozdal et al., 2016; He et al., 2016). We examine the effect of residual blocks, dropout layers, image augmentation, optimizer choice, and loss function choice on semantic sandbar segmentation accuracy. We address the following specific questions:

- Which of the three CNN architectures UNet-1, UNet-Drop, or Res-UNet performs best after 100 training epochs?

- Which loss function and optimizer result in the highest model accuracy after 100 training epochs?
- How many labeled images are needed to train and test a site-specific CNN with a binary segmentation accuracy of  $F1 \geq 0.9$ ?
- How many labeled images, and from how many sites, are needed to train a generalized model with a binary segmentation accuracy of  $F1 \geq 0.9$ ?

This paper presents the entire workflow from image pre-processing, labeling, and CNN training, to model evaluation for a time series of fluvial, remotely sensed oblique imagery. All of our training and testing was performed on a relatively modest desktop computer with a GPU or in Google Colab notebooks using Python, Keras, and Tensorflow. This case study can serve as a starting point for other water scientists in attempting to apply DL to fluvial or coastal image data sets.

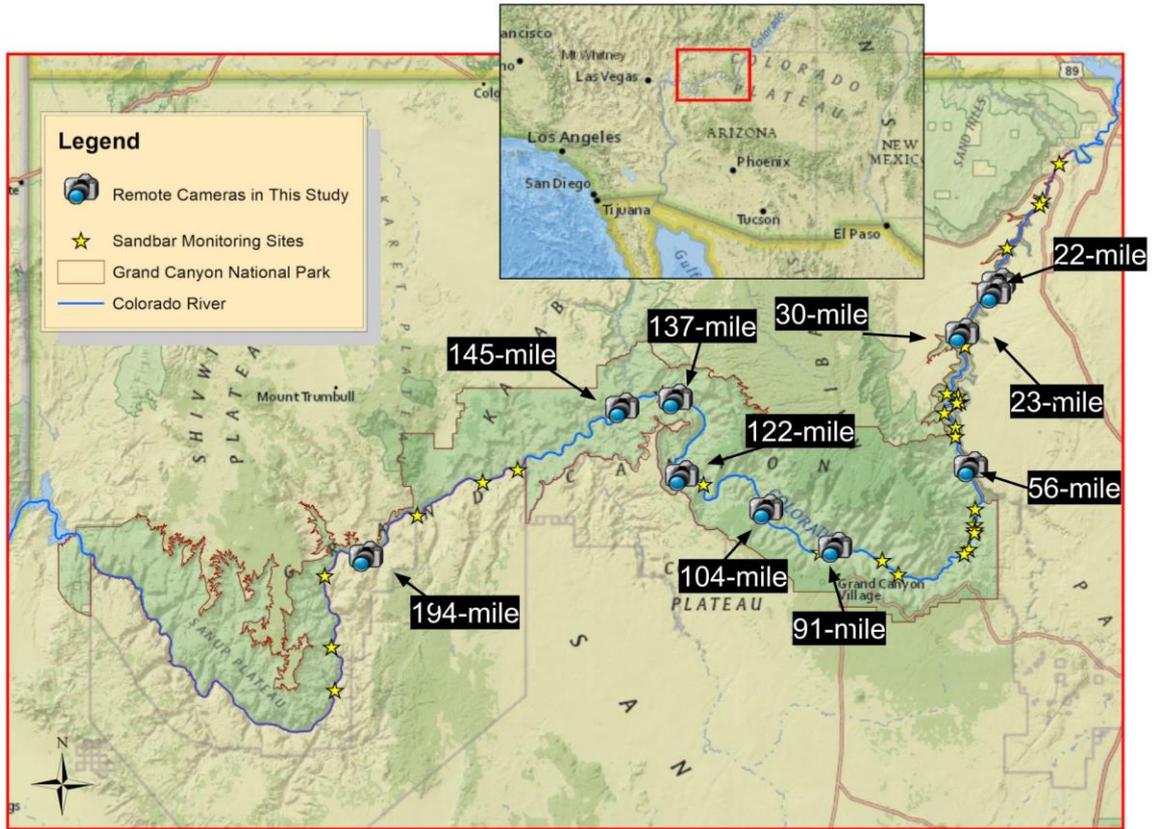
### **3.2. DATA**

Beginning in the 1990s, a network of remote camera systems was installed in Grand Canyon, originally containing film cameras capturing a single image per day at 43 different sandbar monitoring sites. Beginning in 2008, film cameras were replaced with digital cameras capturing up to five images per day at each site. Each of these camera systems contains a solar panel, data logger, 12V battery, digital single lens reflex camera, memory card, and an 18-55mm zoom lens mounted in a weatherproof box. Greater detail about these systems is provided by Bogle et al. (2013), and the specific details of their deployments in Grand Canyon are described by Grams et al. (2018).

There is significant intrasite and intersite variability within the Grand Canyon remote camera data. One of the motivations of this research was to understand how many labeled images from a single site are needed to provide a site-specific CNN with the requisite

intrasite variability to generalize and allow for segmentation of unseen images from that site. In addition, we attempt to understand how many labeled images, from how many different sites, are needed to provide a CNN with the intersite variability necessary to generalize to unseen sandbar sites. Intrasite variability occurs because of annual changes in solar radiation angle, flow, and suspended sediment concentration. Intersite variability is caused by differences in site size, terrain, and camera accessibility. Cameras tend to be placed across the river channel from monitoring sites on talus slopes. Changes in river width, surrounding bedrock type, and vegetation add to this variability.

Sandbar monitoring sites in Grand Canyon are named for the river mile, downstream from Lees Ferry, where they are located. One site of particular interest, which is the subject of the single-site model, is 30-mile. This site was chosen for testing because we have specific morphological questions about this site, and it contains unobstructed views of mostly open sand. Nine additional sites were chosen for CNN testing to provide a range of intersite variability. Imagery from four of those sites and 30-mile are used for model training and validation, and the remaining five are used only for testing model accuracy. A random sample of imagery stratified by month was gathered for each site (up to 125 from each site), and additional imagery from infrequent controlled floods were included. Images are 2592 by 3888 pixels, (.JPG) format with three bands (R-G-B). Imagery depicts a ground area ranging from 0.25 – 2 *hectares* depending on the height and angle of the camera. A map showing the distribution of sites in this study is shown in Figure 3.1.



**Figure 3.1** Map of the study area in Grand Canyon National Park. Data from ten remote camera sites are used in this study and shown on the map. Below the map is a diagram showing how data were split into the various train and test image datasets described in section 3.2.4 below

### **3.3. METHODS**

#### **3.3.1. Image pre-processing**

Oblique images show a 3-dimensional scene projected onto a 2-dimensional surface, where equally sized pixels represent different distances in 3-dimensional space. In order to measure sandbars in oblique imagery, images must be warped such that the pixels in the image plane, represented by X, Y coordinates, have an equal area in a projected coordinate system. We use NAD 83, Arizona Central State Plane FIPS 0202 meters, where locations are defined by Eastings and Northings (E, N). We, hereafter, refer to coordinates in this projected coordinate system as real-world coordinates

Warping images, a process known as rectification, involves several steps. First, camera-lens distortions are removed through camera calibration. Second, a transformation matrix or homography is calculated using ground control points (GCPs) visible in a reference image for each site and surveyed to determine their location in real-world coordinates. Third, registration is performed to map all images in each set to the reference image containing GCPs for which the homography is calculated. Finally, the homography is applied to warp each registered and undistorted image to a nadir view with equal area pixels corresponding to the real-world coordinates.

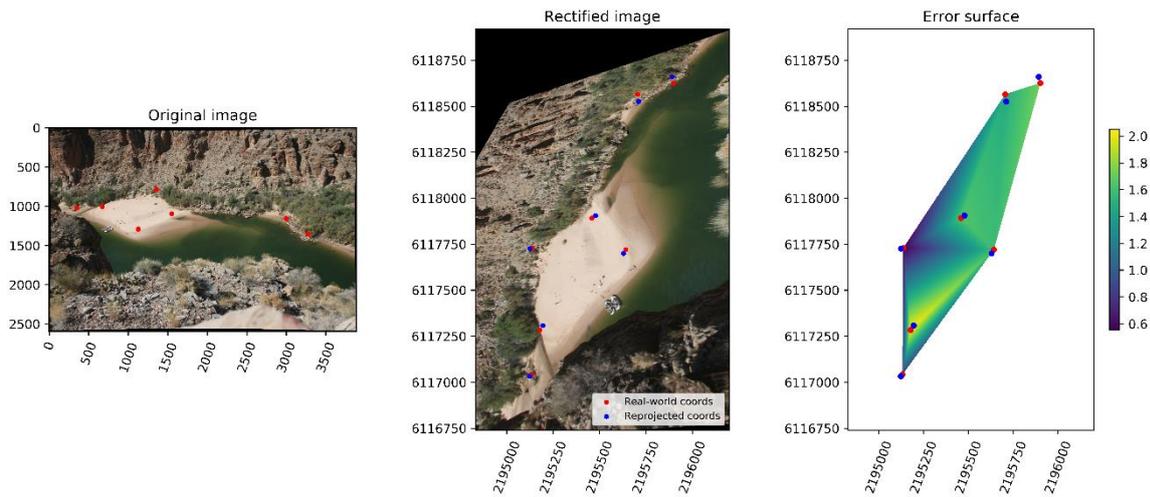
#### **3.3.2. Removing distortion**

Camera calibration is the process of estimating the intrinsic matrix of a camera-lens system (Hartley & Zisserman, 2003). This step is necessary to: correct for tangential and radial distortion inherent in the camera-lens system, determine the focal length in pixels, and to find the principal point or optical center of the image. Ideally, cameras with lenses of a fixed focal length are calibrated in the lab before deployment. The array of remote cameras in

Grand Canyon contains a variety of different camera models and zoom lenses. Focal lengths cannot be fixed beforehand and are adjusted in the field to fit each camera location. Camera calibration took place in situ using a 10x10 checkerboard pattern printed on aluminium. Between 15 and 30 images were taken of the calibration checkerboard at different distances and angles. The camera calibration toolbox in Matlab was used to determine the intrinsic matrix for each camera deployment (Bouquet, 2015). The intrinsic matrix and distortion parameters were used with the `cv2.undistort()` module in OpenCV 4.2 and Python 3.6 to remove distortion from each image.

#### **3.3.2.1. Rectification**

To rectify imagery from sandbar sites, reference images were taken during annual sandbar monitoring trips in 2017 and 2019. Black and white panels, 1x1m in dimension, were photographed from the remote cameras and surveyed with a total station tied to permanent control points with known real-world coordinates. The `cv2.findHomography()` function in the OpenCV module was used to develop a perspective transformation between the X, Y image coordinates of GCPs and their real-world coordinates by minimizing the back-projection error with a simple least-square algorithm (Bradski, 2000). For each homography, an error surface was generated which displays the error between the reprojected image coordinates and the real-world coordinates of each GCP and then values between GCPs were interpolated to provide an estimate of rectification error across each sandbar site as shown in Figure 3.2.



**Figure 3.2** Original reference image on the left for the 30-mile sandbar monitoring site, with the rectified image in the center. Reprojected image coordinates for GCPs can be seen in blue, real-world coordinates in red. The error surface of interpolated error (in meters) between the reprojected and real-world GCP points is shown on the right.

### 3.3.2.2. Registration

Remote cameras deployed in the field are subject to changes in temperature and pressure as well as potential disturbance from, falling rock, wildlife, or backcountry travelers. Remote cameras undergo routine bi-annual maintenance where memory cards are exchanged, and faulty batteries are replaced. Maintenance of this sort and the other disturbances mentioned can lead to camera movement. Each time the camera moves, the relationship between the image plane and the scene changes. In order to minimize the number of rectifications with surveyed GCPs needed for each site, registration, or image alignment, was conducted. Registration is the process of mapping images to a common reference image so that all images share the same field of view. We employed three different registration methods depending on the needs of each specific site.

The first two methods involve correspondences in the Fourier domain. Fast Fourier transform-based (FFT) methods use low-level features such as edges and corners to match images that have been translated and rotated with respect to one another (Reddy & Chatterji, 1996). The first method uses the 'imreg\_dft' module developed by Goklhe (<https://pypi.org/project/imreg/>) and implemented in Python 3.6. The imreg\_dft method is robust and

works most of the time. However, it takes between two and four minutes per image. The second method is noticeably faster, 10 - 60 seconds per image, but less robust to significant changes in illumination and may fail in images with sharp black borders introduced when radial distortion is removed at remote cameras with shorter focal lengths (18 - 24mm). This second method is also based on FFT and is taken from Grams et al. (2018). The third method utilizes the `cv2.findTransformECC()` and `cv2.warpPerspective()` function in the OpenCV module in Python. This third method makes use of the Enhanced Correlation Coefficient algorithm, which is invariant to photometric distortions in contrast and brightness (Evangelidis & Psarakis, 2008). The advantage of this method is that it uses a homography rather than a translation and thus it can deal with warping and scaling associated with small changes in focal length that occur due to lens creep or camera maintenance. This third method is the slowest taking between three and ten minutes per image.

### **3.3.2.3. Labeling imagery**

Labeling imagery is a necessary but time-consuming process needed to generate train, validation, and test data for CNNs. We utilized the labeler found in [https://github.com/dbuscombe-u\\_sgs/dl\\_tools/](https://github.com/dbuscombe-u_sgs/dl_tools/), a tool developed by Buscombe et al. (2018), which accepts manual annotations and expands those annotations using a conditional random field (CRF), a type of classification and graphical modeling known as structured prediction, to label every pixel in the image (Lafferty et al., 2001). The labeling process takes between two and fifteen minutes per image labeling the 475 images in the various test and train sets took approximately 75 hours. Therefore, one of the goals of this study was to determine the minimum amount of training data needed to accurately segment imagery.

This case study involves binary segmentation with a target class (subaerial sand) and all else not in the target class. However, the networks described here can easily be adapted to multiclass problems. Our first attempts at segmentation involved multiple classes, segmenting

pixels into sand, water, boats, tents, vegetation, glare, rocks, sky, etc., to obtain information about vegetation or sandbar use from the data, but eventually a binary approach was selected because there were significant class imbalance issues with small or rarely occurring classes or classes not shown at some sites. Increased numbers of classes led to increased labeling times as well. Increased number of classes are also known to significantly reduce classification accuracies from greater confusion among classes (Congalton and Green, 2019). Labeling an image with five classes took approximately 3- 5 minutes more per class. Pixels within images were labeled as either '1' = *sandbar* (the subaerial or emergent portion of the bar), or '0' = *not sandbar*. The outputs of the labeling process are image-label pairs. Images are in the (.JPG) format, and labels are (.PNG) binary masks containing values of 0 and 255. To input the image-label pairs into the CNN for training and testing, the images and labels were both converted into arrays of unsigned integers (*uint8*). Pixel values in the label array (0-225) were converted to values of 0 or 1, and both the images and labels were resized to 128 x 128 pixels for faster convolution to reduce computing power needed, and to reduce the size of imagery and masks and storage space required. Storage and organization of imagery data is an important concern. During each step of image processing copies of each image were made so that the accuracy of each step could be examined. With a dataset of over 1,000,000 images, each new copy of the image and mask multiplied the amount of storage space required to store the dataset.

### **3.3.3. CNN semantic segmentation**

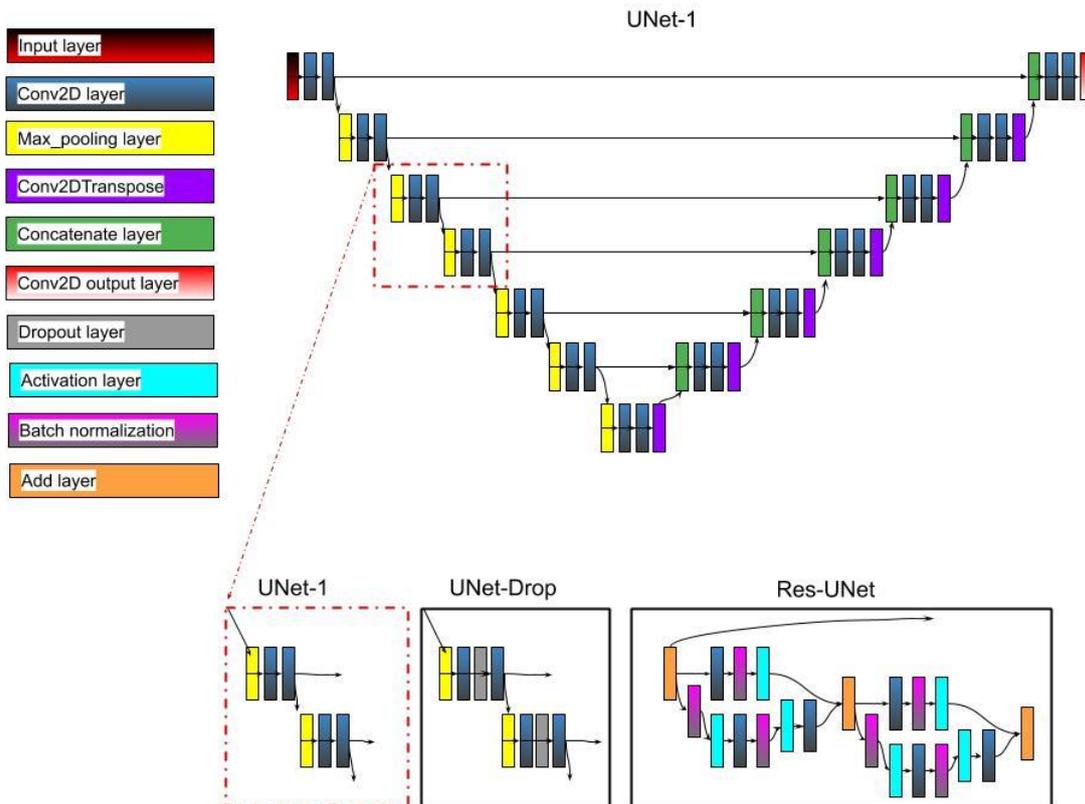
#### **3.3.3.1. Neural network architectures**

We tested three different CNN architectures. All three utilize the UNet encoder-decoder structure. The first model, which we call UNet-1, contains six blocks in the contracting path, each with two Conv2D layers and one Max Pooling layer. The expansive

path also contains six blocks each with two Conv2D layers, one concatenated layer with input from the contracting path, and one conv2D transpose layer before a final Conv2D prediction layer.

The second model we tested 'UNet-Drop' has the same structure as UNet-1 except that it contains dropout layers between each of the Conv2D layers in each block of the encoder and decoder paths. The purpose of a dropout layer is to randomly drop out a given percentage of the data between the convolutions to reduce overfitting and make the model more generalizable (Srivastava et al., 2014). We varied the dropout proportion from  $0.1$  to  $0.3$  throughout UNet-Drop. We would expect a certain amount of overfitting to benefit a model trained and used for segmenting a single-site. However, when it comes to training a multi-site sandbar segmentation model, we expected the dropouts included in UNet-Drop to improve generalizability and accuracy on imagery from unseen sites.

The third network type we tested 'Res-UNet' makes use of residual connections. Residual blocks create short-skip connections as described above. The differences between the three models are highlighted in Figure 3.



**Figure 3.3** A simplified diagram of UNet-1 showing the U-shaped structure, the blocks of convolutional and pooling layers on the encoder path (top left side), and concatenated layers along with conv2DTranspose layers along the decoder path (top right side). Two blocks of the contracting path of all three models tested in this research are shown to highlight their differences (lower). UNet-1 on the left, UNet-Drop in the middle, and Res-UNet on the right. UNet-Drop has the same structure as UNet-1 except for the addition of dropout layers and Res-UNet contains short-skip connections using batch normalization and activation layers. UNet-Drop has the same structure as UNet-1 except for the addition of dropout layers and Res-UNet contains short-skip connections using batch normalization and activation layers.

### 3.3.3.2. Hyperparameters

In the context of DL, hyperparameters are user-defined settings that control how CNNs learn. Learning in this context is an attempt to minimize *loss*--or the difference between predicted classifications and ground-truth data. The loss function is the hyperparameter that defines how loss is calculated, a performance metric. Neural networks may be made up of tens of thousands of different weights and biases that are gradually adjusted as the network is trained and features or patterns are learned. Loss functions are high-dimensional and non-convex, they can be visualized as a landscape where areas of high-

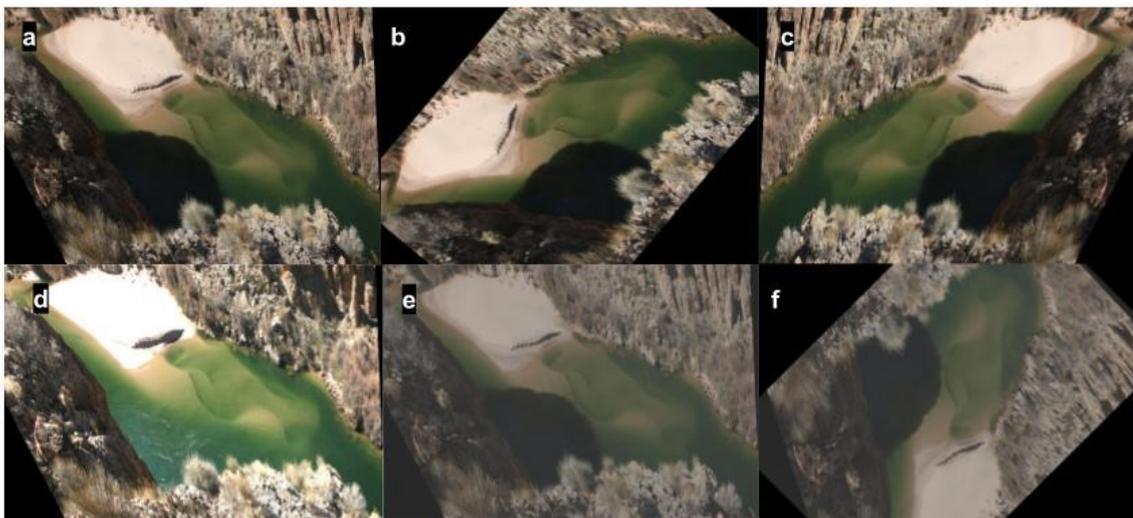
elevation have high loss and areas of low elevation have low loss, the location of any point in this landscape is defined by values of weights and biases (Li et al., 2018). With each pass through the training data, known as an epoch, the loss is re-calculated based on the chosen loss function (Burkov, 2019). The *optimizer* is the hyperparameter that acts as a solver for this minimization problem, providing instructions on which direction to move through the loss landscape to find the global minimum in the landscape rather than local minimums (Burkov, 2019). *Learning rate* is the hyperparameter that specifies the size of the steps you take while moving through the loss landscape (Burkov, 2019).

For this research, we varied the learning rate cyclically during network training between a maximum learning rate of  $0.0001$  and a minimum learning rate of  $0.000001$ . Using cyclic learning rates removes the need to tune for this hyperparameter provided that the range is reasonable and can improve classification in fewer iterations relative to fixed learning rates (Smith, 2017). We tried two popular optimizers; *RMSprop* and *Adam*, which are both variants of Stochastic Gradient Descent (Burkov, 2019). A discussion of the differences between these two can be found in Ruder (2016). We tried two different loss functions, which are commonly used with UNet architectures, trained for binary segmentation; *binary cross-entropy* and *dice-loss* (Ronneberger et al., 2015; Drozdal et al., 2016; Sudre et al., 2017; Diakogiannis et al., 2020). We chose a relatively small batch size of five, because of the memory limitations of our PC. Breuel (2015) tested a range of batch sizes between  $1$  and  $1000$  and found that, with CNNs containing ReLu activation functions, small batch sizes yielded the lowest test set errors. In all three of our models, ReLu activation functions were used within the convolutional blocks and a sigmoid activation function was used in the final, output convolutional layer. During initial testing, our model loss and accuracy began to plateau after about 70 epochs and training longer than 100 epochs did not improve model loss or accuracy.

### 3.3.3.3. *Image augmentation*

Image augmentation is commonly used to artificially increase the number of images in small CNN training datasets to increase variability and reduce overfitting (Ronneberger et al., 2015; Xu et al., 2016; Perez & Wang, 2017). Augmentation reduces overfitting by preventing the network from learning features based on their scale or relative location by flipping, scaling, and rotating them. Keras contains methods for on-the-fly augmentation where augmentations are randomly applied to imagery in-place as training occurs. However, we chose to create separate augmented data sets for the sake of repeatability and control.

We applied random adjustments to contrast and brightness, random rotations, and flips using the Python Image Library (PIL) to create larger augmented data sets. Examples of such augmentations are shown in Figure 3.4. Researchers must be careful to apply augmentations that are realistic for the data and represent the variability within the larger image data set.



**Figure 3.4** Examples of image augmentation performed on a rectified image from the 30-mile site. The original image (a); the rotated image (b); mirrored image (c); random brightness adjustment (d); contrast augmentation (e); mirrored, rotated, contrast augmented image (f). Images undergo a combination of these augmentations to create a much larger dataset of images, which should be representative of conditions in imagery throughout the dataset.

#### 3.3.3.4. *Training, testing, and validation data*

One of the motivations of this research was to determine the number of labeled images needed to train, validate, and test a site-specific model for sandbar segmentation at the 30-mile site, and how many labeled images were needed to train, validate, and test a generalized segmentation model which could be used on many, or all of our sites. To address these research questions, several test and train sets were generated. Image-label pairs were randomly selected from the full set of labeled images; the test sets were held out and only used for final model comparison. *One\_site\_test* contains 25 images or 20% of the total (125) image-label pairs from the 30-mile site.

To test the generalized sandbar model, we created three test sets: *Three\_site\_test*, which contains 10 image-label pairs from each of three sandbar sites (30 total); 30-mile, 22-mile and 23-mile (figure 3.1). *Five\_site\_test* contains 10 image-label pairs from five sites including the aforementioned three sites with imagery from 56-mile and 145-mile added in (50 total). Image-label pairs from all five of these sites are included in the training data. *Ten\_site\_test* contains 10 image-label pairs from the aforementioned 5 sites with additional image-label pairs from 91-mile, 94-mile, 104-mile, 122-mile, and 137-mile included (100 total). Imagery from these last five sites are not included in training data. This final test set is used to evaluate the performance of the generalized sandbar segmentation model on unseen sandbar sites.

