Big Changes in the Big Bend: A data management and shoreline analysis study

By

MELISSA MARIE MORENO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2021

© 2021 Melissa Marie Moreno

To my friends and family that have been there for me and to Michelle Masferrer for inspiring me to make choices that fuel me.

ACKNOWLEDGMENTS

I thank the University of Florida Academic Research and Consulting Services for teaching and encouraging me to learn more in the field of data science. I also thank National Fish and Wildlife Foundation for funding the Lone Cabbage Reef project. I also like to thank all the technicians, research coordinators, and administrators on the Lone Cabbage Reef project that collected, observed, and gathered data. I would also like to thank my committee members and any faculty, staff or student that has helped me.

TABLE OF CONTENTS

page

[ACKNOWLEDGMENTS 4](#_Toc64018306)

[LIST OF TABLES 7](#_Toc64018307)

[LIST OF FIGURES 8](#_Toc64018308)

[LIST OF ABBREVIATIONS 11](#_Toc64018309)

[ABSTRACT 12](#_Toc64018310)

CHAPTER

1 [ESTABLISHING A PROGRESSIVE DATA MANAGEMENT WORKFLOW FOR BIOLOGICAL DATA TO INFORM ADAPTIVE MANAGEMENT DECISIONS 14](#_Toc64018311)

[Introduction 14](#_Toc64018312)

[Terminology 16](#_Toc64018313)

[“Living data” 16](#_Toc64018314)

[Adaptive management 17](#_Toc64018315)

[Version control 17](#_Toc64018316)

[Data Types 18](#_Toc64018317)

[The LCR Project Data Types 18](#_Toc64018318)

[LCR project naming conventions 19](#_Toc64018319)

[Water quality data from autonomous sensors 20](#_Toc64018320)

[Oyster counts and measurements from field sampling by people 20](#_Toc64018321)

[Water quality measured by field-crews 21](#_Toc64018322)

[Establishing a Modern Data Workflow 21](#_Toc64018323)

[Water Quality Workflow 22](#_Toc64018324)

[Adding Water Quality Measurements to our Permanent MySQL Relational Database and Version Control 23](#_Toc64018325)

[Automated Data Checks Through Python and R Scripts 24](#_Toc64018326)

[Oyster Observation Workflow 24](#_Toc64018327)

[Datasheets, Data Entry and Validation of Oyster Data from the Field 25](#_Toc64018328)

[Adding Oyster Observations to a Central Storage and Version Control 26](#_Toc64018329)

[Regularly Updated Data and Adaptive Management 26](#_Toc64018330)

[Discussion 27](#_Toc64018331)

2 [TAILORING GITHUB FOR ECOLOGY 36](#_Toc64018332)

[Introduction 36](#_Toc64018333)

[Terminology overview 37](#_Toc64018334)

[Github and Version and Control 38](#_Toc64018335)

[Challenges Working in one Repository 38](#_Toc64018336)

[Creating a New Github Repository Structure and Workflow 39](#_Toc64018337)

[Github Permissions and Branch Workflow 40](#_Toc64018338)

[Naming Conventions for Repository, Files, and Folders 41](#_Toc64018339)

[Transparency 42](#_Toc64018340)

[Discussion 42](#_Toc64018341)

3 [CASE STUDY: DEER ISLAND, FLORIDA TIME PERIOD SHORELINE ANALYSIS USING DSAS 48](#_Toc64018342)

[Introduction and Background 48](#_Toc64018343)

[Climate Change and SLR 48](#_Toc64018344)

[Characteristics of Sandy Shorelines and Sedimentation 49](#_Toc64018345)

[Suwannee River Sedimentation and Discharge 49](#_Toc64018346)

[Human Development and Impacts 50](#_Toc64018347)

[Big Bend Habitats for Species Richness 51](#_Toc64018348)

[Major Hurricanes Near the Study Area 51](#_Toc64018349)

[Reason for Effort 52](#_Toc64018350)

[Materials and Methods 53](#_Toc64018351)