In order to determine the minimum number of images needed to train a single-site neural network for 30-mile, we created the following training data sets; *One\_site\_25*, *One\_site\_25\_aug*, *One\_site\_50*, *One\_site\_50\_aug*, *One\_site\_100*, *One\_site\_100\_aug*. These contain 25, 50, and 100 image-label pairs with and without image augmentation, respectively, from the 30-mile site. These training sets allow us to investigate the effects of image

augmentation and increasing the number of training image-label pairs, on segmentation accuracy.

To understand how many images, and from how many different sites, a generalized neural network needs to accurately segment imagery from five or more sites, we created the following training data sets; *Three\_site*, *Three\_site\_aug*, *Five\_site*, and *Five\_site\_aug*. Those training sets contain: 40 image-label pairs from each of the sites in *Three\_site\_test* (120 total), and an augmented version of that train set, and 30 image-label pairs from each of the sites in *Five\_site\_test* (150 total) along with an augmented version of that train set. A visual representation of which sites are in the various sets can be found in Figure 3.1 (above).

Image-label pairs within the training datasets are used for CNN training and validation. For each training set, image-label pairs were further subdivided using  $k$ -folds cross-validation into ( $k=5$ ) ‘Fold-n-train’ ( $0.8$  of train images) and ‘Fold-n-validation’ ( $0.2$  of train images) sets. During model training, the Fold-n-train image-label pairs are used in the learning process to fit model weights, while the Fold-n-validation image-pairs provide an unbiased evaluation of the model fit at the end of each training epoch and assist in hyperparameter selection. For each of the five folds of the data, a separate model is trained, then all five fold-specific models and an ensemble are used to evaluate model performance on the test sets. Model performance is evaluated based on the average accuracy of the segmentation results across folds and using the ensemble model on each of the training sets, we use F1-score as the metric for segmentation accuracy (Goodfellow et al., 2016; Burkov, 2019; Carbonneau et al., 2020) (Equation 3.1).

$$\text{(Eq. 3.1) } F1 = 2 \frac{\textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

*FI*-score is also known as the dice similarity coefficient and is numerically equivalent to the dice-loss function we tested during model training. We chose the relatively high accuracy threshold of  $FI \geq 0.9$  to determine segmentation success. For any single image, a value of 75% for overall accuracy is considered acceptable in most remote sensing applications (Congalton & Green, 2019) and (70 – 80%) accuracies are common in fluvial remote sensing studies (Carbonneau et al., 2020). However, when trying to quantify short-term erosion and deposition between images, the errors compound when detecting change between images and thus we chose a higher accuracy threshold (Congalton & Green, 2019).

### **3.3.3.5. Experiments**

Loss function choice and optimizer choice were examined by training CNNs using the single-site training sets (One\_site\_ ...) and evaluating their performance on the validation data. All three UNet models were trained using the single-site training sets and evaluated with the single-site test set (One\_site\_test). Each of the different models were also trained using the six multi-site training sets and evaluated for their performance on the three multi-site test sets.

### **3.3.3.6. Computing hardware and software**

CNNs were built and implemented using the Keras API, Tensorflow 2.1, and Python 3.6. Many of the faster training models with smaller training and validation sets were trained using Google ColabPro and their Tesla P100 PCIE GPUs. However, due to run-time restrictions on Google Colab and long training times for augmented image data sets, a Quadro P2000 GPU was utilized on a Dell precision 3630 desktop PC for all training with the larger augmented data sets.

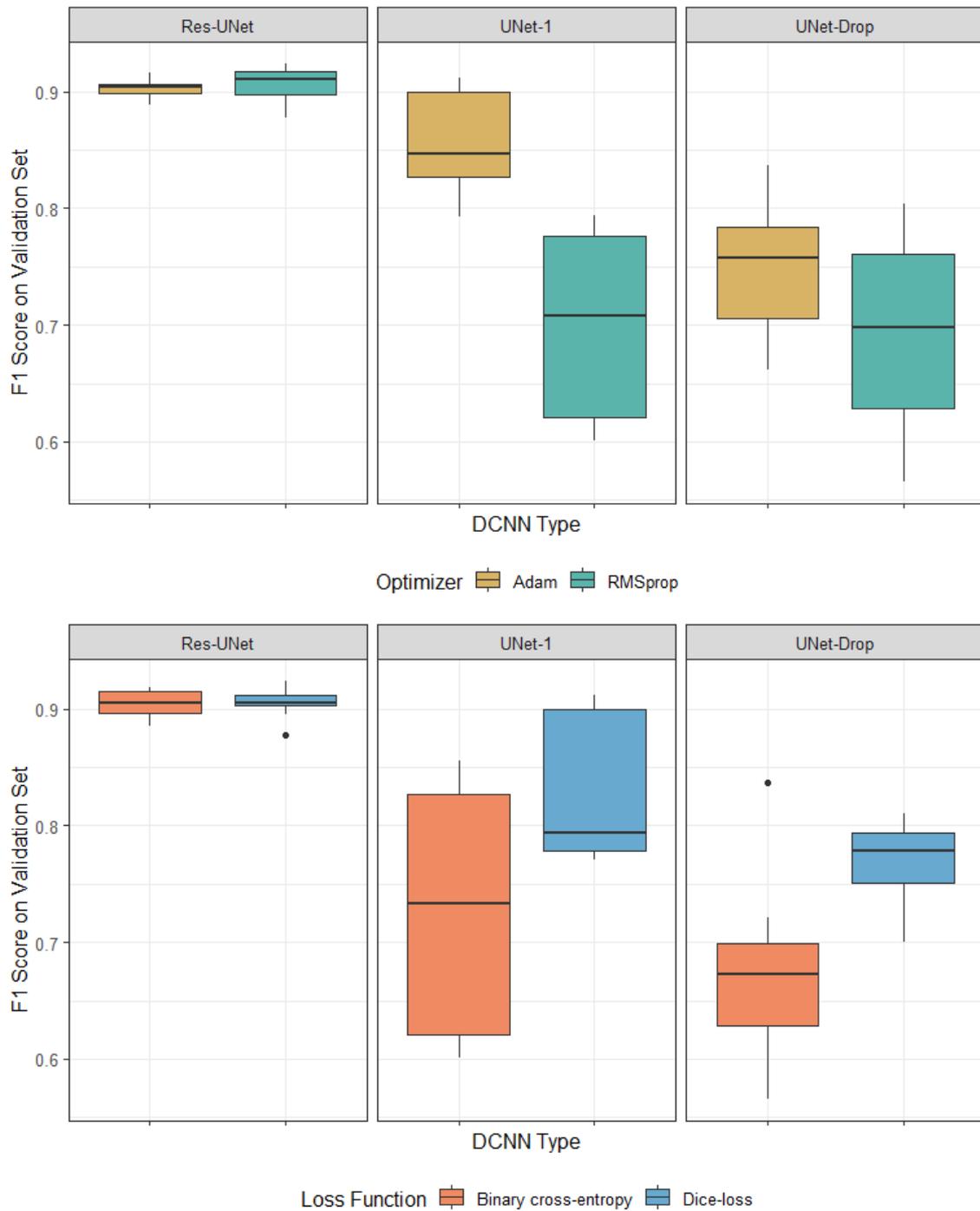
## 3.4. RESULTS

### 3.4.1. Optimizer and loss function

We trained all three models using the 100-image, One\_site\_100 train set to evaluate optimizer performance on the validation data. We found that using Dice-loss as the loss function and Adam as the optimizer resulted in higher  $F1$ -scores on the validation data after 100 epochs with the UNet-1 and UNet-Drop models when averaged across all five folds of the data. The accuracy of the Res-UNet model was not significantly affected by loss function or optimizer choice as shown in Figure 3.5.

### 3.4.2. Architecture performance

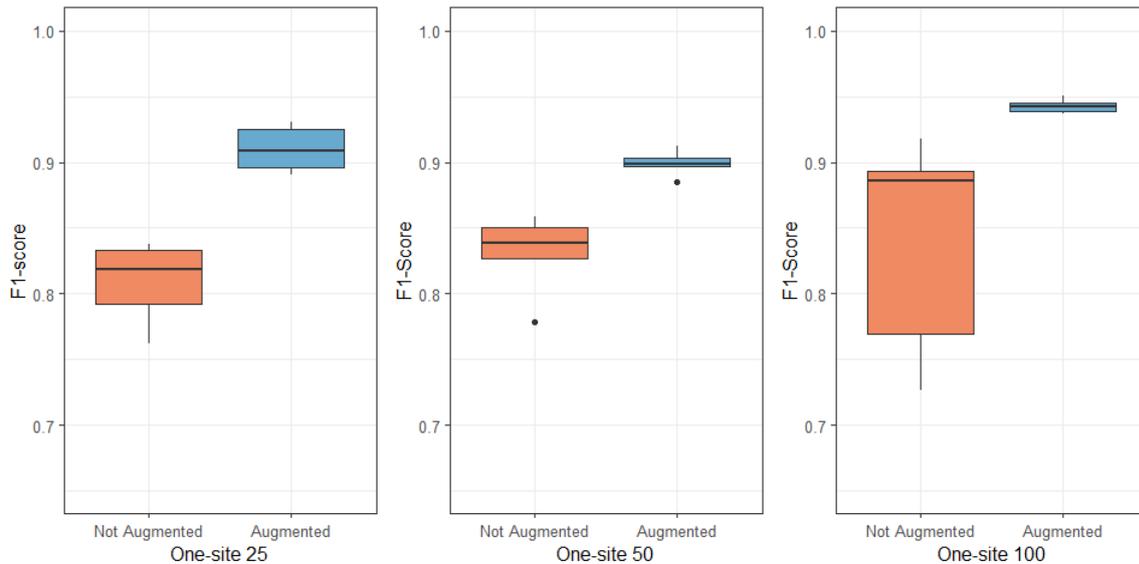
We compared the model performance of each architecture on the same training set, One\_site\_100\_test, to determine which model should be used for single-site segmentation. Res-UNet achieved a marginally higher mean accuracy ( $F1 = 0.89$ ) across five folds after 100 epochs than UNet-1 ( $F1 = 0.87$ ) and significantly better than UNet-Drop ( $F1 = 0.76$ ). When tested on training data including augmented imagery from five sites, similar accuracies were observed Res-UNet ( $F1 = 0.87$ ); UNet-1 ( $F1 = 0.86$ ) and UNet-Drop ( $F1 = 0.79$ )



**Figure 3.5** *F1* scores across five folds of the one-site 100 train and validation set. The choice of loss function and optimizer does not seem to have a significant effect on the performance of Res-UNet after 100 epochs. The use of optimizer Adam does significantly increase performance of UNet-1 after 100 epochs and the use of dice-loss significantly improves the performance of Unet-Drop after 100 epochs.

### 3.4.3. Image augmentation

Models trained with augmented datasets had higher accuracies and lower variability in the F1 scores across folds on the same test sets when compared to models trained with non-augmented data. Figure 3.6 shows the difference in model performance of the Res-UNet models on the single-site training data with, and without, image augmentation.



**Figure 3.6** Accuracy of all five folds of Res-UNet CNNs trained with the 30-mile one-site train data sets containing 25, 50, and 100 images and corresponding augmented train data sets and tested on the One-site test set.

### 3.4.4. Evaluating training sample size

The results in Table 3.1 show that with as few as 25 labeled training images, with augmentation applied, a Res-UNet model could achieve the desired accuracy on a test set with imagery from the same site. Without augmentation, 100 images were sufficient to train a Res-UNet model to achieve the desired accuracy on the single-site test set: `One_site_test`.

None of our multi-site training sets were able to achieve the desired accuracy when evaluated on test sets containing image-label pairs from more than one site as shown in rows 12-22 of Table 3.1. Though the `Five_site_aug` train set, which contains 30 image-label pairs from each of five sites, with augmentation applied, was able to reach accuracies of  $F1 = 0.89$  when evaluated against test sets containing three and five sites whose imagery was also

included in the training sets (Rows 14,15 &17). None of our training sets were able to generalize to sites outside of the training data with the desired accuracy as shown in rows 20-22 of Table 3.1, with the Ten\_site\_test.

**Table 3.1** Selected Accuracies for CNNs trained on various train and test sets, all models shown were trained with optimizer Adam, and loss function dice-loss. F1 scores  $\geq 0.9$  are in bold.

#	DCNN Type	Train Set	Test Set	F1 Ensemble	F1 fold-Mean
1	Res_UNet	one_site_25	one_site_test	0.83	0.64
2	Res_UNet	one_site_25_aug	one_site_test	<b>0.93</b>	<b>0.90</b>
3	Res_UNet	one_site_50	one_site_test	0.89	0.87
4	Res_UNet	one_site_50_aug	one_site_test	<b>0.91</b>	<b>0.90</b>
5	Res_UNet	one_site_100	one_site_test	<b>0.92</b>	<b>0.90</b>
6	Res_UNet	one_site_100_aug	one_site_test	<b>0.95</b>	<b>0.94</b>
7	Res_UNet	three_site_aug	one_site_test	<b>0.93</b>	<b>0.93</b>
8	Res_UNet	three_site_aug	one_site_test	0.88	0.88
9	Res_UNet	five_site_aug	one_site_test	<b>0.93</b>	<b>0.90</b>
10	UNet_1	five_site_aug	one_site_test	<b>0.91</b>	<b>0.90</b>
11	UNet_Drop	five_site_aug	one_site_test	0.85	0.82
12	Res_UNet	three_site_aug	five_site_test	0.80	0.77
13	Res_UNet	three_site_aug	ten_site_test	0.63	0.60
14	Res_UNet	five_site_aug	three_site_test	0.89	0.87
15	UNet_1	five_site_aug	three_site_test	0.89	0.87
16	UNet_Drop	five_site_aug	three_site_test	0.81	0.80
17	Res_UNet	five_site_aug	five_site_test	0.89	0.87
18	UNet_1	five_site_aug	five_site_test	0.88	0.86
19	UNet_Drop	five_site_aug	five_site_test	0.80	0.79
20	Res_UNet	five_site_aug	ten_site_test	0.71	0.68
21	UNet_1	five_site_aug	ten_site_test	0.70	0.68
22	UNet_Drop	five_site_aug	ten_site_test	0.65	0.64

### 3.5. DISCUSSION AND CONCLUSIONS

When trained and tested on the same imagery using the same hyperparameters, the Res-UNet network clearly performed better than UNet-1 and UNet-Drop on a single-site One\_site\_test set. This is consistent with results from Li et al., (2018) that points out that,

when networks become sufficiently deep, loss landscapes can become chaotic and difficult to train; the residual connections effectively flatten the loss landscape. Drozdal et al. (2016), found that networks with both long and short skip connections, like Res-UNet, performed better and converged faster than networks without short-skip connections.

Our results show improved performance when using dice-loss as the loss function in UNet-1 and UNet-Drop, and similar results were found by Drozdal et al., (2016) who showed cleaner segmentation results when comparing dice-loss and binary cross-entropy. We also show improvement when using Adam over RMSprop as the optimizer with UNet- 1 and UNet-Drop. However, Res-UNet does not appear to be as sensitive to optimizer or loss function choice, possibly due to a simplified loss landscape produced by (short-skip) residual connections. A simplified loss landscape might also explain why there was less variability in model performance across different folds of the data with the Res-UNet model relative to UNet-1 or UNet-Drop, as can be seen above in Figure 3.5.

The purpose of including dropout layers in the UNet-Drop architecture was to evaluate if adding dropout layers would make the CNN more generalizable to sites it was not trained on. If this were the case, we would expect the performance of UNet-Drop to be higher than UNet-1 on test sets containing image-label pairs from sites not seen in train sets, such as in the Ten\_site\_test. We did not see the expected increase in ensemble F1-score in UNet-Drop compared to UNet-1 as shown above in rows 21-22 of Table 3.1. More work should be done to test different configurations of dropout layers, and different percentages of dropout.

Our image augmentation method created separate augmented data sets which increased the size and training time significantly. For example, the One-site 100 data set, with augmentations applied grew to 5,400 images, and training time increased from twenty-five minutes per fold on the Tesla P100 GPU available through Google Colab to nearly six hours per fold on a Quadro P2000 GPU on a desktop PC. Due to runtime restrictions, long training

sessions will often cause Google Colab to timeout, which necessitated the use of a desktop PC with a GPU for augmented data sets. While the training-times increased significantly with augmented data, this one-time cost resulted in a model with improved model performance. The results in Figure 3.6 also show that we were able to achieve an accuracy of  $F1 \geq 0.9$  with 100 training images from a single site, but also with fewer images using image augmentation. Researchers must choose whether to spend time creating larger labeled data sets or implementing image augmentation resulting in longer training times.

We found that a site-specific Res-UNet model could be trained and tested with as few as 50 images (25 for training and validation, and 25 for testing) for binary semantic segmentation and achieve an accuracy of  $F1 = 0.93$  (see row 2 of Table 3.1), if image augmentation is applied. Although the generalizability of our results is somewhat limited because we only tested the single-site workflow on one site (30-mile), we expect similar performance at other sites that have similar features. Future work will include single-site testing for several other sites.

Importantly, we also found that our multi-site train set with augmentation did not provide the intersite variability needed to train models to segment multiple sites to our predetermined accuracy threshold, or to generalize to sites outside of the train set. Ensemble accuracy for the Res-UNet trained with Five\_site\_aug and tested on Five\_site\_test achieved an ensemble accuracy of  $F1 = 0.89$  (row 17 of Table 3.1). This leads us to believe one, or all of the following are true: 1) 30-mile single-site network was overfit to the image-label pairs from that site; 2) adding more sites in the same proportion increases model generalizability at the expense of accuracy by introducing more variability into the CNN's 'sandbar' class, or 3) imagery from particular sites might be more difficult to train a CNN to segment than for 30-mile, and that for certain sites more than thirty images are needed.

In this study, we did not assess the optimal size of a test set. We somewhat arbitrarily chose 25 images for a single-site, 30 for three sites (ten from each site), 50 for five sites (10 each), and 100 for ten sites (10 each). Some of the performance loss observed as trained models were tested on the `Five_site_test` and `Ten_site_test` sets likely results from the increased number of images in the unequal sized test sets. We would expect that with each additional image in a test set, the set would have increased variability until, at some point, most of the inter- and intrasite variability in the imagery dataset is captured by the test set, or the test set is of a size that it is representative of the imagery as a whole. We carefully included images capturing the range of lighting, the color of the water, and size conditions at each bar while minimizing computational time and time spent labeling imagery. Future studies should gradually increase the size of the test sets, observing decreases in accuracy with the increased test-set size until accuracy plateaus suggesting much of the intersite or intrasite variability had been captured within the test set. Image augmentation could also be conducted on the test sets to artificially increase the test set size and variability.

The goal of this research was not to create a CNN, which could identify subaerial sand or sandbars in all imagery outside of Grand Canyon. Instead, we trained specialized neural networks end-to-end for our specific dataset and research needs. While certain aspects of our data set are unique to Grand Canyon, large coastal and fluvial imagery datasets are becoming increasingly common, and many have similar attributes. We demonstrate that with relatively little computing power, and with a minimal amount of time spent labeling imagery, an accurate CNN could be trained for binary semantic segmentation with time-series image data containing natural textures.

Future studies might consider comparing the accuracy of encoder-decoder CNNs like those used here with other type of CNNs, such as those used by Buscombe and Richie (2019) and Carbonneau and others (2020). Subsequent segmentation attempts could also experiment

with segmentation before rectification and registration which may improve segmentation by increasing resolution during segmentation and limiting the effects of pixel warping during rectification at the expense of increasing training times and the need for more advanced GPUs. The sequence of steps we utilized in this study was a product of methods evolving through time, storage limitations, and computational limitations and thus more experimentation is warranted to determine the ideal sequence.

As water scientists trying to understand how to apply DL to water resource problems, we found that much of the DL literature is focused on very large image test sets with tens of thousands of unique images and scenes. Little guidance exists on how much labeled imagery is needed to train, validate, and test a CNN tailored to data that is focused on change detection at one or a few sites with significant variability in image angle and lighting conditions. Our research can serve as an example and a starting point for other water scientists attempting to apply DL to their datasets.

### **3.6. ACKNOWLEDGEMENTS, SAMPLES, AND DATA**

Thanks to the U.S. Geological Survey Grand Canyon Monitoring and Research Center (GCMRC) for making the data available and funding this research. This study was supported by the Glen Canyon Dam Adaptive Management Program administered by the U.S. Department of the Interior Bureau of Reclamation.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

This work would not have been possible without the funding and support of the Remote Sensing and Geoinformatics Lab at Northern Arizona University. Thanks to Bob Tusso for camera maintenance and data set management. Thanks to Matt Kaplinski, Joe

Hazel, and everyone else on the annual sandbar monitoring trips for assisting with the collection of GCPs.

Data from this research and the associated code can be found at [https://doi.org/10.5281/zenodo.4081945]. To the best of our knowledge there are no conflicts of interest.

### 3.7. REFERENCES

- Arif, M., Farooq, W., Saduf, A. B., Asif, A., & Khan, I. (2020). *Advances in Deep Learning: Studies in Big Data Volume 57*. Springer. <http://www.springer.com/series/11970>
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481–2495.
- Baumhoer, C. A., Dietz, A. J., Kneisel, C., & Kuenzer, C. (2019). Automated extraction of Antarctic glacier and ice shelf fronts from Sentinel-1 imagery using deep learning. *Remote Sensing*, 11(21), 1–22.
- Bertoldi, W., Piégay, H., Buffin-Bélanger, T., Graham, D., & Rice, S. (2012). Applications of Close-range Imagery in River Research.
- Bogle, R., Velasco, M., & Vogel, J. (2013). An Automated Digital Imaging System for Environmental Monitoring Applications.
- Bouguet, J.-Y. (2015). Camera Calibration Toolbox for Matlab. Bradski, G. (2000). The OpenCV Library.
- Breuel, T. M. (2015). The Effects of Hyperparameters on SGD Training of Neural Networks.
- Burguera, A. (2020). Segmentation through patch classification: A neural network approach to detect *Posidonia oceanica* in underwater images. *Ecological Informatics*, 56(January), 101053.
- Burkov, A. (2019). *The Hundred-Page Machine Learning Book* (1st ed.). Self-Published.

- Buscombe, D. & Carini, R. J. (2019). A data-driven approach to classifying wave breaking in infrared imagery. *Remote Sensing*, 11(7), 1–10.
- Buscombe, D., Ritchie, A., Buscombe, D., & Ritchie, A. C. (2018). Landscape Classification with Deep Neural Networks. *Geosciences*, 8(7), 244.
- Carbonneau, P. E., Belletti, B., Micotti, M., Lastoria, B., Casaioli, M., Mariani, S., Marchetti, G., & Bizzi, S. (2020). UAV-based training for fully fuzzy classification of Sentinel-2 fluvial scenes. *Earth Surface Processes and Landforms*.  
<https://doi.org/10.1002/esp.4955>
- Carbonneau, P.E., Dugdale, S.J., Breckon, T.P., Dietrich, J.T., Fonstad, M.A., Miyamoto, H., Woodget, A.S., 2020. Adopting deep learning methods for airborne RGB fluvial scene classification. *Remote Sens. Environ.* 251, 112107.  
<https://doi.org/10.1016/j.rse.2020.112107>
- Carbonneau, P. E., & Piégay, H. (2012). Fluvial Remote Sensing for Science and Management. *Fluvial Remote Sensing for Science and Management*, 1–440.  
<https://doi.org/10.1002/9781119940791>
- Cheng, G., Han, J., & Lu, X. (2017). Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proceedings of the IEEE*, 105(10), 1865–1883.
- Chu, Z., Tian, T., Feng, R., & Wang, L. (2019). SEA-LAND SEGMENTATION WITHRES-UNET AND FULLY CONNECTED CRF Zhengquan Chu, Tian Tian \*, Ruyi Feng , Lizhe Wang Hubei Key Laboratory of Intelligent GeoInformation Processing, School of Computer Science , China University of Geosciences , Wuhan 430074 ,. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 3840–3843.
- Congalton, R. G. & Green, K. (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, volume 2.
- Converse, Y. K., Hawkins, C. P., & Valdez, R. A. (1998). Habitat relationships of subadult humpback chub in the Colorado River through Grand Canyon: Spatial variability and implications of flow regulation. *Regulated Rivers-Research & Management*, 14(3), 267–284.

- Dexter, L. R., Cluer, B. L., & Manone, M. M. (1995). Using Land-based Photogrammetry to Monitor Sandbar Stability in Grand Canyon on a Daily Time Scale. In Riper, C. V. (Ed.), *Proceedings of the Second Biennial Conference of Research in Colorado Plateau National Parks, Transactions and Proceedings Series*. National Park Service.
- Diakogiannis, F. I., Waldner, F., Caccetta, P., & Wu, C. (2020). ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162(January), 94–114.
- Dolan, R., Howard, A., & Gallenson, A. (1974). Man's impact on the Colorado River in the Grand Canyon. *American Scientist*, 62(4), 392–401.
- Draut, A. E. & Rubin, D. M. (2007). The role of aeolian sediment in the preservation of archaeological sites in the Colorado River corridor, Grand Canyon, Arizona: final report on research activities, 2003-2006.
- Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., & Pal, C. (2016). The importance of skip connections in biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10008 LNCS, 179–187.
- Evangelidis, G. D. & Psarakis, E. Z. (2008). Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10).
- Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A Review on Deep Learning Techniques Applied to Semantic Segmentation, 1–23.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning: Machine Learning Book*. 785. <http://www.deeplearningbook.org/>
- Grams, P. E., Schmidt, J. C., Wright, S. A., Topping, D. J., Melis, T. S., & Rubin, D. M. (2015). Building sandbars in the Grand Canyon. *EOS*, 96.
- Grams, P. E., Tusso, R. B., & Buscombe, D. (2018a). Automated remote cameras for monitoring alluvial sandbars on the Colorado River in Grand Canyon, Arizona.
- Grams, P. E., Tusso, R. B., & Buscombe, D. (2018b). Automated Remote Cameras for Monitoring Alluvial Sandbars on the Colorado River in Grand Canyon, Arizona.

- Guo, H., He, G., Jiang, W., Yin, R., Yan, L., Leng, W., 2020. A multi-scale water extraction convolutional neural network (MWEN) method for GaoFen-1 remote sensing images. *ISPRS Int. J. Geo-Information* 9. <https://doi.org/10.3390/ijgi9040189>
- Guo, Y., Cao, X., Liu, B., & Gao, M. (2020). Cloud detection for satellite imagery using attention-based U-net convolutional neural network. *Symmetry*, 12(6).
- Hartley, R. & Zisserman, A. (2003). *Multiple View Geometry in Computer Vision* (2nd ed.). Cambridge, UK: Cambridge University Press.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, 770–778.
- Heidler, K., Mou, L., Baumhoer, C., Dietz, A., Zhu, X.X., 2021. HED-UNet: Combined Segmentation and Edge Detection for Monitoring the Antarctic Coastline. *IEEE Trans. Geosci. Remote Sens.* 1–14. <https://doi.org/10.1109/TGRS.2021.3064606>
- Jiao, L., Huo, L., Hu, C., & Tang, P. (2020). Refined UNet: UNet-based refinement network for cloud and shadow precise segmentation. *Remote Sensing*, 12(12), 1-28.
- Kearsley, L. H., Schmidt, J. C., & Warren, K. D. (1994). Effects of Glen Canyon dam on Colorado River sand deposits used as campsites in Grand Canyon National Park. *Regulated Rivers*, 9(3), 137–149.
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Lguensat, R., Sun, M., Fablet, R., Mason, E., Tandeo, P., & Chen, G. (2018). EddyNet: A deep neural network for pixel-wise classification of oceanic eddies. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018-July, 1764–1767.
- Li, H., Xu, Z., Taylor, G., Studer, C., & Goldstein, T. (2018). Visualizing the loss landscape of neural nets. *Advances in Neural Information Processing Systems*, 2018-Decem. (Nips 2018), 6389–6399.
- Li, R., Liu, W., Yang, L., Sun, S., Hu, W., Zhang, F., & Li, W. (2018). DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(11), 3954–3962.

- Ling, F., Boyd, D., Ge, Y., Foody, G. M., Li, X., Wang, L., Zhang, Y., Shi, L., Shang, C., Li, X., & Du, Y. (2019). Measuring River Wetted Width from Remotely Sensed Imagery at the Subpixel Scale with a Deep Convolutional Neural Network. *Water Resources Research*, 55(7), 5631–5649.
- Long, J., Shelhamer, E., & Darrel, T. (2015). Fully Convolutional Networks for Semantic Segmentation. In *In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 3431–3440)., Boston, MA.
- Marcus, W. A. & Fonstad, M. A. (2010). Remote sensing of rivers: The emergence of a subdiscipline in the river sciences. *Earth Surface Processes and Landforms*, 35(15), 1867–1872.
- Nagi, A.S., Kumar, D., Sola, D., Scott, K.A., 2021. RUF: Effective Sea Ice Floe Segmentation Using End-to-End RES-UNET-CRF with Dual Loss 1–19.
- NPS. (2006). Grand Canyon: Colorado River Management Plan.
- Nath, A., Mawlong, P., & Saha, G. (2019). River body extraction using convolutional neural network. *Journal of Information and Optimization Sciences*, 40(8), 1741–1751.
- Perez, L. & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning.
- Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carbonneau, P., Dufour, S., Liébault, F., Ruiz-Villanueva, V., & Slater, L. (2020). Remotely sensed rivers in the Anthropocene: state of the art and prospects. *Earth Surface Processes and Landforms*, 45(1), 157–188. <https://doi.org/10.1002/esp.4787>
- Reddy, B. S. & Chatterji, B. N. (1996). An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 5(8), 1266–1271.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351, 234–241.
- Ruder, S. (2016). An overview of gradient descent optimization algorithms, 1–14. Sankey, J. B., Caster, J., Kasprak, A., & East, A. E. (2018). The response of source-bordering

- aeolian dunefields to sediment-supply changes 2: Controlled floods of the Colorado River in Grand Canyon, Arizona, USA. *Aeolian Research*, 32(October 2017), 154–169.
- Shen, C. (2018). A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resources Research*, 54(11), 8558–8593.
- Smith, L. N. (2017). Cyclical learning rates for training neural networks. *Proceedings- 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017, (April)*, 464–472.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple way to prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15, 1929–1958.
- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10553 LNCS, 240–248.
- Valdez, R. A., Hoffnagle, T. L., Mcivor, C. C., Mckinney, T., Leibfried, C., Applications, S. E., & Jun, N. (2012). Effects of a Test Flood on Fishes of the Colorado River in Grand Canyon, Arizona EFFECTS OF A TEST FLOOD ON FISHES OF THE COLORADO RIVER IN GRAND CANYON, ARIZONA. 11(3), 686–700.
- Valdez, R. A. & Ryel, R. J. (1995). Life History and Ecology of the Humpback Chub in the Colorado River in Grand Canyon, Arizona.
- Xu, Y., Jia, R., Mou, L., Li, G., Chen, Y., Lu, Y., & Jin, Z. (2016). Improved relation classification by deep recurrent neural networks with data augmentation. COLING 2016 - 26th International Conference on Computational Linguistics, Proceedings of COLING 2016: Technical Papers, 1461–1470.
- Yan, R., Dong, S., 2019. Optical Remote Sensing Image Waters Extraction Technology Based on Deep Learning Context-Unet. ICSIDP 2019 - IEEE Int. Conf. Signal, Inf. Data Process. 2019 2019–2022. <https://doi.org/10.1109/ICSIDP47821.2019.9173433>

# **CHAPTER 4: MEASURING EDDY SANDBAR SUB-ANNUAL EROSION AND DEPOSITION PATTERNS USING OBLIQUE IMAGE TIME-SERIES OF THE COLORADO RIVER, GRAND CANYON**

**Ryan. E. Lima<sup>a</sup>, Temuulen T. Sankey<sup>b</sup>, Paul. E. Grams<sup>c</sup>, Erich. R. Mueller<sup>d</sup>, Taylor J. Joyal<sup>a</sup>**

<sup>a</sup> School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ

<sup>b</sup> School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ

<sup>c</sup> U.S. Geological Survey, Southwest Biological Science Center, Grand Canyon Monitoring and Research Center, Flagstaff, AZ

<sup>d</sup> Department of Geosciences, Southern Utah University, Cedar City, UT

**KEYWORDS:** Fluvial Remote Sensing, Mass Failure, Oblique Imagery, Alluvial Sandbars, Adaptive Management

## **4.1. ABSTRACT**

Sandbars are salient features in the Colorado River that allow for recreation, provide habitat for riparian vegetation, and create backwater habitat for native fish species. Daily oblique images captured by remote cameras from 41 sandbar monitoring sites on the Colorado River, Grand Canyon were used to examine mass failure occurrence. Then a detailed analysis of spatial variability in the sub-annual response to discharge patterns on daily, monthly, and seasonal timescales was conducted at three sites. In addition, a total of 226 mass failures were identified over ten years at 27 of 41 sandbars examined. Mass failures resulted in a mean area loss of between 0.5 - 30% of the exposed sandbar area at two sites.

We found evidence that mass failure occurrence is related to monthly and seasonal adjustments in discharge, increased discharge variability, and increased daily or monthly mean discharges. The accurate quantification of daily erosion and deposition rates at our three target sites was hampered by errors in area estimates that equalled or exceeded the magnitude of expected rates of change. However, mean monthly area showed significant differences in the magnitude and frequency of cyclic erosion and deposition at the three reattachment bars. Among them, the 30-mile site was the most responsive to controlled floods, but all three sites showed increased area in the six months following controlled floods. These gains in area were generally lost with the onset of high summer flows in June, July, and August. Antecedent size was a more important driver of changes in monthly sandbar area than the monthly flow metrics examined. Our results demonstrate the utility of re-examining an existing fluvial imagery dataset to provide policy-relevant findings and elucidate relationships between sandbar response and discharge patterns on the dam-affected Colorado River through Grand Canyon.

## **4.2. INTRODUCTION**

Fluvial remote sensing studies have increased rapidly as new sensing technologies have emerged (Marcus and Fonstad, 2010; Carbonneau and Piégay, 2012). Advances in deep learning have also improved image segmentation and interpretation within the water sciences (Shen, 2018). While new technologies offer tremendous potential for designing future studies, further exploring existing datasets was identified as an important avenue for the future of fluvial remote sensing (Piégay et al., 2020). For example, large camera networks have been employed to monitor vegetation change and bank erosion across river basins, as well as beach erosion on large sections of coastline (Nieto et al., 2010; Briere et al., 2011; Brinley-Buckley et al., 2016; Splinter et al., 2018). In addition, ground-based camera systems

capturing repeat images or videos have been used to observe coastal and fluvial processes at high spatial and temporal resolutions with minimal time in the field (Lane et al., 2000; Bertoldi et al., 2012).

Today over half of the world's major rivers are affected by dams, which collectively sequester between 25 and 60 Gt of sediment per year (Nilsson et al., 2005; Walling, 2012). Dams reduce sediment supply and alter the flow of their rivers by changing the magnitude and timing of discharge resulting in a broad range of morphological responses (Brandt, 2000; Schmidt and Wilcock, 2008; Williams and Wolman, 1984). The Colorado River in the Grand Canyon is among the most rigorously studied dam-impacted river reaches in the world. A robust long-term, annual monitoring program was established with the passage of the 1992 Grand Canyon Protection Act to ensure that Glen Canyon Dam (GCD) is operated:

*“...in such a manner as to protect, mitigate adverse impacts to, and improve the values for which Grand Canyon National Park and Glen Canyon National Recreation Area were established, including but not limited to natural and cultural resources and visitor use”* (Grand Canyon Protection Act 1992, Public Law 102-575).

In part, a network of remote cameras has been capturing oblique images of sandbar monitoring sites in the Grand Canyon since 1990. Until now, those remote cameras have been primarily used to complement detailed annual topographic surveys with qualitative assessments of imagery examining changes in sandbar size in response to controlled floods downstream of GCD. This research is the most comprehensive attempt at re-visiting the Grand Canyon remote camera dataset to examine changes in fine sediment storage at daily and monthly scales and to measure the effects of controlled floods. Among the resources protected by that law are the eddy sandbars within the Grand Canyon. Sandbars provide substrate for riparian vegetation, backwater habitat preferred by native fish species, and they are the favored camping and picnicking location for up to 25,000 annual river runners

(Kearsley et al., 1994; Converse et al., 1998; Nat. Park Serv., 2006; Vincent and Andrews, 2008; Valdez et al., 2012).

Sandbars have been a subject of ongoing scientific interest and monitoring due their importance to river-based recreation in the canyon, as well as other ecosystem services they provide. The focus of current sandbar monitoring efforts is understanding annual and decadal trends in fine ( $<2mm$ ) sediment storage or measuring sandbar response before and after a handful of controlled floods. Such efforts measure the cumulative effects of discharge patterns which vary on monthly, daily, and hourly scales, providing a conservative estimate of sandbar size through time (Hazel et al., *in review*). However, gaps remain in our understanding of cyclic patterns of sandbar deposition and erosion in response to daily discharge patterns and monthly shifts in discharge.

A more detailed understanding of the link between sub-annual discharge patterns and variability in morphodynamic response can help inform adaptive management efforts and improve future dam operations. The objectives of this research are: 1) to investigate the utility and accuracy of repeat oblique images in quantifying daily and monthly changes in sandbar size; 2) to examine variability in the sub-annual responses of three dynamic sandbars located in similar geomorphic settings to daily and monthly flow adjustments and controlled floods, and 3) to use time-lapse videos from 41 sandbar monitoring sites to examine the frequency and magnitude of mass failure erosional events as well as the flow patterns related to their occurrence. The methods described here could be used to implement detailed monthly analyses at more sites within the Grand Canyon imagery dataset or to re-analyze data from other fluvial or coastal monitoring systems.

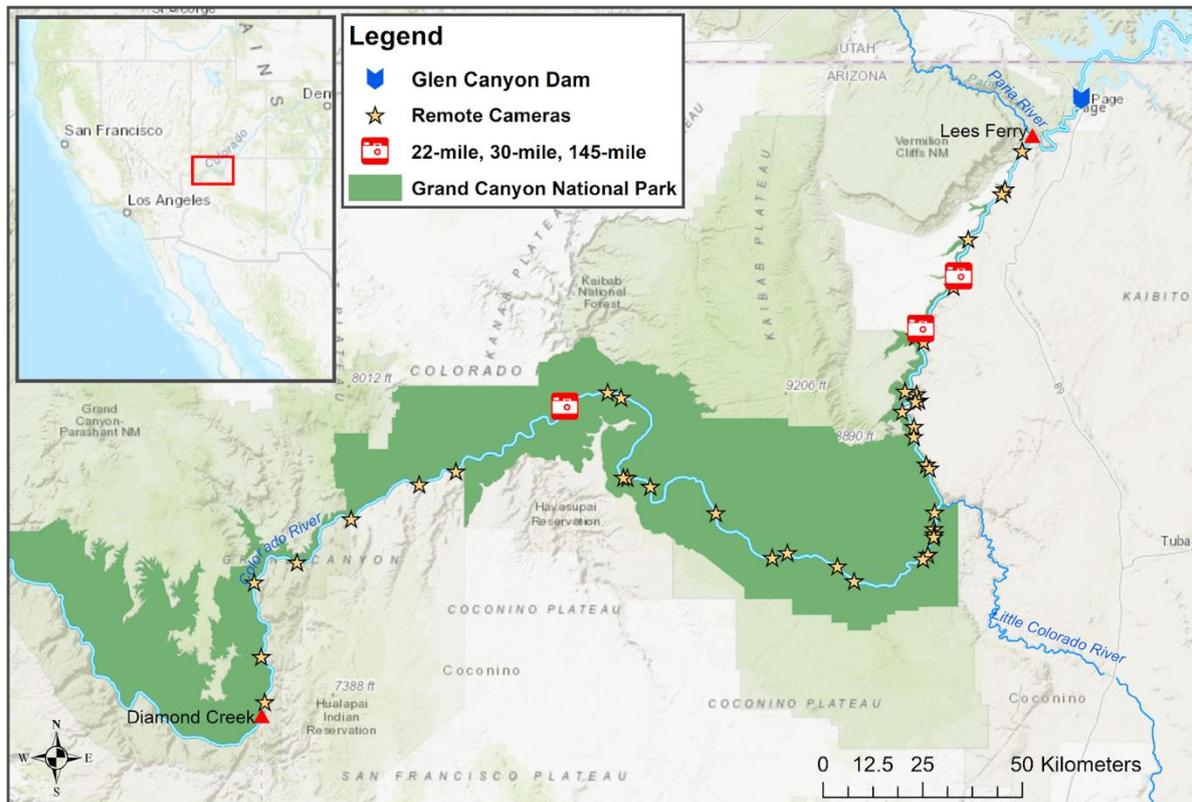
#### 4.2.1. Regional setting and site descriptions

This study is focused on the Colorado River within Grand Canyon National Park (Figure 4.1). The gorge of the Colorado River reaches depths of 2,048 *m* and extends 446 *km* from Lees Ferry in the east to the Grand Wash Cliffs in the west (Annerino, 2000). By convention, sandbar sites within the canyon are specified by their river mile (RM) downstream from Lees Ferry, AZ (RM 0). Our study utilizes long-term sandbar monitoring sites located between Lees Ferry (RM 0) and Diamond Creek (RM 225) including a total distance of 362 *km*.

Remote camera systems include a digital camera, zoom lens, memory card, solar panel, battery, and data logger housed in a weatherproof container (Bogle et al., 2013). The cameras are generally situated across the river from the sandbars on talus slopes providing oblique imagery of monitoring sites. Each camera is programmed to capture five images per day at 2-hour intervals from 8 a.m. to 4 p.m. MST (Grams et al., 2018).

A detailed analysis of segmented imagery was carried out at three sites located at RM 22, 30.7, and 145.9, referred to hereafter as *22-mile*, *30-mile*, and *145-mile* (Figure 4.1). These sites were chosen because they share a similar geomorphic setting and are classified as narrow reattachment bars (Hazel et al., *in press*). Sandbars in this grouping are mostly unvegetated, have the most dynamic response to controlled floods among the monitored sandbar sites, and experience the highest stage changes in the canyon during current dam operations (5-6 *m*) (Mueller et al., 2018). This study aims to understand how discharge patterns caused by the operation of GCD affect sandbar area. We chose sites that have similar annual trends and geomorphic settings to examine the site-specific responses while minimizing inter-site variability (Figure 4.1) more clearly. The dynamics inherent in the response of these narrow reattachment bars also make them ideal for examining links between discharge patterns and changes in sandbar area. We also conducted an analysis of

mass failures using imagery from 2011 - 2021 at 41 long-term sandbar monitoring to assess the frequency of mass failures more broadly throughout the canyon.



**Figure 4.1** Study Area and the Grand Canyon National Park (highlighted in green), which surrounds the gorge of the Colorado River. Glen Canyon Dam is located 24km upstream of Lees Ferry and the Colorado River primarily flows from the upper right to the bottom left of the figure. Yellow stars indicate the 41 sandbar monitoring sites examined in our time-lapse analysis. Red camera icons indicate where more detailed analysis of imagery was conducted at 22-mile, 30-mile, and 145-mile sandbars.

#### 4.2.1.1. *Grand Canyon sandbars*

Rivers with bedrock channels generally have steeper gradients and more significant stage changes relative to discharge than alluvial rivers (Baker, 1984). In such settings, high flows can deposit fine sediment at relatively high elevations, where it becomes less vulnerable to erosion and subsequent transport during lower flows (Baker, 1984; Tinkler and Wohl, 1998). Sandbars thus represent an important storage location for fine sediment.

Sandbars are discrete fluvial bedforms composed primarily of fine sediment. In the Grand Canyon, sandbars occur as persistent geomorphic features along banks within the

debris fan-affected bedrock canyon (Hazel et al., *in press*). Debris flows initiated in high-relief tributaries cause channel constrictions, where high-velocity flow separates from the bank creating areas of low-velocity, recirculating current between the bank and the high-velocity, downstream flow (Webb et al., 1989, Andrews and Vincent, 2007; Mueller et al., 2014). Such zones may contain one or more eddies, rotational flow structures with more or less vertical axes, as well as areas of low-velocity flow that are not organized into rotation (Matthes, 1947; Schmidt and Graf, 1990). Such areas may be stagnant, have a preferential flow direction, or oscillate in several directions (Schmidt and Graf, 1990). Suspended sediment advected into these recirculation zones may be deposited, forming one or more eddy sandbars (Schmidt, 1990; Schmidt and Graf, 1990).

Deposition within recirculation zones occurs where sediment transport capacity is lowest, near the points where flow separates from the bank (the separation point), and where flow reattaches to the bank (the reattachment point) (Schmidt, 1990; Wiele et al., 1999). Deposition occurs at or below the water's surface and may be limited by the amount of space available within the recirculation zone, known as accommodation space (Hazel et al., 1999; Schmidt and Grams, 2011). Therefore, deposition is influenced by sediment availability, accommodation space, and flow structures in the low-velocity recirculation zone that vary from site to site depending on stage, channel geometry, and antecedent bar morphology (Schmidt and Graf, 1990; Mueller et al., 2018).

Before the construction of GCD, the extent and morphology of sandbars in the Grand Canyon were the product of abundant sand supply and large seasonal fluctuations in flow (Beus and Avery, 1992). The closure of GCD in 1963 eliminated fine sediment flux into the canyon from the upper basin of the Colorado River. Today the largest sources of fine sediment to the Grand Canyon are the Paria River, which joins the Colorado River at RM 1, and the Little Colorado River, which enters at RM 62. Together they contribute between 15 -

20% of the pre-dam fine sediment load creating sediment-limited conditions in the Grand Canyon (Topping et al., 2000; Wright et al., 2005).

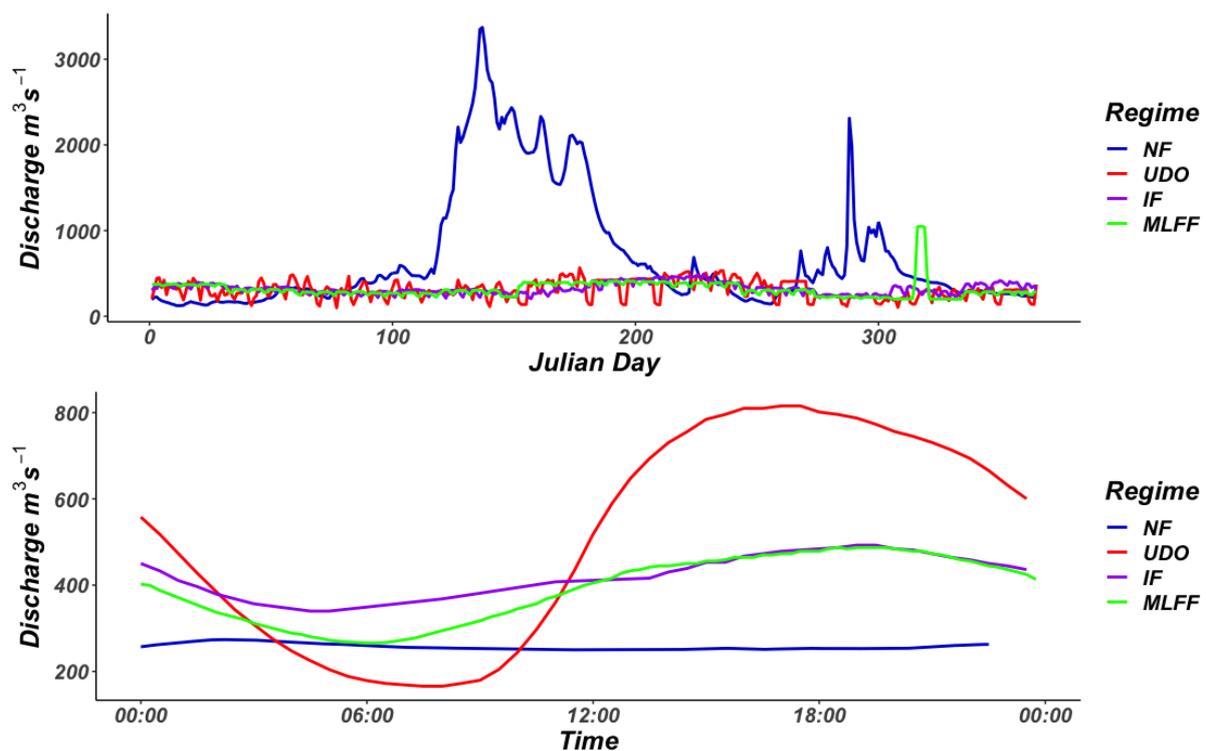
#### **4.2.2. Flow regimes and controlled floods**

The 1963 closure of GCD ended the naturalized flow period (NF) and eliminated snowmelt-driven spring floods, replacing large annual flow fluctuations with large daily variability, which significantly impacted sediment resources downstream. Documentation of the decreasing size and extent of Grand Canyon sandbars due to dam operations began in 1974 (Dolan et al., 1974). The potential loss of ecosystem services, resulting from the decline of sandbars and other dam-induced changes to the downstream environment, led to an adaptive management program and extensive monitoring (U.S. Dept. of Int., 1995).

Annual releases from GCD are targeted at 8.23 million acre-feet (*maf*) or 10.15 billion  $m^3$ , satisfying downstream obligations set by the Law of the River and maximizing storage in Lake Powell (National Resource Council, 1996). Typical monthly releases vary between 0.5 and 1 *maf*, with lower releases occurring in the spring and fall months. Higher releases occur in the summer and winter months to balance demands for hydropower, delivery obligations, recreation, and other downstream resources (National Resource Council, 1996). Daily discharge patterns from GCD are influenced by energy demand and operating rules, which define the maximum and minimum releases, and the rates of discharge change (National Resource Council, 1996). Therefore, we focus on understanding the effects of the dam operations on the sandbar daily, monthly, and seasonal behavior to provide information to the Glen Canyon Dam adaptive management program and insight into the morphodynamics of fluvial bedforms in dam-affected rivers.

Since the end of NF, we identify three other distinct flow periods based on operating rules and resulting dam operations (Fig. 4.2). In the first period from 1963 to 1991, known as unconstrained dam operations (UDO), the hydrograph was primarily shaped by energy

demands and water delivery obligations. High daily flow variability and large down-ramping rates were cited as drivers of sandbar erosion and led to a brief period of reduced flow variability known as interim flows (IF) from 1991 to 1996 (Budhu and Gobin, 1994; Budhu and Gobin, 1995; National Resource Council, 1996; Dexter and Cluer, 1999). During IF, maximum daily discharge was reduced, the down ramp rate was decreased, and minimum discharges were increased, all of which reduced flow variability (Fig. 4.2) (National Resource Council, 1996). Monitoring observations during IF led to the adoption of the current modified low fluctuating flow regime (MLFF) from 1996 to the present. In the MLFF, the maximum allowable daily discharge increased, and sandbar-building controlled floods were introduced (Fig. 4.2) (National Resource Council, 1999; Grams et al., 2015).