[Study Area 53](#_Toc64018352)

[Imagery Selection Process 54](#_Toc64018353)

[Digital Shoreline Analysis System (DSAS) 56](#_Toc64018354)

[DSAS Parameters and Selections 57](#_Toc64018355)

[Results 58](#_Toc64018356)

[Shoreline Analysis for Years 1994-2007 59](#_Toc64018357)

[Shoreline Analysis for Years 2010-2019 59](#_Toc64018358)

[Shoreline Analysis for Years 1994-2019 60](#_Toc64018359)

[Shoreline Analysis for 10 and 20-Year Prediction 61](#_Toc64018360)

[Imagery Observation 62](#_Toc64018361)

[Discussion 62](#_Toc64018362)

4 [Synthesis 82](#_Toc64018363)

[LIST OF REFERENCES 83](#_Toc64018364)

[BIOGRAPHICAL SKETCH 89](#_Toc64018365)

LIST OF TABLES

Table page

[2-1. Table of naming conventions for file types, example, and description of the example 47](#_Toc64018366)

[3-1. Table of metadata for each aerial image used 69](#_Toc64018367)

[3-2. National Agriculture Imagery Program (NAIP) aerial imagery band wavelength ranges in units (µm) 71](#_Toc64018368)

[3-3. Summary statistics calculated by DSAS, Distance: NSM (Net Shoreline Movement) 79](#_Toc64018370)

[3-4. Summary statistics calculated by DSAS, RATE: LRR (Linear Regression Rate) 79](#_Toc64018371)

LIST OF FIGURES

Figure page

[1-1. The adaptive management process for ecological restoration projects. 31](#_Toc64018171)

[1-2. Generalized version control workflow for the LCR project 31](#_Toc64018172)

[1-3. Water quality location map. 32](#_Toc64018174)

[1-4. Data workflow for water quality observations. 33](#_Toc64018175)

[1-5. Diagram of how the tables in our MySQL relational database 34](#_Toc64018176)

[1-6. Data workflow for oyster measurements. 35](#_Toc64018178)

[2-1. Visualization of our main project repository structure and various projects in the same repository.. 45](#_Toc64018180)

[2-2. New LCR project workflow which describes how project repositories update to work with newly added data.. 46](#_Toc64018182)

[3-1. A) Map of Florida with Dixie, Franklin and Taylor counties identified along side the Suwannee River 67](#_Toc64018183)

[3-2. Island degradation of Derrick Key in the Cedar Keys, Florida from 1982 (left) to 2016 (right). 68](#_Toc64018185)

[3-3. Location of Deer Island, Florida. 68](#_Toc64018186)

[3-4. Example of DSAS transect casting 71](#_Toc64018187)

[3-5. DSAS components and operational workflow 72](#_Toc64018188)

[Figure 3-6. Results of DSAS calculations for years 1994 to 2007 73](#_Toc64018189)

[Figure 3-7. Figure of the DSAS statistics for years 1994 to 2007 74](#_Toc64018190)

[Figure 3-8. Results of DSAS calculations for years 2010 to 2019. 75](#_Toc64018191)

[Figure 3-9. Figure of the DSAS statistics for years 2010 to 2019 76](#_Toc64018192)

[Figure 3-10. Results of DSAS calculations for years 1994 to 2019. 77](#_Toc64018193)

[Figure 3-11. Figure of the DSAS statistics for years 1994 to 2019 78](#_Toc64018194)

[Figure 3-12. DSAS shoreline prediction forecast.. 80](#_Toc64018195)

[Figure 3-13. NAIP aerial imagery used for analysis. A) Imagery from 1994 (scale 1:3,000); B) Imagery from 2019 (scale 1:3,000) 81](#_Toc64018196)