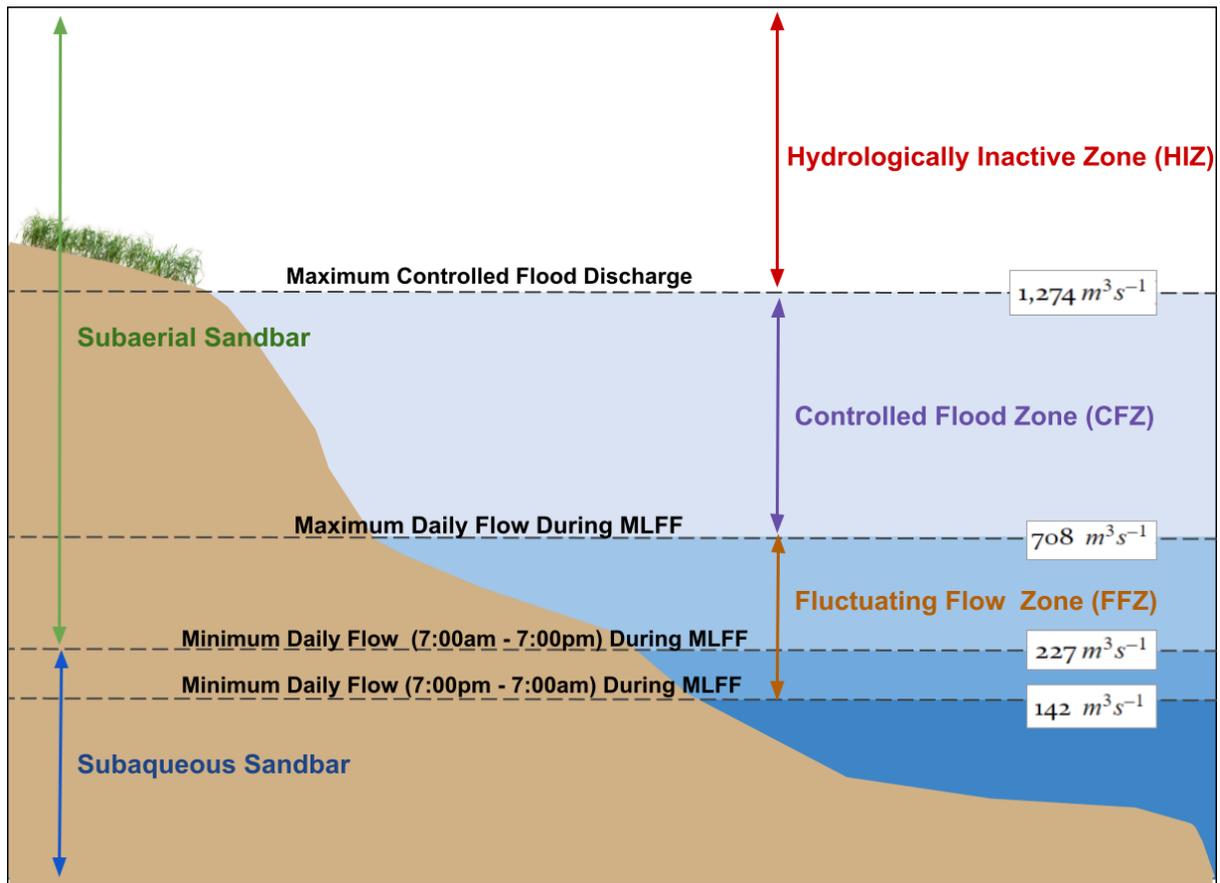


**Figure 4.2** Annual (top) and daily (bottom) hydrographs displaying representative years from each GCD flow period: 1941 (NF - Naturalized Flow), 1990 (UDO - Unconstrained Dam Operations), 1993 (IF - Interim Flows), and 2013 (MLFF - Modified Low Fluctuating Flows). The daily hydrograph represents flows from August 8th of each year. The NF period (blue) has the highest annual and lowest daily flow variability. Daily flow variability is most extreme during the UDO (red) and is the least extreme during IF. The large spike in discharge around Julian days 310-315 in the MLFF (Green) represents a controlled flood, which is considerably smaller than the flooding during NF. Daily variability during the MLFF is higher than during the IF but lower than the UDO.

Controlled floods involve water releases above power plant capacity ( $940m^3s^{-1}$ ) up to a discharge of  $1,274m^3s^{-1}$  for up to 96 hours. While Controlled floods have occurred in both the fall and spring, they typically occur in November in years when a sufficient volume of sand has accumulated in the reach downstream of Lees Ferry. If these conditions exist, fine sediment deposited during tributary floods can be mobilized and deposited at high elevations within eddies, aggrading sandbars (Hazel et al., 2010, U.S. Dept. of Int., 2011; Grams et al., 2015).

#### **4.2.3. Sandbar behavior**

The Colorado River drops 670m as it travels through the canyon. Since we are measuring area change occurring above the water's surface in imagery throughout the canyon, it is convenient to use discharge as a proxy for elevation. We delineate elevation zones or vertical locations on different sandbars which are emergent or submerged at different volumetric flow rates (Fig. 4.3). Daily discharges during MLFF rarely fall below  $227m^3s^{-1}$ ; therefore, annual topographic surveys focus on the portion of the sandbar that is emergent at that discharge-- hereafter described as the *subaerial* sandbar. The portion of the bar inundated at a discharge of  $227m^3s^{-1}$  is referred to as the *subaqueous* sandbar. The fluctuating flow zone (FFZ) is the portion of the bar exposed to fluctuating flows during normal dam operations ( $142m^3s^{-1} - 708m^3s^{-1}$ ) (Fig. 4.3). The controlled flood zone (CFZ) is the portion of the bar only inundated during controlled floods ( $708m^3s^{-1} - 1,274m^3s^{-1}$ ). The hydrologically inactive zone (HIZ) ( $>1,274m^3s^{-1}$ ), which is unaffected by current dam operations containing sediment deposited before dam closure, during floods in the 1980s, or sediment deposited by the wind (Fig. 4.3).



**Figure 4.3** Diagram of flow zones on a conceptualized Grand Canyon sandbar. The difference between subaerial ( $227 \text{ m}^3 \text{ s}^{-1}$ ) and subaqueous ( $<227 \text{ m}^3 \text{ s}^{-1}$ ) portions of the sandbar are highlighted along with elevational zones related to dam operations during the current Modified Low Fluctuating Flow (MLFF) period.

#### 4.2.3.1. Sandbar Erosion

Several erosional processes have been identified in the Grand Canyon, which may act alone or combined with other fluvial processes to erode sandbars. Depending on the process, or processes, such erosion can proceed gradually, resulting in meters of lateral bank retreat over weeks to months or rapidly resulting in meters of erosion over seconds to hours. The action of waves within areas of low-velocity flow can entrain sand in otherwise depositional settings, undercutting or steepening sandbar faces in the FFZ and making them susceptible to other erosional processes (Bauer and Schmidt, 1993). Similar over-steepening can occur from tractive force erosion by nearshore currents (Schmidt and Graf, 1990). Seepage erosion occurs as water exfiltrates from the bar following decreases in stage. Groundwater stage within

sandbars lags relative to the hydrograph, so during rapid reductions in discharge, water escaping the bar may cause increased pore pressure leading to instability, slumping, rilling, or slope failure within the FFZ (Beus and Avery, 1992; Buhu and Gobin, 1994; Budhu and Gobin, 1995). Such processes generally act on sandbars gradually but can lead to more rapid erosion under certain conditions.

Rapid erosional events, identified in imagery and referred to hereafter as mass failures, seem to result from processes originating below the water surface on subaqueous slopes, or on slopes within the low end of the FFZ. Most occur upstream of the reattachment point and are characterized by semi-circular scars consistent with the direction of upstream current in the recirculation zone (Fig. 4.4) (Hazel et al., *in press*). Studies observing mass failures in the Grand Canyon have suggested breaching or static liquefaction as a primary cause (Budhu, 1992; Budhu and Gobin, 1995; Cluer, 1995; Wright and Kaplinski, 2011, Hazel et al., *in press*).



**Figure 4.4** An example of a mass failure at the 41-mile sandbar occurred between Jun 26th and 27th of 2012. Observed mass failures in the Grand Canyon often have this rounded morphology, which is consistent with descriptions in the breaching literature.

Outside of the Grand Canyon, mass failures have been described by several different terms. Such events in fine to medium sand ( $250\mu\text{m}$  -  $0.25\text{mm}$ ) have been attributed to either static liquefaction, in loosely packed sands, or to a phenomenon known as breaching, which

is restricted to more densely packed sand (Yamamuro and Lade, 1997; Van den Berg, 2002; Mastbergen et al., 2019). Liquefaction flow slides are instantaneous failures of large masses of sediment and typically occur on the order of minutes. Conversely, failures by retrogressive breaching generally occur over tens of minutes to hours, during which steep to vertical slopes retreat gradually into the bank (Van den Berg, 2002). First observed during dredging activities, breaching has since been interpreted as the main failure mechanism in several fluvial, tidal, and coastal settings (de Koning, 1970; Mastbergen et al., 2019). While the triggers of breaching events are still poorly understood, they are likely related to some combination of over-steepening of sand deposits and changes in pore pressure associated with fluctuating flows or precipitation events (Mastbergen et al., 2019). Many flow slides or mass failures at eddy sandbars are likely hybrid involving both breaching and liquefaction mechanisms that occur independently of each other (Van den Berg personal communications April 2021).

Laboratory experiments that examined the effects of down ramping flow rates on slope stability found that failures were the dominant form of erosion on steep bar faces ( $\geq 26^\circ$ ) and occurred regardless of ramp rate. Seepage erosion dominated on moderate slopes ( $18^\circ - 22^\circ$ ), and both failure and seepage erosion largely ceased as bars reached an equilibrium slope ( $\leq 14^\circ$ ) (Alvarez and Schmeeckle, 2013). Their results suggest that slope is a significant factor in influencing erosion rates and mechanisms, however, the flow patterns that rework sandbars to achieve steep angles, or trigger mechanisms leading to rapid versus gradual erosion is the subject of this paper. We seek to provide more insight into rates of area change and variability as a result of discharge patterns.

Erosion and deposition are also affected by vegetation, which has been steadily encroaching on the subaerial sandbars since the closure of Glen Canyon Dam and the elimination of snow-melt-driven floods. At many sandbars, area within the HIZ is becoming

increasingly vegetated and stabilized, as well as portions of the CFZ and the FFZ at some sites (Sankey et al., 2015; Kasprak et al., 2018; Hazel et al., *in press*). Because the research focuses on narrow reattachment bars with little vegetation cover, we do not examine the effects of vegetation cover in this study.

#### ***4.2.3.2. Observations from past sandbar studies***

Annual topographic surveys have been conducted using electronic total stations at least once per year at up to 44 different sandbar monitoring sites since 1990 (Hazel et al., 2008; Hazel et al., *in press*). Results from annual surveys have shown that properly timed controlled floods have increased average sandbar volume during the MLFF period throughout the Grand Canyon, though increases in sandbar size resulting from controlled floods are mostly eroded away in the months following controlled floods during normal dam operations (Grams et al., 2010; Hazel et al., 2010; Grams et al., 2015; Hazel et al., *in press*). Surveys indicate that the observed increases in sandbar volume often manifests as increases in vertical sandbar thickness. They show that area is a poor predictor of volume (Hazel et al., *in press*). Despite the poor relationship between area and volume, area measurements can elucidate differences in cyclic erosional and depositional behavior in response to sub-annual flow patterns which cannot be resolved by annual surveys. Furthermore, because one of the goals of maintaining sandbars is to provide substrate for water-based recreation, the lateral area is likely as important of a consideration for recreational values as volume is.

The magnitude of deposition and long-term trends in bar behavior are spatially variable even over within the same reach (Grams et al., 2010; Hazel et al., 2010; Grams et al., 2015). Analysis of topographic and bathymetric data suggested grouping sandbars based on the geomorphic setting of sandbar sites, which explains much of the spatial variability observed in annual data (Mueller et al., 2018). Mueller et al. (2018) found that dynamic

reattachment bars, such as 22-mile, 30-mile, and 145-mile sandbars, are narrow, unvegetated, and experience large stage change relative to other bar types: including wider, more vegetated reattachment bars, separation bars, and undifferentiated eddy sandbars. Sandbars in this narrow reattachment classification also showed the largest increases in vertical bar thickness from controlled floods (Mueller et al., 2018). Despite the usefulness of these sandbar groupings to explain annual trends, observations from the remote camera imagery indicate that at finer temporal resolutions, sandbar behavior may differ significantly even among sandbars within the same geomorphic grouping. Until now, those differences have not been examined in detail during the MLFF flow period. We provide the first detailed quantitative analysis examining cyclic erosion and mass failure occurrence at multiple sites during current dam operations.

Previous studies using remote cameras deployed during the end of UDO and the beginning of the IF flow periods indicated that annual surveys likely underestimate erosion and deposition rates (Cluer, 1995; Dexter and Cluer, 1999). They found a cyclic pattern of erosion and deposition at sandbar sites, where weeks to months of deposition followed similar periods of gradual or rapid erosion (Dexter and Cluer, 1999). While observations of time-lapse imagery show that sand eroded from the FFZ is often redeposited at lower elevations in the eddy, erosion of any type provides an opportunity for fine sediment within the eddy to mobilize and be transported downstream. Observations of mass failures that occurred during repeat bathymetric surveys of two sites during the 2008 controlled flood showed that sediment mobilized during one subaqueous mass failure was deposited in the channel, where it was rapidly transported downstream (Wright and Kaplinski, 2011).

Attempts to model erosion and deposition during controlled floods, which did not include the mechanisms for mass failure, over-predicted low-elevation deposition and led to the formation of slopes at steeper than the angle of repose (Sloff et al., 2009; Sloff et al.,

2012). Mass failures then, at sites where they occur frequently, may be significant mechanisms by which large amounts of fine sediment are transported out of an eddy, and therefore may have a significant impact on the balance of fine sediment at sites where they occur (Dexter and Cluer 1999; Alvarez et al., 2017). Thus, understanding the specific flow patterns and antecedent conditions that favor rapid erosion through mass failure versus gradual erosion is a significant subject of interest for adaptive management of GCD and can improve our understanding of the morphodynamics of fluvial bedforms.

Several lines of evidence suggest that increases in discharge may result in fluvial processes leading to formation of high-angle slopes, or lead to erosion, triggering mass failures. Observed and modeled changes to recirculation zones in Grand Canyon during rising discharge led to the following: the development well-defined upstream currents along the shore (Andrews et al., 1999); lengthening and narrowing of the recirculation zone (Schmidt 1990; Wright and Kaplinski, 2011); reattachment and separation point migration (Schmidt 1990; Wright and Kaplinski, 2011); the appearance of secondary eddies, which scoured deposits (Schmidt, 1990; Cluer, 1995); helical flow patterns (Wright and Kaplinski, 2011); increased high-elevation deposition and low-elevation erosion forming steep subaqueous slopes (Wright and Kaplinski, 2011); and non-periodic eddy pulsations (Rubin and McDonald, 1995; Alvarez et al., 2017). The effects of these discharge-induced changes in flow patterns and deposition are likely to vary between eddies depending on the site-specific geometry and antecedent conditions. For example, differences in reattachment point migration and secondary flow structures were observed at adjacent sites during the 2008 controlled flood. These differences were attributed to the proximity of a downstream constriction at one of the sites (Wright and Kaplinski, 2011) and differences in recirculation variability across discharges has been measured at several other sites (Schmidt, 1990).

Thus, we explored differences in area change and mass failure occurrence in response to the same discharge patterns at three sites in the narrow reattachment sandbar grouping. Because of the connection between discharge patterns and variability in secondary flow structures, we examined metrics related to the flow variability, the flashiness of discharge, the minimum, mean, maximum, and range of discharge. Our goal was to better resolve which aspects of discharge patterns are most strongly related to sandbar area change and to understand what, if any, discharge patterns lead to similar effects at similar sites throughout the canyon providing policy relevant findings to inform dam operations.

### **4.3. MATERIALS AND METHODS**

We segmented 13,579 images from 2011 - 2020 to examine the utility and accuracy of using oblique images to measure daily and monthly changes in subaerial sandbar area at 22-mile, 30-mile, and 145-mile (Figure 4.1) We correlated changes in the mean monthly area with monthly flow metrics. We estimated the magnitude of daily and monthly changes to examine the variability in the monthly sandbar area in years when controlled floods occurred versus years without controlled floods. In addition to the detailed area estimates at the three sites, we examined time-lapse videos created from imagery at 41 sites spanning up to 10 years to characterize general sandbar response to current discharge patterns and controlled floods. Then we identified mass failures as events providing conservative estimates of mass failure frequency. Next, estimates of changes in the area at various elevations and the magnitude of erosion during mass failures were assessed.

#### **4.3.1. Measuring changes in sandbar size**

Quantitative analysis of sandbar dynamics was completed by measuring the emergent portion of the sandbar in each remote camera image. This is the area above the water's surface, which varies between images depending on both the size of the bar and the discharge

in the image. Over a single day, while sandbar size may remain constant, differences in the water surface elevation, or *stage*, reveal more or less of the sandbar affecting the observable size of the emergent sandbar area. Comparing image-derived area estimates, hereafter referred to as *image-areas*, from imagery required each image to be associated with a stage-elevation based on gage data and the travel time of a wave through the canyon. Once the image stage had been estimated, images were separated into discharge bins, so comparisons between images captured changes due to erosion or deposition and minimized changes in area associated with stage change.

#### **4.3.1.1. Estimating discharge**

We estimated discharge in each image by routing the flow from the nearest gage to each site using Wiele and Griffin's (1997) unsteady flow model. This flow model allowed us to predict the lag time, or the time it takes a wave to travel between the nearest gage and each site. The six gages on the Colorado River in the Grand Canyon are 0938000- Lees Ferry at RM30, 09280505-Thirty-Mile at RM30, 09283100-Little Colorado River at RM62, 09402500-Phantom Ranch at RM87, 09404120 - National Canyon at RM166, and 09494200-Diamond Creek at RM225. For each image timestamp, the lag time was added or subtracted, depending on location of the gage relative to each site, providing a *gage-wave time*. The gage-wave time is the time the parcel of water viewed in the image passed the nearest gage. Because discharge data is recorded every fifteen minutes, and the gage-wave time often did not correlate precisely with a gage reading (+/- 7.5 minutes), we used linear interpolation to estimate discharge between the two gage readings most proximate to the gage-wave time providing a more accurate estimate of gage-wave discharge. The segmented sandbar area was then taken as the planar area of the sandbar at the estimated gage-wave discharge. Images were then grouped into discharge bins, used as a proxy for elevation, with a bin width of  $28m^3s^{-1}$  ( $1,000ft^3s^{-1}$ ). A sensitivity analysis demonstrated that the mean changes in water

surface elevation at our three target sites were between 17-20 *cm* for each  $28m^3s^{-1}$  bin.

Based on the mean survey hypsometries of each site, mean area differs between  $22m^2$  and  $92m^2$  over a 17-20 *cm* elevation range.

Several image pre-processing steps were completed before area measurement. Camera calibration was done *in situ* by taking 20-30 photos of a black and white checkerboard pattern on an aluminum plate at multiple orientations and processing those images using the camera calibration toolbox in Matlab. The calibration process determines the intrinsic matrix for the camera-lens system, which contains: the principal point, focal length, and image size, and allows for the removal of radial and tangential distortion with the *undistort()* function in the OpenCV library.

Since images are captured at an oblique angle, pixels within the image show portions of the sandbar whose real-world areas vary due to perspective. Therefore, images had to be warped and rectified to a known coordinate system. Rectified images appear as if viewed from above, with each pixel representing the same area on a known coordinate system. Rectification was done by placing eight to twelve black and white panels (1 *m* x 1 *m* in dimension) on each sandbar; panels were surveyed and used as ground control points (GCPs) to create a homography. A homography is a perspective transformation between the image coordinates (X, Y) of the GCPs in a reference image and their surveyed real-world coordinates (easting, northing). Rectification was done in Python 3 using the *findHomography()* function in the OpenCV library.

The entire image set was then registered or aligned to the reference image in Python 3, using a combination of Enhanced Correlation Coefficient and 2-dimensional fast-Fourier transform algorithms, removing the effect of slight changes in the field-of-view caused by camera maintenance or other disturbance (Evangelidis and Psarakis, 2008; Reddy and Chatterji, 1996; Grams et al., 2018). Registration allows the same homography to be used for

all images, minimizing the number of occasions where surveyed GCPs were needed and allowing for use of a single rectification for each site. For a more in-depth discussion of the methods used in image pre-processing, see Chapter 3 in this dissertation.

#### **4.3.1.2. CNN image segmentation**

The task of measurement also required that areas containing sand within each image be identified through a process known as segmentation. A residual U-Net, a type of Convolutional Neural Network (CNNs), was built, trained, validated, and tested using Python 3, Tensorflow, and Keras for the purpose of binary semantic segmentation to identify pixels that contain subaerial sand from those which do not. Conditional Random Field post-processing was used to improve prediction masks. CNNs with CRF post-processing have shown promise in segmentation tasks involving natural textures (Buscombe and Richie, 2018). Our segmentation had a mean  $F1 = 0.93$  and a mean *Cohen's Kappa* = 0.89 on a test dataset containing 100 images from the 22-mile, 30-mile, and 145-mile sandbars. Detailed descriptions of this neural network as well as considerations for training convolutional neural networks end-to-end for binary segmentation of natural textures can be found in Chapter 3 of this dissertation.

Practically, the accuracy for most images exceeded  $F1 = 0.93$ , but a few images with lower accuracies of  $F1 \leq 0.7$  reduced the average accuracy. These were generally images with significant glare. Rather than spending more time re-training the network, all 13,579 images were manually checked, where we focused on removing very poor segmentations in which portions of water with glare were classified as sand or where large portions of the sandbar were not classified as sand due to overexposure in the image highlights. This eliminated 1,281 images, bringing the total number of images down to 12,298. The final dataset contained 3,365 images at 22-mile, 6,550 at 30-mile, and 2,383 at 145-mile. The disparity in the number of images in part has to do with the success of registration techniques, which fail

more often at 22-mile and 145-mile due to lighting conditions, and increased number of images available at 30-mile. During 2014 and 2015, images were collected every hour at 30-mile, providing twice the number of images during those year.

#### **4.3.1.3. *Daily and monthly sandbar area***

The homographies applied during the rectification process reprojects imagery onto the Arizona State Plane Central coordinate system (FIPS 0202). Rectified images have pixels that measure  $0.01m^2$  in dimension. We summed the pixels, classified as ‘sand’ by our neural network, to provide an estimated area in square meters (image-area). Evaluating the accuracy of image-areas involved comparing those image-areas to survey-derived area measurements, referred to hereafter as *survey-area*, at the same water surface elevation on days when topographic surveys were conducted. Digital elevation models or DEMs of the surveyed bars allowed us to extract the aerial extent, or planar area, of sandbars at any elevation within the subaerial sandbar. Time-series of daily and monthly areas were constructed for each elevation bin, where images were available.

Most of our analysis focus primarily on the elevation bin centered at  $283m^3s^{-1}$  ( $10,000ft^3s^{-1}$ ) for several reasons: 1) this is the lowest discharge which appears in imagery in nearly every month of the photo record, 2) it is the elevation bin with the highest number of images, and 3) it is nearly always in the fluctuating flow zone during normal dam operations, thus it is likely very sensitive to changes in discharge pattern.

#### **4.3.1.4. *Estimating erosion and deposition magnitude***

Daily measurements compare image-areas within the same elevation bin from individual photos. Daily image-areas within each month were averaged to provide a mean monthly area estimate for each elevation bin and a monthly standard deviation in months where multiple images were available. The standard deviation measures the variability in size

over each month and contains errors from individual images making it a liberal estimator of within-month variation. Differences in the estimated area for each month and each daily observation were then compared to the nearest previous value. Positive differences were interpreted as deposition and negative differences as erosion. The change in area was then divided by the number of days or months between area estimates to provide daily and monthly net change rates. The change rates were then standardized (or normalized) using the mean area and standard deviation (or maximum area) in the image-area dataset at the elevation being examined or at the  $227m^3s^{-1}$  ( $8,000ft^3s^{-1}$ ) elevation when comparing multiple elevations. However, in the interest of estimating daily erosion and deposition rates rather than net change, we limited our daily magnitude analysis to observations that were less than ten days apart. We treat these daily net change rates as liberal estimates of erosion rate or deposition rate (depending on the sign (+/-) of the change) due to the error in each observation.

#### 4.3.1.5. Discharge patterns

We utilized 15-minute discharge data from the six gages to link sandbar response with dam operations. We computed the lag-time described in section 4.2.1.1 for each observation to create a local hydrograph for each site. The flow metrics (Table 4.1) were then calculated from local hydrographs. Some metrics were calculated for 24-hours of flow, and others were calculated for monthly flow reflecting the entire calendar month.

**Table 4.1** Discharge ( $Q$ ) metrics calculated for daily and monthly flow periods.

<b>Metric</b>	<b>Description or Formula</b>	<b>Scale</b>	<b>Unit</b>
mean, min., max., median, and range.	Summary statistics were calculated for the discharge period (day or month)	Daily & Monthly	$m^3s^{-1}$
$\Delta mean lag_1$ , $\Delta min lag_1$ ,	The change in these summary statistics relative to the previous month. For example:	Daily & Monthly	$m^3s^{-1}$

$\Delta_{max} lag_1,$ $\Delta_{range} lag_1$	$\Delta_{mean} lag_1 = mean(Q_t) - mean(Q_{t-1})$		
Standardized $\Delta$	The range in daily (or monthly) discharge RDQ (or RMQ) divided by the mean daily (or monthly) discharge MDQ (MMQ). A measure of monthly variability adapted from Bevelhimer et al. (2014) $DS\Delta = \frac{RDQ}{MDQ} \text{ or } MS\Delta = \frac{RMQ}{MMQ}$	Daily & Monthly	Unitless
Coefficient of Variation (CoV)	The standard deviation of discharge over a day (or month) divided by the mean daily (or monthly) discharge. Daily CoV = $\frac{SD \text{ daily } Q}{mean \text{ daily } Q}$ or Monthly CoV = $\frac{SD \text{ monthly } Q}{mean \text{ monthly } Q}$	Daily & Monthly	Unitless
Richards-Baker (R-B) Flashiness Index,  R-B Flashiness index	An index proposed by Baker et al. (2004) which measures oscillations in discharge relative to the total discharge. We also calculated the change in the R-B flashiness index at a lag of 1. R-B flashiness index = $\frac{\sum_{i=1}^n  Q_i - Q_{i-1} }{\sum_{i=1}^n Q_i}$ $\Delta R\text{-B flashiness index} = RB \text{ index}_t - RB \text{ index}_{t-1}$	Daily & Monthly	Unitless

We chose to use summary statistics such as mean, min, max, and range since they are the easiest to conceptualize and interpret. Standardized delta was adapted from the *daily standardized delta* described by Bevelhimer et al. (2014) and applied to both daily and monthly flows. This metric provides a measure of variability based on daily or monthly range scaled to the mean discharge. The coefficient of variation, or CoV, is similar except that it uses the standard deviation of discharge scaled by mean discharge. Richards-Baker Flashiness or the R-B Flashiness index measures the length of the hydrograph line over a given time period (Baker et al., 2004).