LIST OF ABBREVIATIONS

|  |  |
| --- | --- |
| ARCS | University of Florida Academic Research Consulting & Services, http://arcs.uflib.ufl.edu/ |
| CIR/CNIR | Red, green, blue, and infrared four band satellite imagery. |
| DOQQ | Digital Ortho Quarter Quad Tiles |
| DSAS | Digital Shoreline Analysis Systems, created and maintained by USGS for analyzing shorelines. |
| GIS | Geographic Information System, used to gather manage and analyze geographic data. |
| LCR | Lone Cabbage Reef |
| LRR | Linear Regression Rate calculation available in DSAS |
| NAIP | National Agriculture Imagery Program |
| NDVI | Normalized Difference Vegetation Index |
| NEON | National Ecological Observing Network |
| NFWF | National Fish and Wildlife Foundation |
| NSM | Net Shoreline Movement calculation available in DSAS |
| RGB | Red, green, and blue three band satellite imagery |
| SECOORA | Southeast Coastal Ocean Observing Regional Association |
| USGS | United States Geological Survey |
| QA/QC | Quality assurance and quality control in regard to managing gathering, collecting, and reporting data. |
| YSI | Yellow Springs Instrument (manufacturer) of devices that can measure water quality observations |

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Master of Science

BIG CHANGES IN THE BIG BEND: A DATA MANAGEMENT,  
AND SHORELINE ANALYSIS STUDY

By

Melissa Marie Moreno

May 2021

Chair: Pine III, William E

Member: Mossa, Joann, & Adams, Peter N

Special Member: Aufmuth, Joseph L

Major: Interdisciplinary Ecology (ECL) - (IEC\_MS)

A major role played by natural resource management agencies at the state and federal level including Florida Fish and Wildlife Service and US Fish and Wildlife Service is to protect fish, wildlife, and their habitats for the benefit and enjoyment of people. To do this, these agencies have to make decisions such as what lands to acquire for protection or whether to restore a resource that is declining. In the best case these decisions are guided by data that are collected from observations in the field, made available to the resource managers to analyze and interpret, and then used to inform the decision making.

Advances in technology such as expanded remote sensing and animal tracking platforms have triggered rapid expansion of data available for ecologists and natural resource scientists. We have customized a modern data workflow for continuous and discrete long-term ecological data to assist in adaptive decision making. To promote reproducibility in our workflows and reduce data collection errors, we incorporated specific standards into our program including (1) standardizing field datasheets linked to an electronic data entry platform; (2) performing quality assurance and control (QA/QC); (3) creating scripts to analyze data and inform decision making; and (4) use a version control workflow to track changes to data, scripts and documents.

Making these data and scripts available to multiple users is challenging. I developed a repository structure using GitHub so that files and data may be available to all members of the team, other collaborators, and publishers. I do this by 1) evaluating our previous repository structure and workflow; 2) creating a new and consistent structure and workflow among all project repositories; 3) and establishing and maintaining a file naming convention which encompasses any file that could be in a repository. This repository structure takes into account the need to manage “living data” and considers the need for ecological efforts to be transparent.

Other types of data useful for making decision include information from imagery captured from different platforms over time that can provide evidence of changes in landscape features. As an example, climate change perpetuation and sea level rise have led to concerns by resource managers and the public related to shoreline dynamics such as changes in water levels and erosion of shoreline features along the Gulf of Mexico. Shoreline dynamics in areas with high human coastal development have been widely studied because of the economic costs to these properties from changes in sea level. However in areas with low human population density, changes in shoreline features are less studied because these areas are either assumed to not be undergoing change or because their economic value is lower. In this study we used seven NAIP (National Agriculture Imagery Program) aerial images, from 1994 to 2019, of our study area near Cedar Key, FL. The cloud-free images were collected during relatively similar mean river discharge levels and during (mostly) the same season. We assessed the shoreline changes using ESRI’s ArcMap© spatial data analysis extension DSAS (Digital Shoreline Analysis Systems) on three different time periods in from the imagery, 1994-2007, 2010-2019 and 1994-2019. From this analysis we determined the greatest areas of impact and possibly speculate on possible factors that may contribute to an escalated shoreline change rate during a selected time frame. The data acquisition and workflow for the DSAS is documented following a similar structure as earlier chapters to provide a framework for repeating these analyses with additional imagery in the future.

CHAPTER 1

ESTABLISHING A PROGRESSIVE DATA MANAGEMENT WORKFLOW FOR BIOLOGICAL DATA TO INFORM ADAPTIVE MANAGEMENT DECISIONS

Introduction

Traditional field biology programs, many of which are designed to monitor animal populations and their environments, have experienced a substantial evolution in data collection, management, and storage technology in recent years. Changes include new sensor technology, data collection methods, and data observing platforms that are being used in large-scale monitoring programs including SECOORA (Southeast Coastal Ocean Observing Regional Association) and NEON (National Ecological Observing Network). As an example, advancements in sensor technology have allowed for significant changes in water quality monitoring such as transitioning from discrete single location and single point in time sample collections to real-time continuous observations at multiple locations (Martinelli et al., 2016). While the scale and technological capacity of many monitoring programs has increased, these monitoring programs are still most often conceived, planned, and used by personnel trained as biologists and not data scientists. The lack of training in basic data management, curation, and workflow of data generated from these types of data collection platforms was demonstrated in a recent NSF (National Science Foundation) survey (Lowndes et al., 2017)which highlighted that of the 704 scientists who participated in the survey, “data skills” (e.g. multi-step workflows, ability to store, share and publish data) was identified as the largest unmet need (Barone et al., 2017).

The US Gulf of Mexico region is undergoing a large restoration effort to reverse observed declines in key ecosystem components including seagrass, fish communities, and oyster reefs using funding from the consolidated Deepwater Horizon settlements (see https://www.nfwf.org/gulf/Pages/home.aspx as an example). These restoration projects vary in spatial scale, but, like other restoration efforts, these projects have data collection and evaluation efforts that occur frequently throughout the project. Several of the restoration programs in this funding program require basic adaptive management concepts be used to guide restoration actions (Zedler, 2017). Under this framework, decisions related to restoration actions are made iteratively based on stating, testing, and updating hypotheses based on observed outcomes (Figure 1-1). In a restoration context, this information can be used to inform the restoration actions such type of substrate to use in an oyster restoration project or monitoring program design as the project is ongoing, increasing efficiency by maximizing return on investment from restoration dollars. Doing so requires a data management plan designed to improve restoration actions by maximizing learning from previous and ongoing restoration efforts (Tompkins & Adger, 2004).

One example restoration effort funded by NFWF (National Fish and Wildlife Foundation) as part of the consolidated Deepwater Horizon settlements is the Lone Cabbage Reef (LCR) oyster reef project in the northeastern Gulf of Mexico. The primary goal of this project is to restore specific oyster reefs to historic levels so that they may be resilient to changing sea level and river discharge. This project generates data from multiple sources including continuous autonomous water quality data via sensors and observations of oyster populations by field biologists. These data are generated at different time steps with sensor data obtained at hourly time intervals from multiple spatial locations and biological data collected at discrete time intervals from multiple spatial locations. For both cases, there is a need to capture and process data to meet standards and then complete routine analyses of these data to ensure they are useful for informing project objectives and questions. This is critical because this project by design uses adaptive management principles to inform the restoration through an interactive process of collecting data, analyzing these data, and informing restoration actions from these analyses. For this project to efficiently operate in an adaptive management framework, we developed a system which captures data as it is collected, guides the data to analyses, documents data and analyses decisions via version control, and archives and makes these data available for long-term reproducible exploration. Here we describe this data management system and the structure and decisions made in implementing the system to improve data quality and reduce the likelihood of data collection and errors in analyses.