#### 4.3.1.6. Application-specific considerations

Remote imagery only captures flows that occur within the 8-hour daily imagery window. Because of the time required for water to travel from GCD downstream to each site,

the daily maximum or minimum discharge is not always captured in imagery. Seasonal flow patterns related to power generation result in higher flows in December, January, July, and August, and lower flows in March, April, October, and November, with the flows in other months falling somewhere in-between. Therefore, the image-area time-series at any elevation chosen will have some gaps.

At certain times of the year, imagery from various sites contains heavy glare due to the sun's angle reflecting off the sand or water, making image segmentation difficult. Our segmentation procedure also struggled with imagery during high flows when very little of the sandbar was showing, reducing the imagery analyzed during peak flows of controlled floods. We thus focused our analysis on changes in the area at the water surface elevation with the most imagery throughout the year at our target sites, the  $283m^3s^{-1}$  discharge elevation.

#### **4.3.2. Mass failure time-lapse analysis**

Time-lapse videos were created for 41 different sandbar sites using one image per day captured between 11:00 - 13:00 MST to minimize stage change between images and brightness differences. We utilized the OpenCV module in Python 3 to downsize imagery, extract the EXIF data, annotate the image with the timestamp, and combine the images into a video (\*.MP4 or \*.AVI) at a rate of seven frames per second.

Time-lapses show a week's worth of bar behavior each second. Gradual erosion or deposition happens over several seconds, representing weeks to months of imagery. Mass failures are easily observable as large changes in lateral area between frames and appear instantaneous in time-lapses. While it is likely that large, rapid changes in sandbar area are the result of different conditions, mechanisms and processes discussed in section 4.1.3.1, without detailed field measurements, we treat all noticeably large single-day changes in area as mass failures. Since most of these failures appear to initiate in the subaqueous portion of

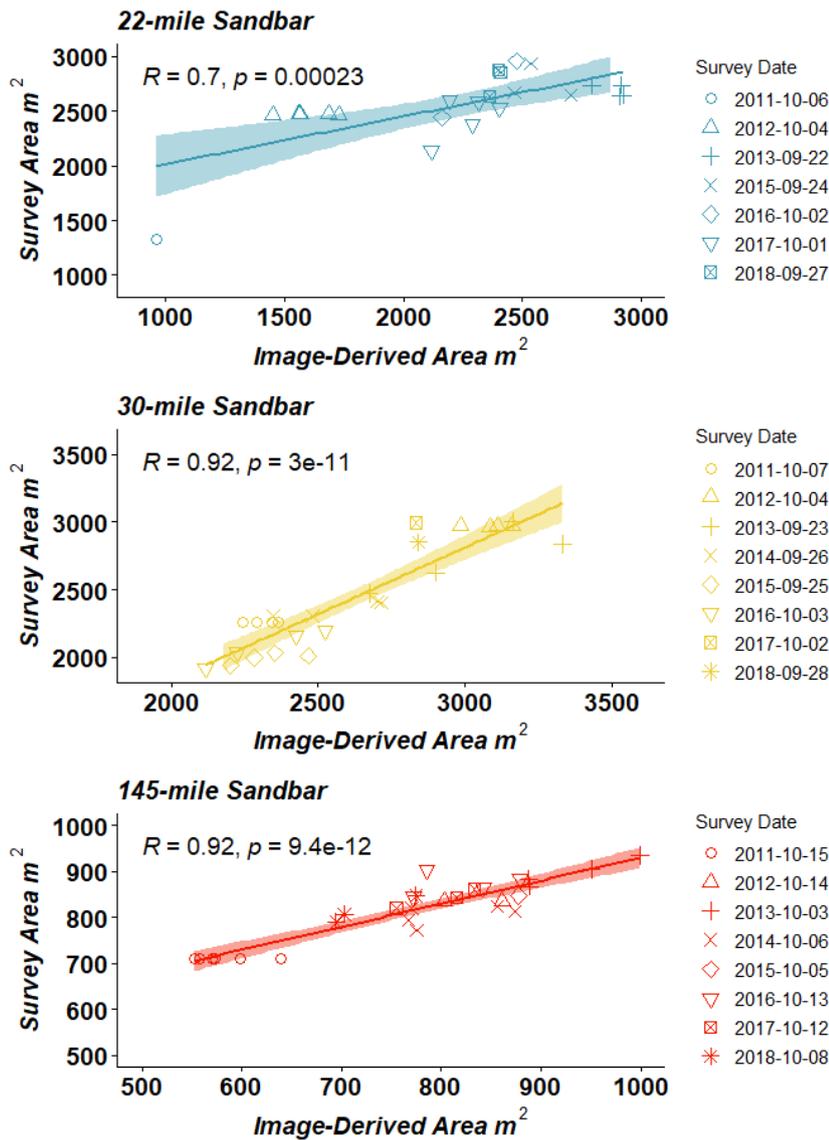
the bar, with varying effects on the subaerial part of the bar, some of the mass failures could not be detected by image-area analysis, preventing the development of a quantitative threshold for mass failures. Identification of mass failure events was conducted by qualitative visual assessment of time-lapse videos. The timing of each of these events was documented based on the video timestamp, and the imagery surrounding the event was scrutinized to confine the timing of mass failures to a 24-hour window from noon to noon, which we term the *failure day*. We calculated the daily flow from noon-noon because often we could not tell the precise time of mass failures that occurred overnight. We only know that they occurred between the first image on the day the failure was identified and the last image from the day before.

Localized flow from each failure day was summarized into daily flow metrics described above in section 4.2.1.6. Flow metrics were compared between three different intervals: a) Failure days, b)  $lag_2$  the days which preceded failure days by 48 hours, and c) randomly selected days (*random days*) from the study period 2008 - 2021 with failure days removed. The  $lag_2$  group was created to compare failure days to flows with similar seasonal patterns as failure days occurring two days prior. We treat the  $lag_2$  and random groups as control groups.

## 4.4. RESULTS

### 4.4.1. Accuracy of image-derived area

Image-areas were strongly correlated with the survey-areas at the three selected sites with correlation coefficients ranging between 0.7-0.96 (Figure 4.5). *RMSE* values were relatively small at  $243m^2$ ,  $151m^2$ , and  $26m^2$  at sites 22-mile, 30-mile, and 145-mile, respectively, comprising roughly 8.8%, 4.8%, and 2.9% of the maximum surveyed site area.



**Figure 4.5** Correlations between image-areas and survey-areas at the three selected sites. Imagery, excluding mass failure events, from within 1-3 days of annual surveys were processed, segmented, and image-areas were measured. Image-area estimates were compared to survey-areas at the same elevation to estimate image-area accuracy at each site. Shaded areas surrounding lines show the 95% confidence interval of the Loess smoothed linear model.

Image-derived area estimates contain several sources of error from image processing: error in camera calibration, registration, rectification, and semantic segmentation. In addition, there is error in the timestamp based on the accuracy of the camera's internal clock and error in the estimated travel time of the imaged discharge to the nearest gage (Table 4.2).

**Table 4.2** Sources of error in image-derived area estimates. Errors affect either area estimation or the elevation to which that area is assigned because of errors in gage-wave time. The errors vary by site, elevation, and camera.

Source of Error	Description of Error	Estimated Error
<b>Errors in Area Estimation</b>		
<b>Camera Calibration</b>	The distance between a pattern keypoint detected in a calibration image and the corresponding world point projected into the same image	0.30 - 0.71 pixels
<b>Registration</b>	Translation error may occur between any subject image and the reference image to which it is being registered. Significant differences in lighting between a subject image and a reference image can cause misregistration.	Unknown; images with obvious registration errors were excluded.
<b>Rectification</b>	Mean squared distance between surveyed ground control point locations and re-projected ground control points using homography	0.1 - 2 meters error in the location or size of each pixel when re-projected
<b>CNN Segmentation</b>	Binary classification error using a test set containing 25 images from each sandbar site. Lima et al., ( <i>in review</i> )	mean $F1 = 0.93$ , mean (~7% of max site area) Cohen's Kappa = 0.89
<b>Errors in Image Discharge Estimation<sup>1</sup></b>		
<b>Time-stamp</b>	Error in the camera's internal clock	+/- up to 10 minutes or $18 m^3 s^{-1} (\leq 59m^2)^2$
<b>Lag-Time</b>	Error in estimating the time wave travel times in Grand Canyon based on the Wiele and Griffin (1997) unsteady flow model	$\leq 1$ hour as discharge in increasing ( $\leq 113m^3 s^{-1}; \leq 371m^2$ ) <sup>2</sup> & $\leq 3.2$ hours as the discharge is decreasing. ( $\leq 227m^3 s^{-1}; \leq 746m^2$ ) <sup>2</sup>

<sup>1</sup> Lima et al. (2019) compared estimated-water surface elevation using timestamps and lag-time to surveyed water surface elevation at the 30-mile site. The data had an  $R$ -squared = 0.957 and an RMSE = 0.053 m, or ~5cm

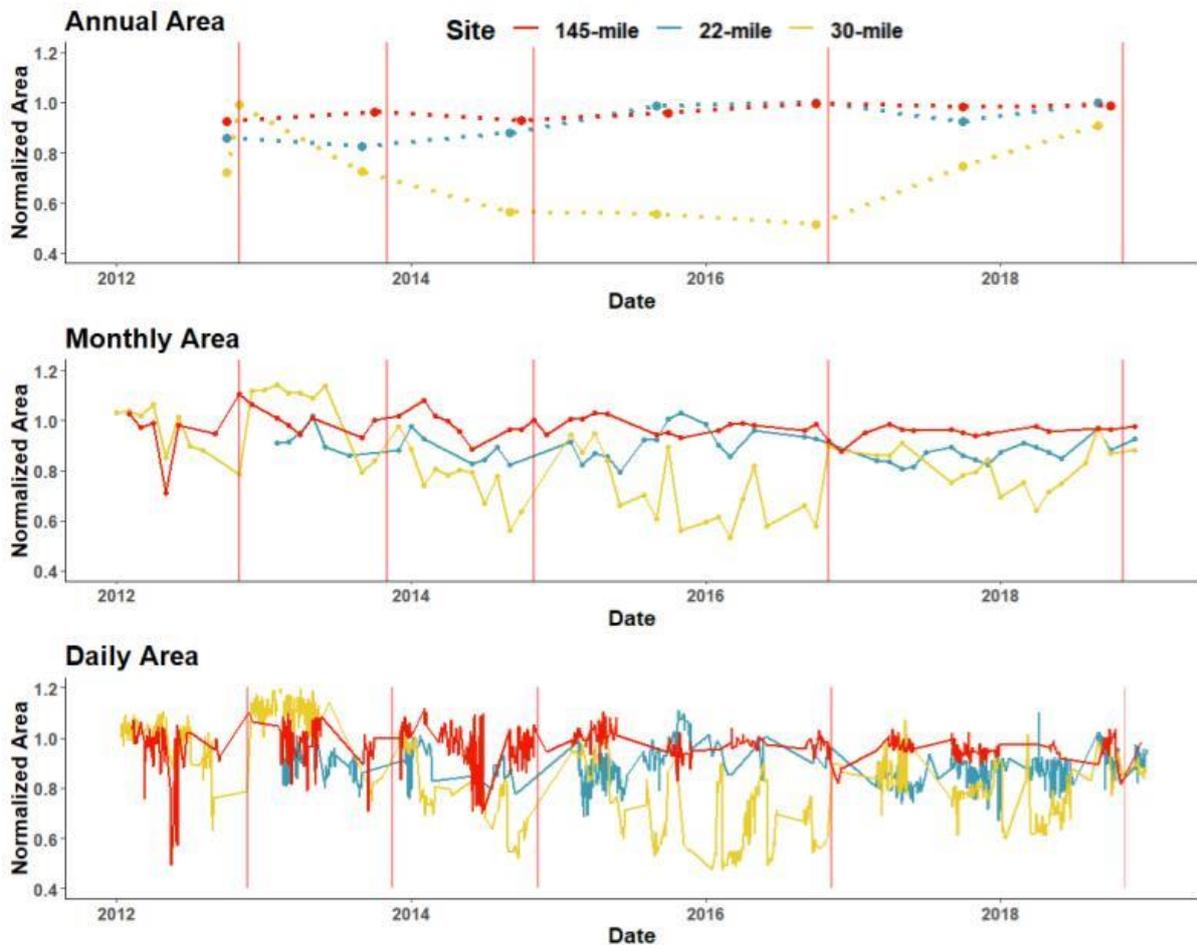
<sup>2</sup>Area-error estimates related to discharge change are based on area differences at different water surface elevations from average surveyed hypsometries from 22-mile, 30-mile, and 145-mile.

#### **4.4.2. Image-area erosion and deposition rates at 22-mile, 30-mile, and 145-mile sandbars**

Rates of daily erosion and deposition at the  $283m^3s^{-1}$  stage, derived from image-areas differenced at a lag of 1 day were the same order of magnitude at all three sites. Observations recorded deposition slightly more often than erosion at 30-mile and 145-mile sandbars, but the opposite was true at 22-mile sandbar. The daily erosion rate at 22-mile was  $-117m^2$  or  $-3.7\%$  of max sandbar area, and the daily deposition rate was  $106m^2$  or  $3.4\%$  of max sandbar area, where 49% of observations showed erosion, and 48% of observations showed deposition. Only 3% of observations occurred over intervals longer than ten days and thus were excluded. At 30-mile sandbar, the daily erosion averaged  $-120m^2$ , or  $-2.6\%$  of the maximum area, with a deposition rate of  $112m^2$ , or  $2.5\%$  of max sandbar area, where 46% of observations at 30-mile showed erosion, and 50% of observations showed deposition. At 145-mile, 47% of observations showed erosion, which averaged  $-34m^2$  or  $-3.4\%$  of the max area, while 48% of days showed deposition with an average of  $35m^2$  of deposition or  $3.5\%$  max sandbar area.

##### **4.4.2.1. Time-series of net area change**

We examined the time-series of sandbar area at the  $283m^3s^{-1}$  discharge elevation to show differences in intersite variability and behavior at three timescales. The annual time-series was created with survey-area measurements; monthly time-series using averaged image-areas measurements for each month; and daily time-series using individual image-areas (Fig. 4.6).



**Figure 4.6** A comparison of image-derived normalized sandbar area from annual, monthly, and daily time-series at the elevation associated with the  $283\text{m}^3\text{s}^{-1}$  discharge at the 22-mile, 30-mile, and 145-mile sandbars. Values were normalized by dividing each value by the maximum survey-area for the  $283\text{m}^3\text{s}^{-1}$  elevation from 1990 -2018. The monthly and daily area values exceed 1 in some cases because annual surveys do not always capture the maximum sandbar extent. Vertical red lines indicate controlled floods.

These three time-series show similar trends, but a more detailed picture of cyclic erosion and deposition is apparent in the daily and monthly series. The annual series shows that 30-mile steadily decreased in size after the 2013 controlled flood until the trend reversed following the 2016 controlled flood. However, daily and monthly series show that the decline in size was not as large as is implied by the annual data and that a period of high variability was occurring with pronounced peaks in sandbar area reaching near maximum size.

When examining the standard deviation of measurements at all three scales and the absolute value of the average change between observations at each scale, significant

differences in area change between sites are apparent (Table 4.3). Within annual surveys, the 30-mile sandbar experiences 2 -4 times as much net change in its  $283m^3s^{-1}$  area between subsequent years as the other two sites (Table 4.3). The standard deviation of mean monthly areas and the average net change in monthly mean areas are 2-3 times higher at 30-mile than at 145-mile and 22-mile sandbars (Table 4.3). Daily measurements, by contrast, show less of a disparity in the average change between images, but do show over twice as large of a standard deviation at 30-mile than the other two sites (Table 4.3).

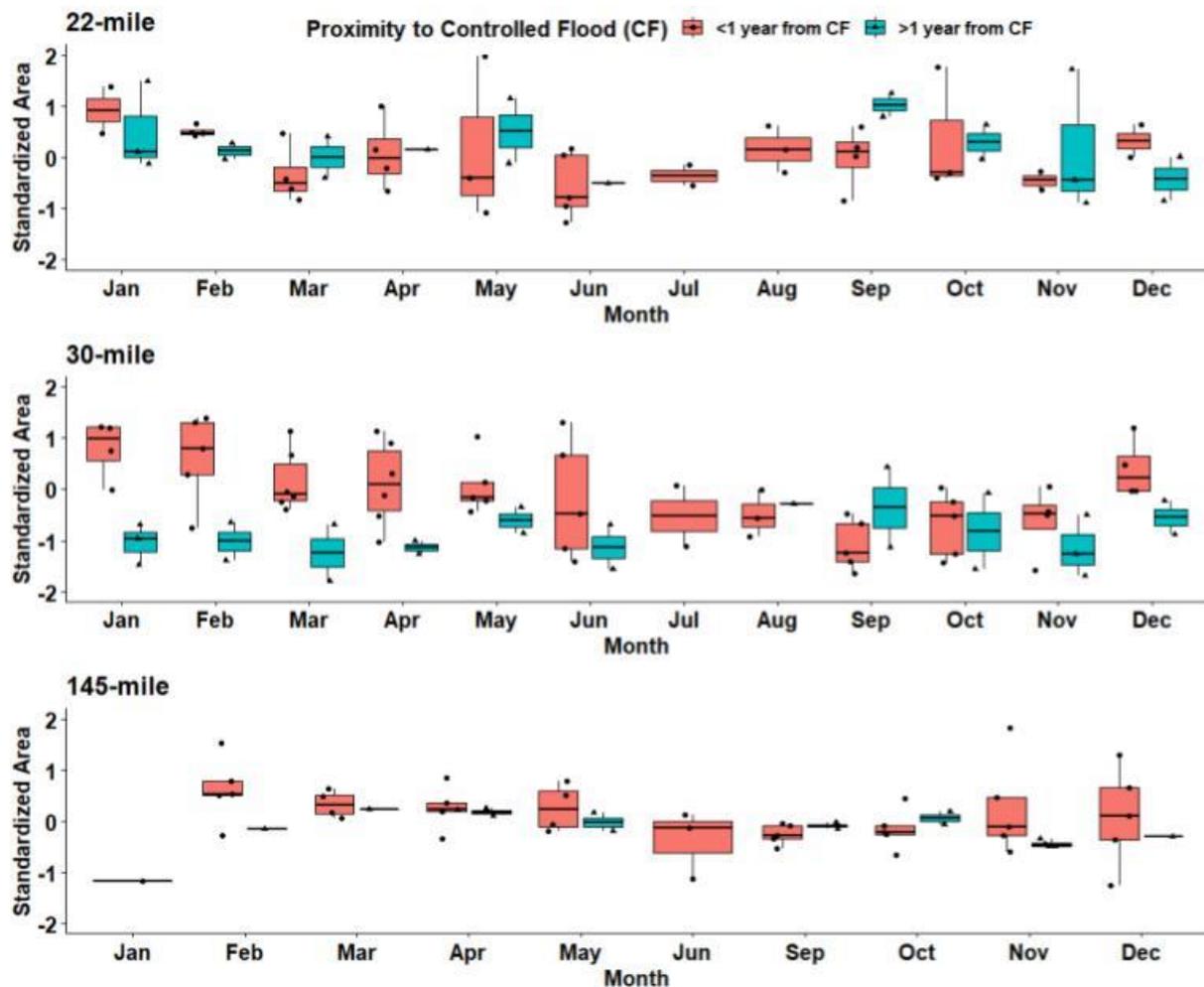
**Table 4.3** Variability in  $283m^3s^{-1}$  area at three timescales: annual, monthly, and daily. We present two measures of variability for each timescale: 1) The standard deviation across all measurements at each scale (Rows 1,3,& 5) and 2) The mean absolute value of change between observations at each timescale with a lag of 1 (Rows 2,4,& 6). These values are reported in Area ( $m^2$ ) and normalized area, or the size divided by the maximum surveyed area at  $283m^3s^{-1}$ .

	22-mile		30-mile		145-mile	
	Area ( $m^2$ )	Norm. Area	Area ( $m^2$ )	Norm. Area	Area ( $m^2$ )	Norm. Area
<b>1</b> SD Annual Surveyed Area	183	0.071	603	0.164	26	0.028
<b>2</b> Mean Annual Area Change	171	0.067	751	0.204	29	0.032
<b>3</b> SD Mean Monthly Image Area	148	0.058	603	0.164	49	0.056
<b>4</b> Mean Monthly Image Area Change	127	0.050	368	0.102	40	0.045
<b>5</b> SD Daily Image Area	184	0.072	751	0.194	66	0.074
<b>6</b> Mean Daily Image Area Change	136	0.053	159	0.043	49	0.055

#### 4.4.2.2. Response to controlled floods

Image-area estimates allowed us to examine differences in monthly area in years with and without controlled floods. The 30-mile site has the largest response to controlled floods at the  $283m^3s^{-1}$  stage elevation (Fig. 4.7). At all sites, the area is higher in the month

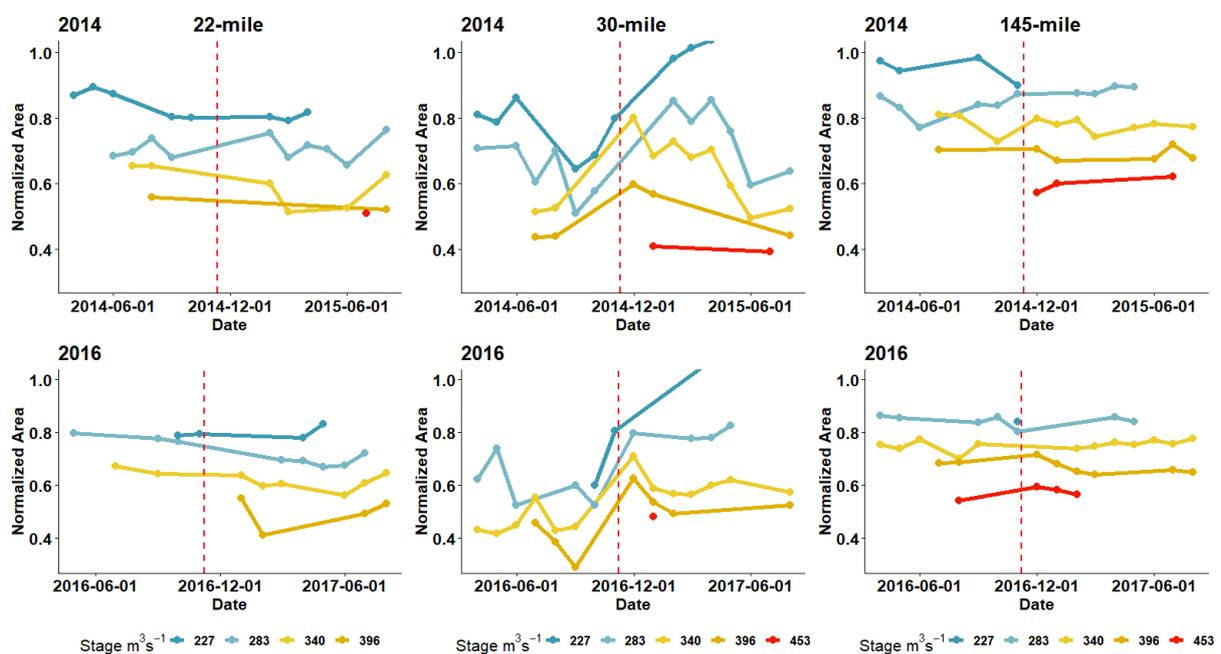
immediately following controlled floods (December), but that difference increases and peaks between January and February and then the differences begin to dampen. The effects of suspended sediment concentration are also important in determining response in future analyses.



**Figure 4.7** Standardized mean monthly sandbar area at the elevation associated with  $283m^3s^{-1}$  in years with controlled floods (red) vs. years without controlled floods (blue) at 22-mile, 30-mile, and 145-mile sandbars. Dots indicate individual mean monthly values. Controlled floods typically occur in November. The months June and July are absent from 145-mile because images of the  $283m^3s^{-1}$  rarely occur during daylight hours at that site during the summer.

A closer look at monthly changes in area at several discharge elevations before and after the 2014 and 2016 controlled floods at all three sites show just how much more dynamic 30-mile is in its response to controlled floods than 145-mile or 22-mile (Fig. 4.8). At all three

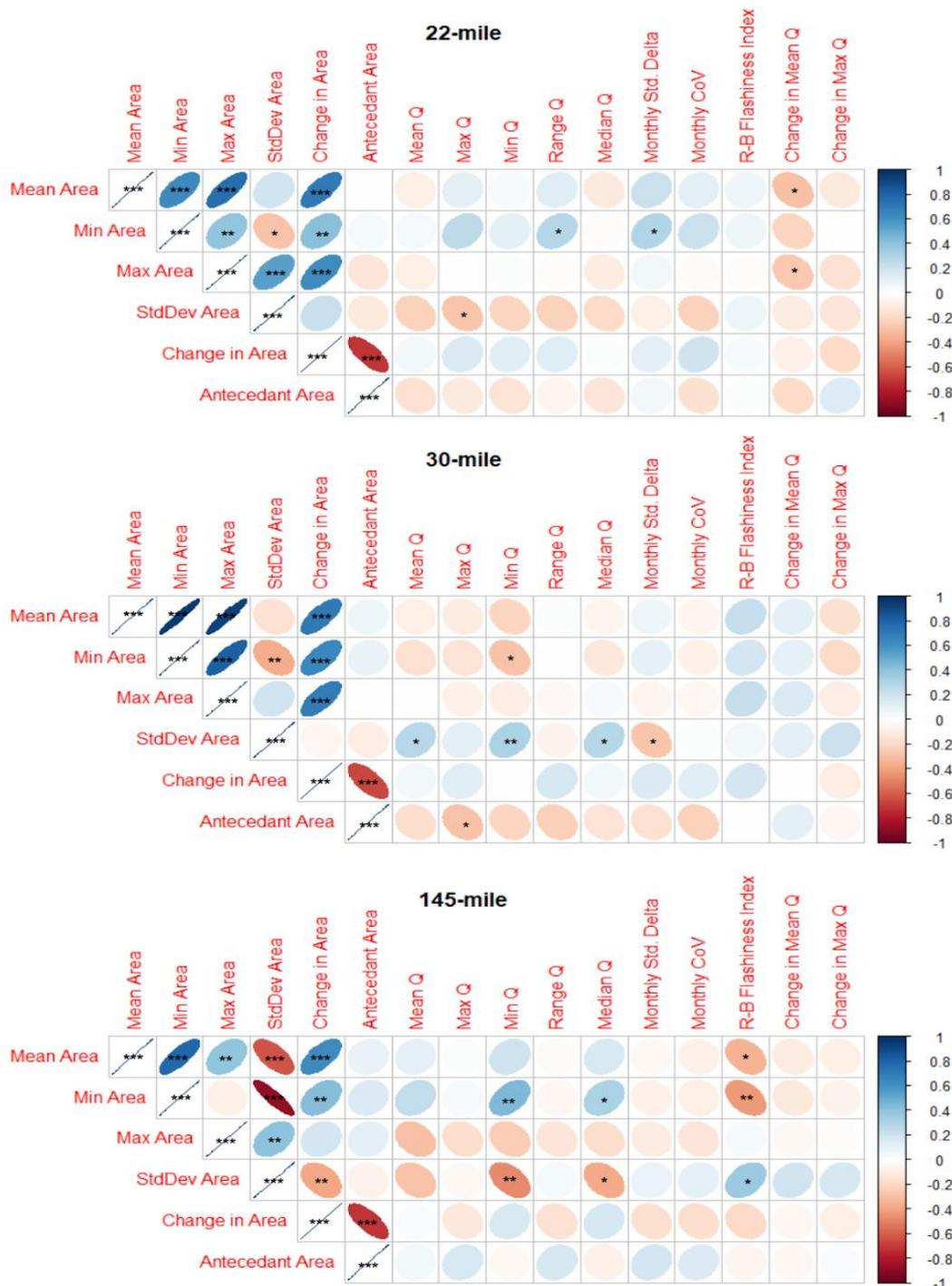
sandbars, max area following a controlled flood occurs in the first monthly observation following the flood at high elevations ( $340m^3s^{-1}$  and  $396m^3s^{-1}$ ), while maximum area at lower elevations ( $227m^3s^{-1}$ ) is lagged by 1-4 months. The 30-mile sandbar is smaller in terms of normalized area before controlled floods and then grows significantly after controlled floods, whereas the response at 22-mile and 145-mile are less extreme, starting from a larger size and gaining small amounts of area at multiple elevations.



**Figure 4.8** Time-series of image-derived mean monthly area at 22-mile, 30-mile and 145-mile before and after the 2014 and 2016 controlled floods (red dashed lines). Five water surface elevations within the FFZ are shown. Imagery, and thus mean monthly area measurements were not available at some elevations during some months at 145-mile because the peaks and troughs of daily discharge did not always arrive at 145-mile during daylight hours. Area gains are lagged by a few months at lower elevations and coincide with area loss at higher elevations suggesting the downslope movement of fine sediment following controlled floods.

#### 4.4.2.3. Monthly net area changes and monthly flow

Our second objective aims to correlate sandbar behavior with monthly flow patterns, which reflect seasonal changes in power demand and dam operations. We compared monthly flow metrics described in Table 4.1 to monthly measures of sandbar area and net change relative to the previous month at the elevation associated with  $283m^3s^{-1}$  (Fig. 4.9).



**Figure 4.9** Correlation matrices for 22-mile, 30-mile, and 145-mile showing relationships between mean monthly area, minimum monthly area, the standard deviation of area measurements within a month, change in mean area relative to the previous month (which can be either positive or negative), normalized area from the previous month (antecedent area which ranges from 0 to 1), and flow variables: monthly mean discharge (Q), maximum monthly Q, minimum monthly Q, the range of monthly Q, median Q, monthly standardized delta, monthly coefficient of variation, monthly Richards-Baker Flashiness Index, change in monthly mean Q relative to the previous month, and change in maximum monthly Q relative to the previous month. Descriptions of these metrics are found in Table 4.1. and in section 4.1.6. Significant relationships between variables are shown with stars (*p.values*: \* = 0.05; \*\* = 0.01, \*\*\* = 0.001).

For all sites, there is a significant and strong negative relationship between change in area and antecedent bar size (Fig. 4.9) suggesting that the larger a bar was the previous month, the more erosion or negative change in area is occurring, and vice-versa. As indicated by the R-B flashiness index, the flashiness of discharge has significant negative relationships with mean and minimum monthly area at 145-mile. However, this flow metric is not significantly related to metrics of sandbar area at the other two sites. Increases in the mean discharge relative to the previous month appear to have a significant negative relationship with mean and maximum monthly area at 22-mile only. The minimum monthly discharge does not show any significant relationships with area metrics at 22-mile but does appear significant at 30-mile and 145-mile, where it is significantly related to both monthly area variation (the standard deviation of monthly area measurements) and the minimum areal extent-- though the relationships are opposite between 145-mile and 30-mile. The range of monthly discharge has a positive and significant relationship with the minimum monthly area measured at 22-mile only. The median monthly discharge shows a significant relationship with the standard deviation of area measurements within a given month at 30-mile and 145-mile, but the relationships are opposite.

#### **4.4.3. Mass failures in time-lapse analysis**

Analysis of time-lapse videos of 41 sandbar sites in the period between 2008 - 2021 revealed a total of 226 mass failure events occurring at 27 sites and no observed failures at 14 sites. There were seven sites, where ten or more mass failures were observed. Across sites from Lees Ferry (RM-0) to Diamond Creek (RM-225), there does not seem to be a longitudinal trend in failures. The lengths of these records varied between sites depending on image availability as shown in Table 4.4.

**Table 4.2** Number of observed mass failures in time-lapse videos and length of the time-lapse record. Mass failures were observed at 27 of 41 sites. RM is the site location in river miles downstream of Lees Ferry. Time-lapse start (TL-start) and TL\_end are the start- and end-dates of time-lapses, Total days in record (Rec. Len. Days) and mean recurrence interval (Mean Rec. Int. Days) is the number of failures divided by Rec. Len. Days. If no mass failures were observed an NA is in the frequency interval column.

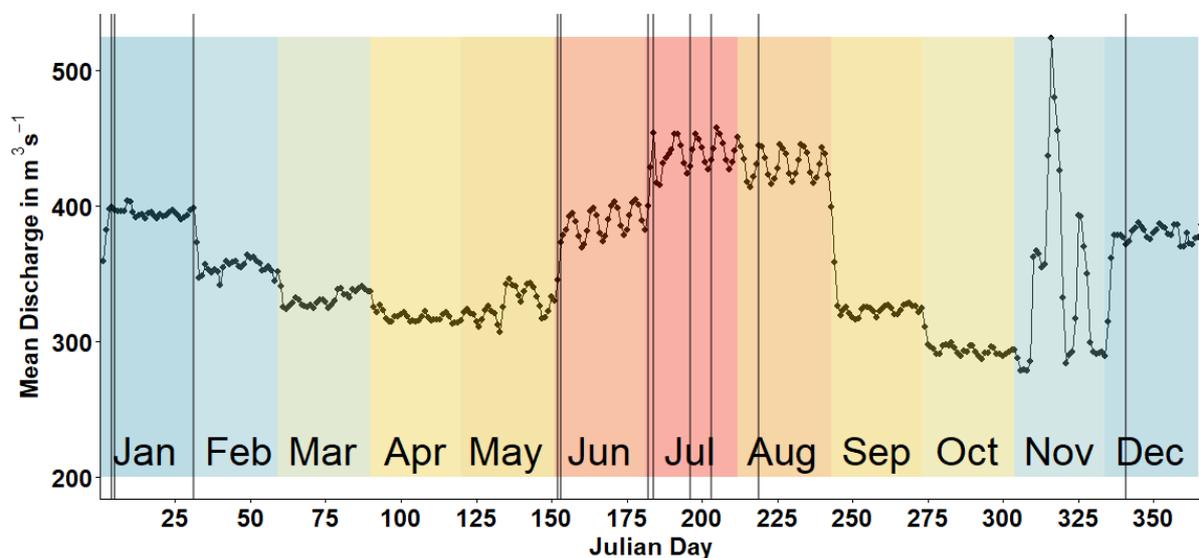
Site#	Name	RM	TL_start	TL_end	Mass Failures	Rec. Len. Days	Mean Rec. Int. Days
0025L	Cathedral Wash	2.5	2010/02/18	2017/02/11	4	2,550	638
0081L	Jackass Camp	8.1	2012/02/18	2017/02/12	0	1,821	NA
0089L	9-Mile	8.9	2012/10/04	2016/05/06	0	1,310	NA
0166L	Hot Na Na Wash	16.6	2014/05/03	2015/09/21	2	506	253
0220R	22- mile	22	2010/02/19	2020/02/01	0	3,634	NA
0235L	Lone Cedar	23.5	2013/02/17	2020/02/01	3	2,540	847
0307R	30-mile	30.7	2010/01/01	2021/02/21	50	4,069	81 days
0319R	South Canyon	31.9	2012/02/02	2016/12/31	1	1,794	1,794
0333L	Redwall Cavern	33.3	2013/09/24	2016/05/13	0	962	NA
0414R	Buckfarm	41.4	2010/02/21	2021/02/19	29	4,016	138 days
0434L	Anasazi Bridge (43-mile)	43.4	2012/10/07	2016/05/19	0	1,320	NA
0445L	Eminence	44.5	2012/01/01	2021/02/19	6	3,337	556
0450L	Willie Taylor	45	2008/03/02	2021/02/21	15	4,739	316
0476R	Saddle	47.6	2011/10/10	2021/02/21	2	3,422	1,711
0501R	Dinosaur	50.1	2011/10/10	2021/02/20	5	3,421	684
0515L	51-mile	51.5	2010/02/06	2021/02/21	16	4,033	252
0559R	Kwagunt Marsh	55.9	2011/01/01	2016/05/16	3	1,962	654
0566R	Kwagunt Beach	56.6	2013/09/27	2017/02/16	2	1,238	619
0629R	Crash Canyon	62.9	2013/02/22	2015/06/14	0	842	NA
0651R	Carbon	65.1	2012/10/09	2021/02/23	12	3,059	255
0658L	Above Lava Chuar	65.8	2008/02/10	2017/02/20	1	3,298	3,298
0661L	Palisades	66.1	2011/02/24	2016/05/16	0	1,908	NA
0688R	Tanner	68.8	2011/10/11	2016/05/17	1	1,680	1,680
0817L	Grapevine	81.7	2014/10/28	2017/02/15	0	841	NA
0846R	Clear Creek	84.6	2010/10/01	2021/02/24	0	3,799	NA
0917R	Above Trinity Camp	91.7	2012/10/11	2016/05/18	0	1,315	NA

0938L	Granite	93.8	2012/10/01	2021/02/25	9	3,069	341
1044R	Emerald	104.4	2012/10/01	2021/02/26	7	3,070	439
1194R	Big Dune	119.4	2011/01/01	2021/02/26	0	3,709	NA
1227R	122-mile	122.7	2010/02/27	2020/02/01	3	3,626	1,209
1233L	Upper Forster	123.3	2014/05/17	2021/02/27	2	1,239	1,239
1377L	Football Field	137.7	2010/02/28	2021/02/26	4	4,016	1,004
1459L	Above Olo (145-mile)	145.9	2010/02/27	2019/11/27	24	3,560	148
1671L	Lower National	167.1	2013/03/01	2016/11/12	0	1,352	NA
1726L	Below Mohawk (172-mile)	172.6	2011/01/01	2021/02/27	10	3,710	371
1833R	Below Chevron	183.3	2011/03/05	2021/02/28	0	3,648	NA
1946L	Hualapai Acres	194.6	2010/03/02	2016/11/15	0	2,450	NA
2023R	202-mile	202.3	2011/01/01	2016/11/15	1	2,145	2,145
2133L	Pumpkin springs	213.3	2013/10/06	2021/04/13	7	2,746	392
2201R	middle 220-mile	220.1	2013/10/06	2021/03/01	1	2,703	2,703
2255R	Above Last Chance	225.5	2010/03/03	2016/10/14	1	2,417	2,417

After identifying the approximate timing of each failure, we used lag-time estimates for each site to identify the discharge at Lees Ferry coinciding with the mass failure which allowed us to compare dam release patterns to mass failure occurrence, rather than the timing of when water from the dam arrived at each site, hours or days later. This allowed us to compare dam release patterns to mass failures rather than Mass failure occurrences, which tend to be concentrated in high-flow months and are conspicuously absent from lower flow months within the decadal hydrograph. Mass failures are sometimes clustered, with one failure following shortly after another.

Mass failures occur most often during water released from GCD on the first day of the month (8% of all mass failures) when flows are adjusting to the new monthly patterns, which is 2.5 times as frequent as would be expected if mass failures happened at the same rate each day. Mass failures most often occur on Mondays, Wednesdays, or Thursdays, with 17.3% of

mass failures occurring on each of those days, compared to Saturdays, which corresponds with just 8.0% of mass failures and lower weekend flows. The day of the year when failures were most common was Julian day 152 (2.7%), which is June 1, followed by Julian day 153 (2.2%), June 2 (Fig. 4.10). In comparison, the mean percentage of mass failures per Julian day was 1.4%. A total of 16.4% of all mass failures occurred in the month of January, followed by 13.7% in August and 12.8% in June. Meanwhile, October, November, and February had just 3.5% of mass failures each.

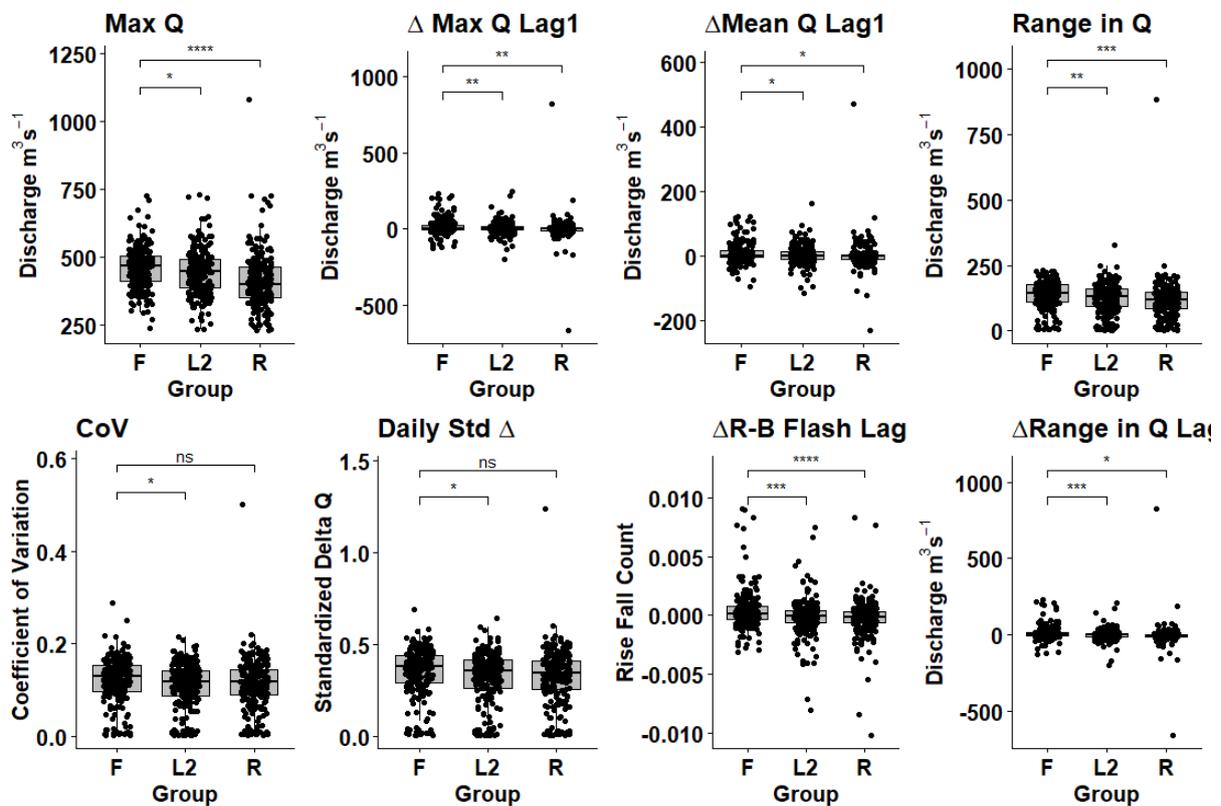


**Figure 4.10** Discharge patterns and mass failure frequency in the Grand Canyon. Average daily discharge from USGS gage 09380000 at Lees Ferry on the Colorado River from 2011 - 2021 was averaged by Julian Day. Vertical lines indicate Julian days with three or more mass failure observations. Varying colors indicate different months.

Monthly, annual, and seasonal patterns occur during the MLFF with the highest flows occurring in the summer months of June, July and August. High flows are also common in the winter, specifically in December and January. February is a transition month between high winter flows and lower spring flows (March and April), and May is a transition month between lower spring flows and high summer flows. September and October comprise the low fall flows, and November is typically low, but also has controlled floods. There is also a

subtle weekly flow pattern characterized by higher flows during the week and lower flows on the weekend when power demand is reduced.

We examined the relationship between daily flow patterns and mass failure occurrence (Fig. 4.11). There are significant differences in flow pattern on days when failures occurred compared to randomly selected days and two days prior to failures ( $lag_2$  days). Days when failures occurred generally have a higher mean, max, and range of discharges as well as higher variance (CoV), daily-std-delta, and increased flashiness (R-B Flashiness index) relative to the previous day, than  $lag_2$  or random days.



**Figure 4.11** Flow metrics on failure days (F),  $lag_2$  days (L2), and randomly selected days (R). Flow metrics shown in Table 4.1. Section 4.2.1.6. were calculated to determine if there is a link between dam operations under the modified low fluctuating flow protocol and the occurrence of mass failure events observed at 21 different sites (of 41) between 2010 - 2021. All flow metrics were calculated using 24-hour flow periods from noon - noon local time at each sandbar. We compare flow metrics on 226 days when mass failures were observed (F) to 226 randomly selected days within the same timeframe (R) and to 226 days which each preceded the failure day by two days (L2). We examined summary statistics for each 24-hour flow period as well as the Coefficient of Variation (CoV), the daily standard delta, and the Richards-Baker Flashiness index along with changes ( $\Delta$ ) in those metrics

relative to the previous 24-hour period indicating a shift in weekly or monthly. (*p.values*: \* = 0.05; \*\* = 0.01, \*\*\* = 0.001, *ns* = not significant)

#### **4.4.3.1. Mass failures at 30-mile and 145-mile sandbars**

No mass failures were observed at the 22-mile site during the study period. However, 50 mass failures were observed at 30-mile from 2010 - 2021, and 24 mass failures at 145-mile from 2010 - 2019. The detailed area analysis conducted at these two sites allowed us to measure mass failure magnitude at multiple elevations. Change in net area in a 7-day period in which a mass failure was observed resulted in up to a 40% decrease in sandbar area at both sites with mean area losses between 0 - 20% at 145-mile and mean losses in area at 30-mile ranging from 0 - 30% depending on the elevation examined. Mass failure magnitude is then larger on average at 30-mile and more frequent and these large magnitude changes in area primarily occur at elevations at the lower end of the FFZ. The magnitude of area changes at 145-mile is slightly smaller, but failures result in erosion over a wider range of elevations within the FFZ.

## **4.5. DISCUSSION**

### **4.5.1. Image-area and spatial variability among sandbars**

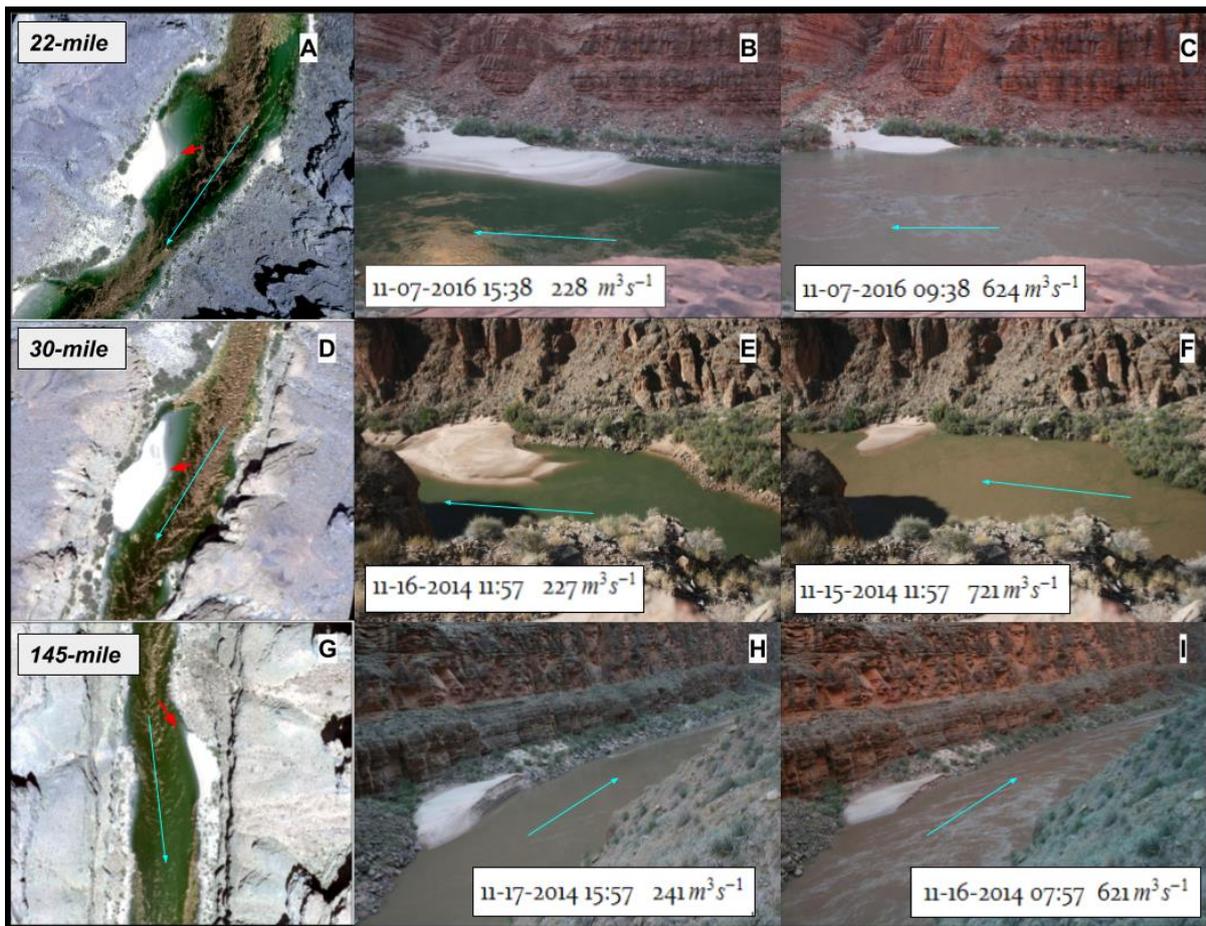
We quantified daily erosion and deposition rates by examining differences in image-area at the  $283m^3s^{-1}$  stage elevation. Errors associated with each image-derived area measurement ranged from 2.9% (145-mile) - 8.8% (22-mile) of site area, which is promising. However, these uncertainties are the same magnitude as expected area changes making it difficult to separate the signal from the noise. This is particularly true at 22-mile, which appears to have the lowest monthly and annual variability of the three sites, but the highest error in its daily image-derived area estimates. The lower accuracy for 22-mile is likely due to shifts in focal length (1-2 *mm*), which were recorded in the image metadata and occurred in 2017 and 2013. Calibration and collection of ground control points took place after the 2017

shift. Ideally, calibration data and ground control is collected each time the focal length changes, but this was not the case, and therefore error in the registration, rectification, and calibration likely increases at the 22-mile site going back in time.

In addition to the varying accuracies across sites, there are also different numbers of observations at each site, which may have biased these daily estimates. With more frequent observations at 30-mile, changes between images would undoubtedly be smaller. Also, there were fewer  $283m^3s^{-1}$  images available at the 145-mile because the minimum daily flows do not typically arrive at that site during daylight hours in July or January.

Monthly measures of area, since they are the average of several daily measurements, serve as more reliable estimates of sandbar area. Differences in the mean monthly area allow for estimates of net change in area and show the cyclic patterns of erosion and deposition. The monthly area time-series (Fig. 4.9) show that monthly net area change shifted from erosion to deposition at 30-mile at nearly twice the rate as 22-mile or 145-mile. The average change in area between subsequent months was twice as large at 30-mile as at 22-mile or 145-mile (Table 4.3). This is likely due, in part, to the increased frequency of mass failures at 30-mile, a lack of mass failures at 22-mile, and the lower magnitude of mass failures at lower elevations such as  $283m^3s^{-1}$  at 145-mile as evidenced in Figure 4.13.