Terminology

“Living data”

Living data” are defined as data which are continuously collected and updated (Yenni et al., 2018). These types of data are critical to adaptive learning to inform restoration and management actions. Examples of learning as part of a restoration project includes small changes like shifting the location of an autonomous sensor, to larger changes such as revamping of sampling programs because of low statistical power. Living data can inform these decisions, but living data are challenging to work with from a data management perspective because the data (by design) change as new data are collected. In a restoration context as these data are collected, they must be processed, and analyses of these data to be completed to help draw inferences on how the system of interest is responding to the restoration action. This idea of iteratively integrating new data, analyses, and comparing these outcomes with previously stated objectives is not new and is a central aspect of the adaptive management process for natural resources as originally described (Holling 1978; Walters 1986).

Adaptive management

By design, an adaptive management system requires rapid feedback loops between data collection, analyses, and interpretation to drive the process of updating knowledge, examining management and restoration options, making decisions and implementing actions (Nie & Schultz, 2012). This process is repeated (Figure 1-1) to improve management actions such as identifying the best restoration approach. Data used must meet quality assurance/quality control (QA/QC) protocols to identify and correct inconsistencies and errors in field or sensor observations before these data are used in an analysis. Errors in these data, or delays in producing the data in a framework useable for analysis, can quickly lead to a breakdown in the adaptive learning process either in terms of slowing the analyses and limiting their utility for timely decision making, or worse, erroneously informing the decision-making process because of errors in data management or analyses.

Version control

Version control software is a type of software that helps to manage documents, scripts of computer code, and other developmental information documents that are shared and iteratively updated over time in an ecological project (Figure 1-2**)**. The key purpose of using version control software is to document and confirm that changes in content are intended and planned. The advantages of using version control (1) a version control system saves all versions of a file, (2) version control records who made what changes to specific files and makes the user write detailed notes about what they changed (3) allows these changes to be undone if needed, (4) can facilitate reproducibility and transparency of project code and decision making (Ram, 2013). Version control can be incorporated into a data workflow using software such as Github (https://github.com).

Data Types

The LCR project collects data on multiple parameters to measure ecosystem response to oyster reef restoration. One response metric are observations of water conductivity and temperature collected hourly from autonomous sensors. These types of data are measured and recorded by the sensor and are output in a standard format than can be interpreted for analyses directly by a computer. A second metric are counts of oysters at locations where restoration has been done (restored sites) and sites where restoration has not been done (wild oyster bars). Oyster counts are made by people conducting the fieldwork during winter low tide events. These data are collected by people and then must be entered into a computer as a standard data form before these data can be analyzed. We created a data management workflow to efficiently process and analyze data from both of these data streams. These data streams are then consolidated, and used to actively inform decision-making for the project such as the amount of sampling trips needed to optimize oyster density estimates. We use software and tools that are open source, widely available and familiar to many field biologists such as program R (Lefcheck, 2016) and Microsoft Excel. This chapter documents this workflow and provides an example for use in other restoration and conservation projects.

The LCR Project Data Types

The LCR project generates data from autonomous sensors, and human observed counts/measurements. Several types of data are collected at various frequencies (seasonally, bi-monthly) and each data type requires a specific plan to monitor data workflow from collection to analyses.

LCR project naming conventions

A critical component of our data management plan was to create a naming convention standard for every sampling location. Every oyster transect sampling location is identified by its location, which we abbreviate (e.g., Lone Cabbage= LC, Horseshoe Beach = HB), and then a number added to the end to identify the location as separate from other sampling locations. Each sampling location is then recorded in a master sampling list to ensure that we can go back to the same location and to also reference the same location in the exact way throughout multiple sampling events. Even though the oyster transect sampling locations are randomized when selected, it is an integral part of our sampling workflow to establish the naming convention of the sampling location prior to the sampling event.

Another aspect of our naming convention standards, which directly relates to data management, are the way we name project files. We use a standard of referencing the date the file was created and what the file is so that every project member will be able to decipher the subject matter of the file without having to view its content. For example, our water quality sensor download files are named in a specific format YYYYMMDD\_sitelocation\_sensortype.file (e.g., 20200902\_wq7\_diver.MON). The file name is important to identify the date and site in a clear fashion especially when we are investigating sensor readings which may be corrupt or uncalibrated. This file naming format has saved time for project team members because all files are uniform and consistent in their naming, making it easier for each team member to follow the naming convention guidelines.