Differences in the rate and magnitude of change at these three similarly classified sites (Mueller et al., 2018) are not surprising considering the differences in erosion and deposition rate and morphological evolution observed at two adjacent eddies in the same geomorphic grouping during the 2008 controlled flood (Wright and Kaplinski, 2011). Wright and Kaplinski (2011) suggested differences in site geometry caused by a downstream constriction at one site might be responsible for the discrepancy. Aerial imagery of the three sites illustrates that while all three are reattachment bars, there are major differences in the site characteristics.



**Figure 4.12** Imagery of 22-mile (A-C), 30-mile (D-F), and 145-mile (G-I). The first column of images (A,D,G) contains aerial imagery taken in 2013. The middle column (B,E,H) contains oblique imagery from a remote camera capturing the subaerial sandbar at the low end of the FFZ. The column on the right contains oblique imagery showing the subaerial sandbar near the high end of the FFZ. Red arrows are placed at the estimated reattachment point based on visual assessment of bar morphology. Blue arrows indicate the direction of downstream flow.

At the 145-mile site, a large portion of the subaerial bar is protected on both the upstream and downstream ends by rocks. The erodible portion of 145-mile within the FFZ is narrow and slopes steeply to the water which may cause it to aggrade and degrade more vertically than the other sites (Fig. 4.12G-I). The 22-mile sandbar has a higher percentage of its subaerial area within the FFZ and a conspicuous channel constriction mantling its downstream (Fig. 4.12A-C). Such constrictions might limit reattachment point migration and other forms of variability with regards to secondary flow structures during higher discharges (Wright and Kaplinski, 2011). The 30-mile sandbar has the largest extent of its area within

the fluctuating flow zone, sticking out further from the banks than the other two (Fig. 4.12D-F). The deposit is surrounded on three sides by water maximizing the area where erosion can occur. Its recirculation zone does not appear to be restricted like at 22-mile, thus it might have more variability in secondary flow structures at different discharges.

Correlations between monthly flow patterns and monthly area showed a significant and strong relationship between antecedent area and change in sandbar size from month to month (Fig. 4.9). Consistent with Mueller et al (2018), the significant and strongly negative relationship indicates that bars with large antecedent areas are primed to enter a phase of erosion, while small bars with ample accommodation space are primed to begin a phase of deposition. Meanwhile, correlations between other flow patterns varied between sites both in significance and direction. This suggests sandbar area change is tied more closely to antecedent sandbar size than to the flow metrics we examined. It also suggests that depending on the site-specific characteristics and antecedent conditions, the same discharges can have significantly variable effects on sandbar area at sandbars of the same geomorphic type, giving rise to the spatial variability observed in annual surveys. It may be useful to classify these sandbars into smaller sub-groupings based on average hypsometry to better predict sandbar behavior.

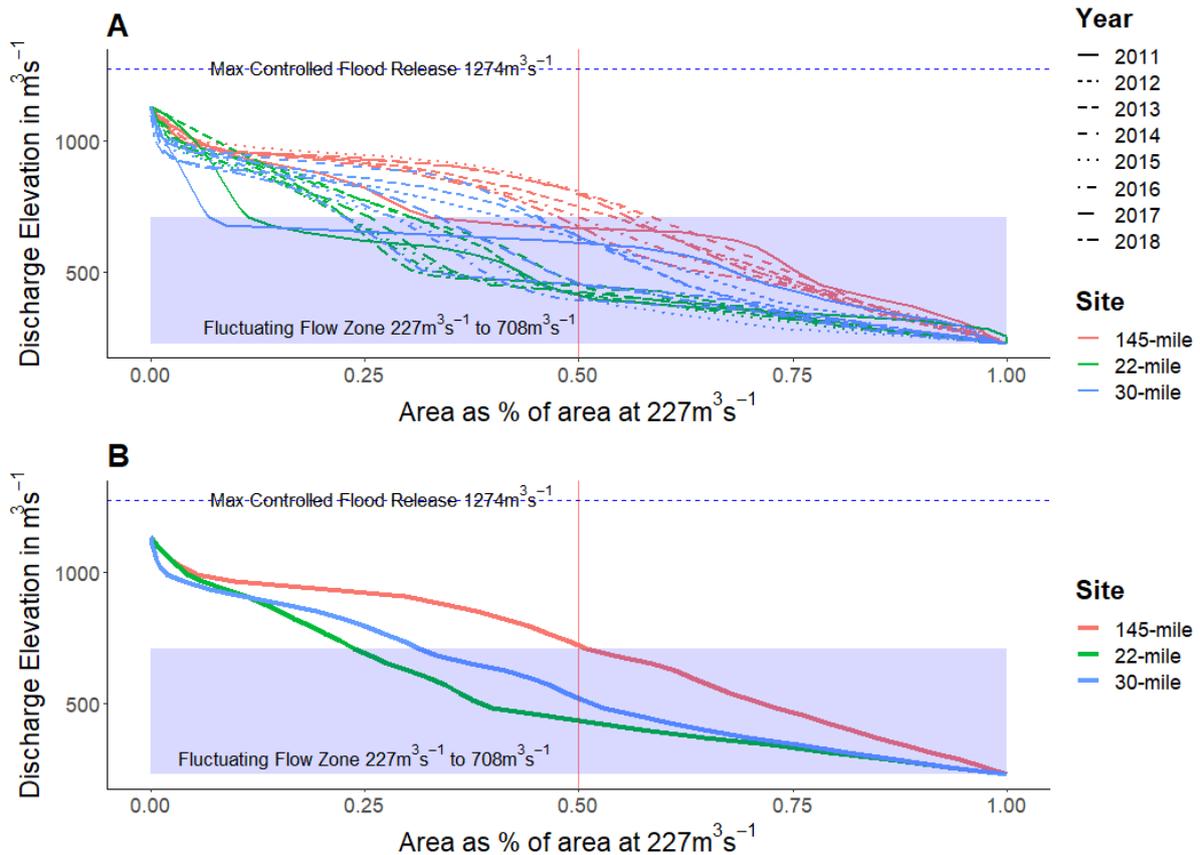
The response of these three sites to controlled floods also differs. Sandbar area is nearly twice as large at 30-mile during January and February in years following controlled floods than in years without controlled floods, while a measurable but more subdued response occurs at 22-mile and 145-mile (Fig. 4.7). However, across all sandbars, the effects of controlled floods on sandbar areas seem to disappear following the onset of high summer flows (Fig. 4.7 and 4.8). Caution must be taken before assuming that controlled flood responses at 22-mile and 145-mile are as similar as our results suggest. The increased steepness of 145-mile makes it more likely that controlled floods result in increased thickness

rather than area change at that site. Examining the trends in 30-mile's area after the 2014 and 2016 controlled floods shows that the peak of low-elevation area ( $227m^3s^{-1}$ ) occurs 4-6 months after the controlled flood. This change corresponds with decreases in the area at higher elevations ( $340m^3s^{-1}$  and  $396m^3s^{-1}$ ), suggesting the downslope movement of sand during normal dam operations.

#### **4.5.2. Mass failures at 22-mile, 30-mile, and 145-mile sandbars**

Mass failures appear to be significant erosional events in the cyclic behavior at 30-mile and 145-mile, but were not observed at 22-mile. The mean magnitude of area loss during mass failures at 30-mile ranged from 0.5% - 30% of the sandbar area at multiple elevations. This is up to 1.5 times larger than the average annual change in  $283m^3s^{-1}$  area (20.4%) and up to twice the average monthly change in  $283m^3s^{-1}$  area (16.4%) at 30-mile (Table 4.3). The magnitude of mass failures at 145-mile was similar in range but the means were between 0.5% - 20% of area at multiple elevations, though more high-elevation erosion seemed to result from failures. This increased tendency for higher elevation area loss at 145-mile might be due to the steepness of the narrow portion of the bar in the FFZ, providing an opportunity for mass failures to affect high-elevation sand deposited steeply over the failure area.

Alvarez and Schmeekle (2013) suggested that failure occurrence was primarily influenced by slope angle. Our results along with the hypsometric curves of these bars offer further evidence. Lima et al. (2019) (Chapter 2 of this dissertation) showed that hypsometric curves, which illustrate the normalized subaerial area by elevation (Fig. 4.13), could be used to examine the overall slope and morphology of sandbars in the Grand Canyon.



**Figure 4.13** Hypsometric curves showing the percent of  $227\text{m}^3\text{s}^{-1}$  area at elevations within the FFZ and CFZ derived from annual topographic surveys. (A) shows the hypsometric curves for each survey (2011 - 2018), and (B) shows the hypsometry averaged across all surveys (2011 - 2018). The blue box shows the FFZ. A dashed horizontal line shows the elevation of the maximum controlled flood  $1274\text{m}^3\text{s}^{-1}$ . A red vertical line indicates 50% of the sandbar  $227\text{m}^3\text{s}^{-1}$  area. The elevation where each curve crosses this line indicates the stage elevation where half or more of the subaerial sand is stored.

Based on annual surveys, the 145-mile sandbar is the steepest of the three sandbars on average and in each survey. The 22-mile sandbar is the least steep on average, with 30-mile in the middle. The lack of failures at 22-mile may be due in part to its reduced overall steepness and in part due to the limitations in flow variability caused by the downstream constriction. The 145-mile site, which has the steepest hypsometry, may have fewer mass failures than 30-mile because overall bar steepness is less important than the steepness of the portion of the bar within the FFZ or the steepness of the subaqueous bar. On average, over half of the subaerial extent of the 145-mile sandbar is above the FFZ (Fig 4.13B), while only around 28% of the 30-mile subaerial extent is above the FFZ. The high elevation area of 145-

mile above the FFZ is sheltered on both the upstream and downstream ends (Fig. 4.12G). We take this evidence to suggest that site-specific factors and resulting average bar morphology have a large impact on whether mass failures occur or not.

#### **4.5.3. Time-lapse mass failure analysis**

Mass failures were observed in time-lapses at 27 of 41 sites in our study (Table 4.4). The lack of observed mass failure events at 14 of the sites is likely due, in part, to the limited length of time-lapses, turbid water hiding sub-aqueous failures from detection, and more oblique angles preventing observation of sub-aqueous failures occurring at low elevations.

At sites where mass failures were observed, their mean recurrence intervals ranged from 74 days to 3,298 days with a mean recurrence interval for all sites of 976 days. This is considerably longer than the 371 day mean recurrence interval during the MLFF found in a study examining mass failure event frequency at seven sandbar sites from 1990 – 1993 (Cluer, 1995). The previous study identified failures at the end of the UDO and beginning of IF regimes by visually identifying them in photos. Of the three sites that are in both our study and the previous study (43-mile, 51-mile, and 172-mile), 51-mile has a shorter mean recurrence interval (684 days) during MLFF than in the previous study (1021 days). At the 43-mile site, mass failure events were documented with a recurrence interval of 177 days, while none were observed at that site during our study. The mean recurrence interval at 172-mile was 117 days in the previous study but 371 days in our analysis.

The discrepancy between recurrence intervals in the past and present could be due to differences in visual identification criteria, changes in camera viewing angle showing different portions of the bars in past studies versus the present study, or due to differences in bar behavior. The previous study also spanned just three years, making it difficult to accurately measure recurrence intervals. The current MLFF period started in 1996, and the current protocol for controlled floods began in 2011, giving the sandbars time to adjust to the

discharge patterns common in the MLFF. In contrast, the previous study was conducted as a transition was occurring between the UDO and IF. If sandbars had adjusted to the discharges in the UDO, the shift to IF could have altered the cyclic patterns of sandbars as they were adjusting to a new flow regime. Our results show that mass failure frequency increases with the onset of monthly changes in discharge (Fig. 4.10), most often occurring on the first day of a new month, when flows are adjusting. It is also important to consider that since controlled floods have been successful at increasing sediment storage within eddies, mass failures could be increasing in frequency due to increased sediment storage within eddies.

There are several potential mechanisms through which changes in flow might lead to conditions triggering mass failures. Previous studies have suggested that mass failures occur in response to high down ramp rates and coincided with weekend lows or seasonal shifts to lower water (Carruth and Cluer, 1991; Cluer, 1991; Budhu, 1995; Cluer, 1995). The breaching literature also describes mass failures in tidal environments coinciding with very low tides (Masterbergen et al., 2019). However, our analysis of mass failures during the MLFF found failures more often coincided with increases in discharge. Mass failure occurrences were clustered in the high-flow winter and summer months and on weekday water, which is higher on average than weekend water (Figs. 4.10). January had the highest number of mass failures, which may be because January is the first high flow month following November controlled floods, which result in steeper sandbar configurations. We also saw a clustering of some mass failures, where multiple failures at the same site happen in a short period of time. This may be due to a dramatic increase in instability following an initial failure or increased variability in secondary flow structures due to rapid changes in morphology.

We also found significant differences with regard to several flow metrics between days when failures occurred (Failure days) and our control groups ( $Lag_2$  days and Random

days) (Fig. 4.11). Variance in flows, captured by the CoV and daily standard delta metrics, were all significantly higher on failure days than on  $Lag_2$  days. We found that maximum discharge, increases in maximum discharge relative to the previous day, range in discharge, and increases in mean discharge were all significantly higher on days when failures occurred than on the control group days ( $lag_2$  and random).

The increases in discharge and subsequent mass failures are likely connected to changes in secondary flow structures at increased discharges. Discharge-induced changes, such as the appearance of secondary eddies, the formation of helical flow patterns, and adjustments in reattachment and separation point location observed during the 2008 controlled flood could be related to site-specific discharge thresholds, which vary based on site-geometry and antecedent sandbar conditions. The appearance of a secondary eddy coinciding with a mass failure has been observed at least once in the field during an increase in discharge from weekend to weekday water (Cluer, 1995). Perhaps at sites like 30-mile, significant changes in flow structure, leading to steeper deposition and more erosive forces within the eddy, can happen at the high end of fluctuating flow during high-flow months.

While the specific triggers that lead to mass failures remain elusive, we have established a connection between seasonal flow adjustments, increases in discharge, and discharge variability with mass failure occurrence throughout the Grand Canyon. Since mass failures are likely a significant mechanism for the transport of sediment out of long-term storage within the eddy, and because failures result in size change larger in magnitude than either the monthly or annual variability, they will remain a subject of management interest. Future research in this area would benefit from 3-D modeling at a wide range of discharges, repeat bathymetric surveys, and Acoustic Doppler current profiling during seasonal flow transitions.

Careful examination of time-lapses at all 41 sites adds context to the more detailed results at 22-mile, 30-mile, and 145-mile. Time-lapses elucidate patterns and processes that are often difficult to quantify (Brinley-Buckley et al., 2017). While there is significant variability in the responses of different sandbar sites to the current regime (MLFF) and to controlled floods, there does appear to be a generalized pattern at work. Controlled floods appear to be the most significant driver of morphological change at sites in years when they occur. Controlled floods most often increase deposit thickness, creating large convex-shaped sandbars though sometimes changing little with regards to aerial extent. Over the following months, gradual erosion and gravitational collapses of steepening banks begin to erode sand at the high end of the FFZ and in the CFZ, which is either carried downstream or redeposited at lower elevations in the FFZ or on the subaqueous sandbar forming a low-elevation platform. This process often creates a concave feature within the fluctuating flow zone, which retrogressively erodes bankward, often simultaneously as deposition is occurring on the low-elevation platform. Low-elevation aggradation continues until the sand fills the lateral accommodation space, which is also stage-dependent. Occasionally the channel-proximate edge of the subaqueous sandbars experiences mass failures, which often coincide with seasonal or monthly shifts in discharge occurring more often in high-flow months. Such failures open up accommodation space which is then gradually filled in by deposition, presumably with sand from higher elevations on the bar or sediment actively advecting into eddies.

In general, controlled floods seem to recharge eddies with sediment, while the cumulative effects of daily, weekly, monthly, and seasonal flow patterns in the MLFF rework that sediment creating bars, which are increasingly stable at specific flows but become destabilized when monthly flows shift. Thus, gradual erosion acts to mobilize sediment, conveying it primarily downslope, but likely also downstream. As sediment moves out of the

eddy and out of long-term storage, controlled floods are needed to periodically enrich eddies with new sediment (Grams et al., 2015; Hazel et al., *in press*).

Shifts from erosion to depositional phases do seem to coincide with seasonal or monthly changes in flow and with mass failure occurrence. Deposition appears to happen in both vertical and horizontal directions simultaneously and the sandbars exhibit continued growth until most of the horizontal accommodation space is filled. Without the ability to expand laterally and bounded by high-velocity mainstream current, deposition occurs vertically, creating steeper and steeper slopes. High-angle deposition may also be occurring more often in high-flow months since steeper deposition has been noted during controlled floods (Wright and Kaplinski, 2011). Such steep slopes prime the sandbar for a shift to an erosional phase that may be triggered by adjustments in seasonal flow or daily flow patterns. Spatial variability in bar behavior is likely related to both 1) site-specific hydraulic geometry, the site-specific discharge threshold for increased variability of secondary flow structures, and 2) antecedent conditions including the available accommodation space, sandbar slopes, and whether the site is experiencing an erosional or depositional phase.

#### **4.6. CONCLUSIONS**

We demonstrate that, despite its challenges, the enormous dataset of Grand Canyon sandbar imagery, which now exceeds a million images, can be utilized in a variety of ways to quantify sandbar response. First, imagery and image-areas can be used to examine spatial variability between sites. A detailed analysis of three reattachment bars revealed high spatial variability in the magnitude and frequency of cyclic behavior. It demonstrated that site-specific hydraulic geometry and antecedent conditions are likely more significant drivers of sandbar area change than discharge patterns, which might affect sandbars differently depending on those site-specific conditions. Secondly, image-derived monthly area estimates

track annual trends and reveal the frequency and magnitude of cyclic changes of eddy sandbars. However, errors in image-area and gage-wave time estimation make it challenging to quantify daily erosion or deposition rates with much certainty. Thirdly, image-derived time-lapses allow us to conservatively measure mass failure frequency and add to our understanding of complex processes that are difficult to quantify

Mass failures might be the most significant process for transporting fine sediment out of long-term storage in eddies. Our findings demonstrate that mass failures can result in erosion at magnitudes greater than average annual variability, most often corresponding with increases in daily, monthly, or seasonal discharge. Whether increases in discharge primarily lead to steeper deposition, which results in inevitable slope-induced mass failures, act as triggering mechanisms for failures on already unstable slopes, or both, is unclear. However, dam operators might consider making seasonal adjustments to make flow more gradual or consider equalizing differences between seasons to some degree to reduce the risk of mass failure events.

Information gathered from remote cameras can inform dam managers of sandbar behavior during the MLFF and contextualize that information with the detailed and more extended time-series. Findings like ours have the potential to influence dam operations and improve the management of downstream resources. Our results could be helpful in understanding other dam-affected and debris fan-affected canyons throughout the Western United States. The methods described here can also be applied to other large coastal or fluvial oblique image datasets to quantify sub-annual changes.

#### **4.7. ACKNOWLEDGEMENTS**

This study was supported by the U.S. Geological Survey's Grand Canyon Monitoring and Research Center. None of this would have been possible without data and guidance

provided by Joe Hazel, Matt Kaplinski, Daniel Buscombe, and Robert Tusso. Thanks to those who volunteered their time to manually check segmentation quality on over 13,000 images: Laura Lima, Charles Mogen, Hannah Kolus, Debbie Wells, and Rebecca Lima. Thanks to all the field technicians who assisted in collecting ground control points and camera rectifications during the annual sandbar monitoring trips, including Keith Kohl, Katie Chapman, Karen Kostner, Robert Ross, Bryan Cooperrider, Ryan Seumtewa, Megan Daubert, Amy Rust, and many more. A special thanks to Dr. Janrik Van den Berg for repeated correspondence discussing breach failures and hybrid flows.

**Disclaimer:** Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

#### 4.8. REFERENCES

- Alvarez, L. V., Schmeeckle, M.W., 2013. Erosion of river sandbars by diurnal stage fluctuations in the Colorado River in the Marble and Grand Canyons: full-scale laboratory experiments. *River Res. Appl.* 22, 1085–1095. <https://doi.org/10.1002/rra>
- Alvarez, L. V., Schmeeckle, M.W., Grams, P.E., 2017. A detached eddy simulation model for the study of lateral separation zones along a large canyon-bound river. *J. Geophys. Res. Earth Surf.* 122, 25–49. <https://doi.org/10.1002/2016JF003895>
- Annerino, J., 2000. *Canyons of the Southwest: A Tour of the Great Canyon Country from Colorado to Northern Mexico*. University of Arizona Press, Tucson, AZ.
- Andrews, E.D., Vincent, K.R., 2007. Sand Deposition in Shoreline Eddies Along Five Wild and Scenic Rivers, Idaho. *River Res. Appl.* 23, 7–20. <https://doi.org/10.1002/rra>
- Baker, D. B., Richards, R. P., Loftus, T. T., & Kramer, J. W. (2004). A new flashiness index: Characteristics and applications to Midwestern rivers and streams. *Journal of the American Water Resources Association*, 40(2), 503–522. <https://doi.org/10.1111/j.1752-1688.2004.tb01046.x>

- Baker, V.R., 1984. Flood Sedimentation in Bedrock Fluvial Systems, in: Koster, E.H., Steel, R.J. (Eds.), *Sedimentology of Gravels and Conglomerates*. Canadian Society of Petroleum Geologists, pp. 87–96.
- Bauer, B.O., Schmidt, J.C., 1993. Waves and Sandbar Erosion in the Grand Canyon: Applying Coastal Theory to a Fluvial System. *Ann. Assoc. Am. Geogr.*  
<https://doi.org/10.1111/j.1467-8306.1993.tb01946.x>
- Bertoldi, W., Piégay, H., Buffin-Bélanger, T., Graham, D., Rice, S., 2012. Applications of Close-range Imagery in River Research, in: Carbonneau, P., Piégay, H. (Eds.), *Fluvial Remote Sensing for Science and Management*. Wiley and Sons, Oxford, UK, pp. 341–366.
- Beus, S.S., Avery, C.C., 1992. The Influence of Variable Discharge Regimes on Colorado River Sand Bars Below Glen Canyon Dam. Flagstaff, AZ.
- Bevelhimer, M.S., Mcmanamay, R.A., O'Connor, B., 2014. Characterizing Sub-daily Flow Regimes: Implications of Hydrologic Resolution on Ecohydrology Studies. *River Res. Appl.* 30, 132–133. <https://doi.org/10.1002/rra>
- Bogle, R., Velasco, M., Vogel, J., 2013. An Automated Digital Imaging System for Environmental Monitoring Applications.
- Brandt, S.A., 2000. Classification of Geomorphological Effects Downstream of Dams. *Catena* 40, 375–401. [https://doi.org/10.1016/S0341-8162\(00\)00093-X](https://doi.org/10.1016/S0341-8162(00)00093-X)
- Briere, C., Elshoff, I., Morelissen, R., Santinelli, G., 2011. ARGUS imagery video technique for monitoring coasts. Ed. 2, Tanger, Maroc 25–28.  
<https://doi.org/10.5150/cmcm.2011.006>
- Brinley-Buckley, E.M., Allen, C.R., Forsberg, M., Farrell, M., Caven, A.J., 2017. Capturing change: The duality of time-lapse imagery to acquire data and depict ecological dynamics. *Ecol. Soc.* 22. <https://doi.org/10.5751/ES-09268-220330>
- Budhu, M., 1992. Mechanisms of erosion and a model to predict seepage-drive erosion due to transient flow., in: Beus, S.S., Avery, C.C. (Eds.), *The Influence of Variable Discharge Regimes on Colorado River Sand Bars Below Glen Canyon Dam*. Northern Arizona University, Flagstaff, AZ, pp. 12–42.

- Budhu, M., Gobin, R., 1994. Instability of Sandbars in Grand Canyon. *J. Hydrol. Eng.* 120, 919–933.
- Budhu, M., Gobin, R., 1995. Seepage-Induced Slope Failures on Sandbars in Grand-Canyon. *J. Geotech. Eng.*
- Buscombe, D., Ritchie, A., 2018. Landscape Classification with Deep Neural Networks. *Geosciences* 8, 244.
- Carbonneau, P.E., Piégay, H., 2012. Fluvial Remote Sensing for Science and Management. *Fluv. Remote Sens. Sci. Manag.* 1–440. <https://doi.org/10.1002/9781119940791>
- Carruth, R.L., M.C., C., Cluer, B.L., 1991. Documentation of Beach Failure Caused by Slumping and Seepage Erosion at two sites on the Colorado River in Grand Canyon, in: American Geophysical Union. Transactions Fall Meeting. American Geophysical Union, San Francisco, CA, p. 223.
- Cluer, B.L., 1991. Catastrophic Erosion Events and Rapid Deposition of Sandbars on the Colorado River, the Grand Canyon, Arizona, in: American Geophysical Union. Transactions Fall Meeting. American Geophysical Union, San Francisco, CA, p. 223.
- Cluer, B.L., 1995. Cyclic Fluvial Processes and Bias in Environmental Monitoring, Colorado River in Grand Canyon. *J. Geol.* 103, 411–421. <https://doi.org/10.1086/629777>
- Cluer, B.L., 1995. Cyclic Fluvial Processes and Bias in Environmental Monitoring, Colorado River in Grand Canyon. *J. Geol.* 103, 411–421. <https://doi.org/10.1086/629777>
- Converse, Y.K., Hawkins, C.P., Valdez, R.A., 1998. Habitat relationships of subadult humpback chub in the Colorado River through Grand Canyon: Spatial variability and implications of flow regulation. *Regul. Rivers-Research Manag.* 14, 267–284. [https://doi.org/Doi 10.1002/\(Sici\)1099-1646\(199805/06\)14:3<267::Aid-Rrr501>3.0.Co;2-5](https://doi.org/Doi%2010.1002/(Sici)1099-1646(199805/06)14:3<267::Aid-Rrr501>3.0.Co;2-5)
- de Koning, I.J., 1970. Neue Erkenntnisse beim Gewinnen und Transport von Sand im Spiilprojekt Venserpolder, VDI Tagung “Bauen im Ausland.” Hamburg.
- Dexter, L.R., Cluer, B.L., 1999. Cyclic erosional instability of sandbars along the Colorado River, Grand Canyon, Arizona. *Ann. Assoc. Am. Geogr.* 89, 238–266. [https://doi.org/Doi 10.1111/1467-8306.00144](https://doi.org/Doi%2010.1111/1467-8306.00144)

- Dolan, R., Howard, A., Gallenson, A., 1974. Man's impact on the Colorado River in the Grand Canyon. *Am. Sci.* 62, 392–401.
- Evangelidis, G.D., Psarakis, E.Z., 2008. Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization. *IEEE Trans. Pattern Anal. Mach. Intell.* 30. <https://doi.org/10.1109/TPAMI.2008.113>
- Grams, P.E., Schmidt, J.C., Wright, S.A., Topping, D.J., Melis, T.S., Rubin, D.M., 2015. Building sandbars in the Grand Canyon. *Eos (Washington, DC)*. 96. <https://doi.org/10.1029/2015EO030349>
- Grams, P.E., Tusso, R.B., Buscombe, D., 2018. Automated remote cameras for monitoring alluvial sandbars on the Colorado River in Grand Canyon, Arizona, Open-File Report. Reston, VA. <https://doi.org/10.3133/ofr20181019>
- Hazel, J.E., Grams, P.E., Schmidt, J.C., Kaplinski, M., 2010. Sandbar response following the 2008 high-flow experiment on the Colorado River in Marble and Grand Canyons. *U.S. Geol. Surv. Sci. Investig. Rep.* 2010-5015 52.
- Hazel, J.E., Kaplinski, M.A., Parnell, R.A., Manone, M.M., Dale, A., 1999. Topographic and Bathymetric Changes at Thirty-Three Long-Term Study Sites, in: Webb, R.H., Schmidt, J.C., Marzolf, G.R., Valdez, R.A. (Eds.), *The Controlled Flood in Grand Canyon*. American Geophysical Union, Washington, D.C., pp. 161–183.
- Hazel, J.E., Kaplinski, M., Parnell, R.A., Kohl, K., Schmidt, J.C., 2008. Monitoring fine-grained sediment in the Colorado River Ecosystem, Arizona-Control network and conventional survey techniques. *U.S. Geol. Surv. Open-File Rep.* 2008-1276 15.
- Hazel, J.E., Kaplinski, M., Grams, P., Hamill, D., Buscombe, D., Mueller, R., Ross, R.P., Jewell, S., *In press*. Multi-Decadal Sandbar Response to Flow Management Downstream from a Large Dam: the Colorado River in Marble and Grand Canyons, Arizona.
- Kasprak, A., Sankey, J.B., Buscombe, D., Caster, J., East, A.E., Grams, P.E., 2018. Quantifying and forecasting changes in the areal extent of river valley sediment in response to altered hydrology and land cover. *Prog. Phys. Geogr.* 42, 739–764. <https://doi.org/10.1177/0309133318795846>

- Kearsley, L.H., Schmidt, J.C., Warren, K.D., 1994. Effects of Glen Canyon dam on Colorado River sand deposits used as campsites in Grand Canyon National Park. *Regul. Rivers* 9, 137–149.
- Lane, S.N., 2000. The measurement of river channel morphology using digital photogrammetry. *Photogramm. Rec.* 16, 937–961. <https://doi.org/10.1111/0031-868X.00159>
- Marcus, W.A., Fonstad, M.A., 2010. Remote sensing of rivers: The emergence of a subdiscipline in the river sciences. *Earth Surf. Process. Landforms* 35, 1867–1872. <https://doi.org/10.1002/esp.2094>
- Mastbergen, D.R., Beinssen, K., Nédélec, Y., 2019. Watching the beach steadily disappearing: The evolution of understanding of retrogressive breach failures. *J. Mar. Sci. Eng.* 7. <https://doi.org/10.3390/jmse7100368>
- Matthes, G.H., 1947. Macroturbulence in natural streamflow. *Trans. - Am. Geophys. Union* 28, 255–262.
- Mueller, E.R., Grams, P.E., Schmidt, J.C., Hazel, J.E., Alexander, J.S., Kaplinski, M., 2014. The influence of controlled floods on fine sediment storage in debris fan-affected canyons of the Colorado River basin. *Geomorphology*. <https://doi.org/10.1016/j.geomorph.2014.07.029>
- Mueller, E.R., Grams, P.E., Hazel, J.E., Schmidt, J.C., 2018. Variability in eddy sandbar dynamics during two decades of controlled flooding of the Colorado River in the Grand Canyon. *Sediment. Geol.* 363, 181–199. <https://doi.org/10.1016/j.sedgeo.2017.11.007>
- National Research Council, 1996. *River Resource Management in the Grand Canyon*. The National Academies Press, Washington, D.C.
- National Research Council, 1999. *Downstream: Adaptive Management of Glen Canyon Dam and the Colorado River Ecosystem*. National Academy Press, Washington, D.C.
- National Park Service, 2006. *Grand Canyon: Colorado River Management Plan*. Grand Canyon, AZ.
- Nieto, M.A., Garau, B., Balle, S., Simarro, G., Zarruk, G.A., Ortiz, A., Tintoré, J., Álvarez-Ellacuría, Gómez-Pujol, L., Orfila, A., 2010. An open source, low cost video-based

- coastal monitoring system. *Earth Surf. Process. Landforms* 35, 1712–1719.  
<https://doi.org/10.1002/esp.2025>
- Nilsson, C., Reidy, C. a, Dynesius, M., Revenga, C., 2005. Fragmentation and Flow Regulation of the World’s Large River Systems. *Middle East* 308, 405–8.  
<https://doi.org/10.1126/science.1107887>
- Reddy, B.S., Chatterji, B.N., 1996. An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration. *IEEE Trans. IMAGE Process.* 5, 1266–1271.
- Rubin, D.M., McDonald, R.R., 1995. Nonperiodic eddy pulsations. *Water Resour. Res.* 31, 1595–1605.
- Sankey, J.B., Ralston, B.E., Grams, P.E., Schmidt, J.C., Cagney, L.E., 2015. Riparian vegetation, Colorado River, and climate: Five decades of spatiotemporal dynamics in the Grand Canyon with river regulation. *J. Geophys. Res. G Biogeosciences* 120, 1532–1547. <https://doi.org/10.1002/2015JG002991>
- Schmidt, J.C., 1990. Recirculating Flow and Sedimentation in the Colorado River in Grand Canyon, Arizona. *J. Geol.* 98, 709–724.
- Schmidt, J.C., Graf, J.B., 1990. Aggradation and Degradation of Alluvial Sand Deposits, 1965 to 1986, Colorado River, Grand Canyon National Park, Arizona. Washington, D.C
- Schmidt, J.C., Wilcock, P.R., 2008. Metrics for assessing the downstream effects of dams. *Water Resour. Res.* 44. <https://doi.org/10.1029/2006WR005092>
- Schmidt, J.C., Grams, P.E., 2011. The High Flows, in: Melis, T.S. (Ed.), *Effects of Three High-Flow Experiments on the Colorado River Ecosystem Downstream of Glen Canyon Dam, Arizona: Circular 1366*. U.S. Geological Survey, Reston, VA, pp. 25–50.
- Shen, C., 2018. A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resour. Res.* 54, 8558–8593.  
<https://doi.org/10.1029/2018WR022643>
- Sloff, K.C.J., Nieuwboer, B., Logan, B., Nelson, J.M., 2012. Effect of sediment entrainment and avalanching on modeling of Grand Canyon eddy bars 791–798.

- Sloff, K., Wright, S., Kaplinski, M., 2009. High-resolution three-dimensional modeling of river eddy sandbars, Grand Canyon, USA. *River, Coast. Estuar. Morphodynamics RCEM* 2009 883–890.
- Splinter, K.D., Harley, M.D., Turner, I.L., 2018. Remote sensing is changing our view of the coast: Insights from 40 years of monitoring at Narrabeen-Collaroy, Australia. *Remote Sens.* 10. <https://doi.org/10.3390/rs10111744>
- Tinkler, K., Wohl, E., 1998. A primer on bedrock channels, in: Tinkler, K., Wohl, E. (Eds.), *Rivers Over Rock: Fluvial Processes in Bedrock Channels*, Geophysical Monograph 107. American Geophysical Union, Washington, D.C., pp. 1–18.
- Topping, D.J., Rubin, D.M., Nelson, J.M., Iii, P.J.K., Corson, I.C., 2000. Colorado River sediment transport 2. Systematic bed-elevation and grain-size effects of sand supply limitation. *Water Resour. Res.* 36, 543–570.
- U.S.C. 102nd., 1992. Reclamation Projects Authorization and Adjustment Act of 1992, 102-575 et seq 1801-1809. (United States);  
<https://www.govinfo.gov/content/pkg/STATUTE-106/pdf/STATUTE-106-Pg4600.pdf>
- U.S. Department of the Interior, 1995. Operation of Glen Canyon Dam Colorado River Storage Project, Arizona: Final Environmental Impact Statement.
- Van den Berg, J.H., Van Gelder, A., Mastbergen, D.R., 2002. The importance of breaching as a mechanism of subaqueous slope failure in fine sand. *Sedimentology* 49, 81–95.  
<https://doi.org/10.1111/j.1525-139X.2006.00168.x-i1>
- Vincent, K.R., Andrews, E.D., 2008. Depositional settings of sand beaches along whitewater rivers. *River Res. Appl.* 24, 771–788. <https://doi.org/10.1002/rra>
- Valdez, R.A., Ryel, R.J., 1995. Life History and Ecology of the Humpback Chub in the Colorado River in Grand Canyon, Arizona. *Trans. Proc. Ser.*
- Walling, D.E., 2012. The role of dams in the global sediment budget, in: *Erosion and Sediment Yields in the Changing Environment: Proceedings of a Symposium Held at the Institute of Mountain Hazards and Environment 11-15 October*. IAHS, Chengdu, China, p. 356.

- Webb, R.H., Pringle, P.T., Rink, G.R., 1989. Debris flows from tributaries of the Colorado River, Grand Canyon National Park, Arizona. U.S. Geol. Surv. Prof. Pap. 1492 39.
- Wiele, S.M., Griffin, E.R., 1997. Modifications to a one-dimensional model of unsteady flow in the Colorado River through the Grand Canyon. Water-Resources Investig. Rep. 97-4046 7.
- Wiele, S.M., Andrews, E.D., Griffin, E.R., 1999. The effect of sand concentrations on depositional rate, magnitude, and location in the Colorado River below the Little Colorado River. *Control. Flood Gd. Canyon Geophysica*, 131–145.
- Williams, G.P., Wolman, M.G., 1984. Downstream effects of dams on alluvial rivers. U.S. Geol. Surv., Prof. Pap.; (United States); J. Vol. 1286.  
<https://doi.org/10.1126/science.277.5322.9j>
- Wright, S.A., Melis, T.S., Topping, D.J., Rubin, D.M., 2005. Influence of Glen Canyon Dam Operations on Downstream Sand Resources of the Colorado River in Grand Canyon, in: Gloss, S., Lovich, J., Melis, T. (Eds.), *The State of the Colorado River Ecosystem in Grand Canyon*. United States Geological Survey, Reston, Virginia.
- Wright, S.A., Kaplinski, M., 2011. Flow structures and sandbar dynamics in a canyon river during a controlled flood, Colorado River, Arizona. *J. Geophys. Res. Earth Surf.* 116, 1–15. <https://doi.org/10.1029/2009JF001442>
- Yamamuro, J.A., Lade, P. V., 1997. Static liquefaction of very loose sands. *Can. Geotech. J.* 34, 905–917. <https://doi.org/10.1139/t97-057>

## **CHAPTER 5: CONCLUSIONS AND MANAGEMENT IMPLICATIONS**

Ryan E. Lima

### **5.1. INTRODUCTION**

The remote camera network in Grand Canyon is one of the longest running and most extensive fluvial monitoring datasets in existence. The purpose of this study was to re-examine this under-utilized dataset to measure sub-annual sandbar morphodynamics and relate them to current dam operations. This chapter includes a summary of the key findings and contributions from each of the previous chapters, along with limitations and questions for future research. Finally, the implications of these findings for the adaptive management of Glen Canyon Dam are discussed.

### **5.2. RESEARCH QUESTIONS AND KEY FINDINGS**

While annual topographic surveys still provide the most reliable method for monitoring year-to-year and decadal trends in sandbar size, they provide little insight into the sub-annual topographic evolution of sandbars (Cluer, 1995; Hazel et al., *in press*). In Chapter 2, I attempted to determine if accurate estimates of sandbar volume could be quantified using oblique imagery. Results showed that careful selection of imagery over periods where a wide range of discharges were captured, such as the rising and falling limb of HFEs could be used to estimate volume at one site. These results provide evidence that sandbar volume may be estimated more frequently using pairs of image-derived areas at different elevations. If combined with the deep learning approach to image segmentation described in Chapter 3, such analyses could extend image-derived volume estimation to more sites at higher frequencies.

A fundamental question answered in Chapter 2 was whether imagery could be used in other ways to measure changes in elevational storage and measure the effectiveness of controlled floods. This research demonstrates that hypsometric curves can be a useful tool in this regard. Hypsometric curves provided a useful visual metric of overall sandbar morphology, allowed for monitoring the effects of HFEs and revealed patterns of elevational sediment distribution through time (Strahler, 1952; Jakobsson, 2002; Bahr et al., 2015, Karran et al., 2017).

Analyses in Chapter 2 were limited by my initial inexperience in programming for computer vision, the lack of reliable rectifications and camera calibrations, and lack of a method for rapid segmentation of imagery in the early days of this study. Monitoring efforts in the canyon would greatly benefit from more time and resources dedicated to image-derived volume estimation and hypsometric analysis. The next logical step in this line of inquiry, other than extending it to more sites, would be developing a method for extracting elevation relief ratio from image-derived hypsometric curves, which would provide a single metric useful in comparing sandbar morphology throughout the canyon (Pike et al., 1971). The analyses completed in Chapter 2 were critical to understanding the limitations of this dataset and exploring what information could feasibly be extracted from oblique remote imagery if segmentation methods could be improved.

Chapter 3 focused on improving our ability to rapidly segment imagery. The segmentation method employed in the second chapter was time consuming and infeasible for segmenting a dataset of over 1,000,000 images. Studies employing deep learning approaches to segmentation of remotely sensed imagery have shown great promise (Buscombe and Richie 2018; Carbonneau et al., 2020; Piégay et al., 2020). However, the majority of deep learning segmentation methodologies used in the earth sciences are designed for use with satellite or UAV-based imagery, while relatively few have been developed for use with

oblique, ground-based imagery applied to change detection in images with natural textures. The third chapter demonstrates a successful application of deep learning to the rapid binary semantic segmentation of sandbars in Grand Canyon remote camera imagery. This study contributes to the literature by providing an example of the steps needed to train, test, validate, and deploy a convolutional neural network to achieve sufficiently accurate ( $F1 = 0.9$ ) binary semantic segmentation in oblique ground-based imagery. Methods such as these can easily be adapted for use with similar fluvial and coastal monitoring datasets. The research in this third chapter evolved slowly and consumed a majority of the time and resources dedicated to this dissertation. My lack of programming expertise at the start of the study, led to innumerable dead ends. However, as my competency in programming increased, the work progressed with increasing speed and sophistication. With more time, computing power, and the availability of assistants to aid in image labelling the accuracy of sandbar segmentation models could be improved substantially.

Future attempts to improve sandbar segmentation in the canyon should compare the encoder-decoder CNNs that we employed with other types of fully-connected neural networks, which make use of transfer learning. Furthermore, steps should be taken to apply segmentation prior to registration and rectification which could increase accuracy further. The model developed in this study was tested on, at most, ten different sandbar sites and utilized training data from just five sites. The creation of a generalized sandbar segmentation tool would benefit from training sets including imagery from a larger sample of sandbar sites.

With improved methods for delineating sandbar area in remote camera imagery, several questions about sandbar dynamics could finally be examined in chapter 4. Results showed that area estimates with errors on the order of 2.9%-8.8% can be extracted from oblique imagery. Such errors make it difficult to quantify daily or monthly erosion rates, but daily area time-series and mean monthly area time-series are consistent with trends observed

in annual surveys and reveal nuanced sub-annual behavior. The variability in area measured at daily and monthly scales was the same order of magnitude as annual variability. Changes in net monthly area were strongly correlated with antecedent sandbar area at three dynamic reattachment bars. This study revealed that sandbar behavior is likely influenced more by site-specific characteristics and antecedent conditions than discharge patterns in the MLFF. For example, one of our sites experienced about twice as many cycles of erosion and deposition than the others. While differences in sandbar response to current dam operations and HFEs have been observed before in several studies, we are the first to quantify daily and monthly variability through nearly ten years of the MLFF (Hazel et al., 2010; Wright and Kaplinski, 2011; Hazel et al., *in press*). Our area analyses were limited to a few elevations and just three sites, but future work should extend to sites occupying different geomorphic classifications to capture spatial variability at sandbars throughout the canyon.

While uncertainty remains in what triggers mass failure events at Grand Canyon sandbars, the time-lapse analysis in Chapter 4 revealed that there is a significant connection between mass failures and increases in flow magnitude and variability that coincide with seasonal shifts in discharge. Results show that mass failures can result in area loss on the order of 20-40% of maximum subaerial sandbar area. Rates and magnitudes of mass failure varied among the three sites examined in detail. This variance is likely related to the influence of local controls on sandbar steepness and secondary flow structures.

### **5.3. MANAGEMENT IMPLICATIONS**

The impetus behind this dissertation attempts to improve our understanding of morphodynamic processes occurring at scales that are difficult to monitor in order to improve the operation of a dam. Dams are among the largest pieces of infrastructure on the planet.

Collectively dams hold back nearly 15% of total annual runoff globally disrupting the conveyance, and sequestering between 25 and 60Gt of sediment each year, resulting in a broad range of morphological and ecological responses (Williams and Wolman, 1984; Brandt, 2000; Nilsson et al., 2005; Walling, 2012). At the same time, dams provide critical services to society storing water, controlling floods, and generating electricity. Operations at Glen Canyon Dam are influenced by a program of adaptive management that seeks to mitigate impacts to, and improve the ecological and cultural values of, the downstream environment (Grand Canyon Protection Act 1992, Public Law 102-575; U.S. Department of the Interior, 1995). The effect of discharge patterns on sandbar morphodynamics is one consideration among many, but this dissertation contributes the following policy-relevant findings in that regard.

The sub-annual responses of sandbars to daily and monthly discharge patterns is complex and spatially variable even among sites in similar geomorphic settings, and therefore it may be useful to re-think sandbar classifications considering not only annual trends but sub-annual behavior or hypsometry. Monthly discharge patterns alone cannot explain rates of mean monthly area change or sandbar behavior. It would also be useful to conduct repeat topographic and bathymetric surveys at multiple sites during controlled floods to provide more insight into secondary flow structure variability across a range of discharges. This dissertation presents evidence that mass failure occurrence is linked to daily increases in discharge magnitude and variability that occurs as discharge shifts from low-flow to high-flow months and mass failures increase in frequency during higher weekly flows and in the first days of high flow months. Adjusting seasonal flow transitions to occur more gradually, over a week rather than a single day, could reduce the frequency of mass failures at sites where they occur, although more explicit testing of this hypothesis is needed. Remote

cameras such as those in the Grand Canyon have proved a vital tool in sandbar monitoring and show great potential for improving our understanding of sub-annual sandbar morphodynamics and answering outstanding questions in the field of fluvial geomorphology.

#### 5.4. REFERENCES

- Bahr, D.B., Pfeffer, W.T., Kaser, G., 2015. A review of volume-area scaling of glaciers. *Rev. Geophys.* 95–140. <https://doi.org/10.1002/2014RG000470>
- Brandt, S.A., 2000. Classification of geomorphological effects downstream of dams. *Catena* 40, 375–401. [https://doi.org/10.1016/S0341-8162\(00\)00093-X](https://doi.org/10.1016/S0341-8162(00)00093-X)
- Buscombe, D., Ritchie, A., 2018. Landscape Classification with Deep Neural Networks. *Geosciences* 8, 244.
- Carbonneau, P.E., Dugdale, S.J., Breckon, T.P., Dietrich, J.T., Fonstad, M.A., Miyamoto, H., Woodget, A.S., 2020. Adopting deep learning methods for airborne RGB fluvial scene classification. *Remote Sens. Environ.* 251, 112107. <https://doi.org/10.1016/j.rse.2020.112107>
- Cluer, B.L., 1995. Cyclic Fluvial Processes and Bias in Environmental Monitoring, Colorado River in Grand Canyon. *J. Geol.* 103, 411–421. <https://doi.org/10.1086/629777>
- Hazel Jr., J.E., Grams, P.E., Schmidt, J.C., Kaplinski, M., 2010. Sandbar response following the 2008 high-flow experiment on the Colorado River in Marble and Grand Canyons. *U.S. Geol. Surv. Sci. Investig. Rep.* 2010-5015 52.
- Hazel, B.J.E., Kaplinski, M., Grams, P., Hamill, D., Buscombe, D., Mueller, R., Ross, R.P., Jewell, S., *in press*. Multi-Decadal Sandbar Response to Flow Management Downstream from a Large Dam : the Colorado River in Marble and Grand Canyons, Arizona.
- Jakobsson, M., 2002. Hypsometry and volume of the Arctic Ocean and its constituent seas. *Geochemistry, Geophys. Geosystems* 3, 1–18. <https://doi.org/10.1029/2001gc000302>

- Karran, D.J., Westbrook, C.J., Wheaton, J.M., Johnston, C.A., Bedard-Haughn, A., 2017. Rapid surface-water volume estimations in beaver ponds. *Hydrol. Earth Syst. Sci.* 21, 1039–1050. <https://doi.org/10.5194/hess-21-1039-2017>
- Nilsson, C., Reidy, C. a, Dynesius, M., Revenga, C., 2005. Fragmentation and Flow Regulation of the World's Large River Systems. *Middle East* 308, 405–8. <https://doi.org/10.1126/science.1107887>
- Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carbonneau, P., Dufour, S., Liébault, F., Ruiz-Villanueva, V., Slater, L., 2020. Remotely sensed rivers in the Anthropocene: state of the art and prospects. *Earth Surf. Process. Landforms* 45, 157–188. <https://doi.org/10.1002/esp.4787>
- Pike, R.J., Wilson, S.E., 1971. Elevation-relief ratio, hypsometric integral, and geomorphic area-altitude analysis. *Bull. Geol. Soc. Am.* 82, 1079–1084. [https://doi.org/10.1130/0016-7606\(1971\)82\[1079:ERHIAG\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1971)82[1079:ERHIAG]2.0.CO;2)
- Strahler, A.N., 1952. Hypsometric analysis of erosional topography. *Bull. Geol. Soc. Am.* 63, 1117–1142. [https://doi.org/10.1130/0016-7606\(1952\)63\[1117:HAAOET\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1952)63[1117:HAAOET]2.0.CO;2)
- Walling, D.E., 2012. The role of dams in the global sediment budget, in: *Erosion and Sediment Yields in the Changing Environment: Proceedings of a Symposium Held at the Institute of Mountain Hazards and Environment 11-15 October*. IAHS, Chengdu, China, p. 356.
- Williams, G.P., Wolman, M.G., 1984. Downstream effects of dams on alluvial rivers. *U.S. Geol. Surv., Prof. Pap.*; (United States); J. Vol. 1286. <https://doi.org/10.1126/science.277.5322.9j>
- Wright, S.A., Kaplinski, M., 2011. Flow structures and sandbar dynamics in a canyon river during a controlled flood, Colorado River, Arizona. *J. Geophys. Res. Earth Surf.* 116, 1–15. <https://doi.org/10.1029/2009JF001442>
- U.S.C. 102nd., 1992. Reclamation Projects Authorization and Adjustment Act of 1992, 102-575 et seq 1801-1809. (United States); <https://www.govinfo.gov/content/pkg/STATUTE-106/pdf/STATUTE-106-Pg4600.pdf>

U.S. Department of the Interior, 1995. Operation of Glen Canyon Dam Colorado River Storage Project, Arizona: Final Environmental Impact Statement.

## **APPENDIX**

### **UNITS AND ACRONYMS**

#### **ACRONYMS**

ANN	Artificial neural network
CFZ	Controlled flood zone
CNN	Convolutional neural network
DEM	Digital elevation model
DL	Deep learning
EXIF	Exchangeable image file format
FCN	Fully-convolutional network
FFZ	Fluctuating flow zone
FFT	fast Fourier transform
GCD	Glen Canyon Dam
GCP	Ground control point
GPU	Graphical processing unit
HIZ	Hydrologically inactive zone
IF	Interim flows
MLFF	Modified low fluctuating flows
NF	Naturalized flows
PC	Personal computer
RM	River mile

ROD

Record of Decision

UAV

Unmanned aerial vehicle

## UNIT CONVERSIONS

<b>Multiply</b>	<b>By</b>	<b>To obtain</b>	<b>Quantifies</b>
cubic feet per second ( $ft^3s^{-1}$ )	0.028316846592	cubic meters per second ( $ft^3s^{-1}$ )	Volumetric flow rate
acre-foot ( <b>af</b> )	1233.48	cubic meters ( $m^3$ )	Volume
acre-foot ( <b>af</b> )	1,000,000	million acre-feet ( <b>maf</b> )	Volume
ton( <i>t</i> ) U.S.	1102311310.9244	gigaton( <i>Gt</i> )	Mass
acre	0.000247105	square meter ( $m^2$ )	Area
mile ( <i>mi</i> )	1.609	kilometer ( <i>km</i> )	Distance

## DATA AVAILABILITY AND ASSOCIATED CODE

Lima, Ryan E. (2021, July 29). ESTIMATING CHANGES IN FINE-SEDIMENT STORAGE AT EDDY-SANDBARS ON THE COLORADO RIVER, GRAND CANYON, AZ USING OBLIQUE IMAGERY FROM REMOTE CAMERAS (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.5146763>

DOI [10.5281/zenodo.5146763](https://doi.org/10.5281/zenodo.5146763)