Water quality data from autonomous sensors

We collect hourly water quality observations from 11 different sites around Lone Cabbage reef (Figure 1-3). These observations are downloaded from autonomous sensors approximately every two weeks. Maintenance of these sensors and their protective housing are completed to ensure continuous stream of data by reducing data errors due to biofouling or equipment loss. These “living data” have the highest frequency of occurrence (most number of observations) and require strict data management protocols (Box 3) both in terms of launching and maintaining the sensors and in importing the data files to maintain database integrity.

Oyster counts and measurements from field sampling by people

The lowest lunar tides of the year in the area around Lone Cabbage Reef occur during winter, and these low-tide events de-water oyster reefs which allow teams of people to count and measure oysters to document status and trends of oyster populations (Moore et al. 2020). These count data are recorded in the field on datasheets and then entered into a computer by people through a dual data entry system where each data record is entered independently into the computer. The dual entry system also includes data validation drop downs that ensure that the user is only entering appropriate data. For example, if a user is entering the height of an oyster to be 1000 mm the data validation will reject this entry because there is a range of acceptable oyster heights that it will allow and 1000 mm falls outside this range. Any mismatch or errors in the dual data entry worksheets will then be reconciled by the research coordinator. This double data entry system was created to reduce the chance of data entry errors and human introduced errors (Box 4 3A).

Water quality measured by field-crews

During water quality service trips, we also collect water quality measurements using a hand-held YSI (Yellow Springs Instrument) device to provide a supplemental check on our autonomous sensor observations. These measurements are recorded once during the water quality service trip for each site location. These observations are the least intensive data type as their frequency is low, and they are manually entered in the MySQL database (Box 3).

While conceptually each of these data types appear to be similar, because of differences in the frequency the data are collected and the collection method (with a machine or by hand), each data stream must be managed differently. Addressing the variety of concerns which have been discovered through regularly updating these data types may also address many data management challenges which researchers may confront.

Establishing a Modern Data Workflow

Data collected in the field are transferred and stored in a relational database for QA/QC and analyses. Database development efforts for this project started prior to data collection through development of database “blueprints” via white board exercises to clarify (1) database goals, (2) data types and data sources, and (3) relationships among data types within the database. Blueprinting development efforts were led by University of Florida Academic Research Consulting & Services (ARCS, <http://arcs.uflib.ufl.edu/>). A key database need identified in blueprinting was the ability in the database to track observations at a particular site in space, and not focus on tracking observations recorded by an individual sensor, which could change locations over time. The workflow we have developed for water quality management (Box 3) addresses goals and special concerns identified through whiteboarding but requires open source computational tools, some level of knowledge of computational tools (e.g., MySQL and R) and version control (e.g., GitHub), which are tools essential for basic data management. While this example is specific to the LCR project, we feel that the workflow developed could be implemented in similar restoration efforts.

Water Quality Workflow

Extensive details on the MySQL import process are provided in the project management library (Moreno et al, 2020, MYSQL workflow for the LCR Oyster Project). A step by step guide and overview is provided below (Figure 1-4):

1. Datasheets are standardized and include pre-populated fields including the location and date to minimize error.

2. Water quality hourly sensor observations are downloaded in the field to a field laptop while simultaneously notes are made on paper datasheets related to field weather conditions and equipment status. Water quality observations using the YSI device are also taken and recorded on the same data sheet.

3.A. Water quality sensor files are then uploaded into a secure University of Florida internal server and a trigger starts the Python import process into the MySQL relational database, which permanently stores raw files as an archive. The YSI measurements are manually entered into our MySQL relational database in its appropriate table.

3.B. QA/QC R scripts pull and process the water quality observations to check for flatlined or out of bound measurements (i.e., outside of expected range).

3.C. Processed data, edited scripts, and documents are then stored and updated unto GitHub. Standardized GitHub workflows are used during collaborative projects to ensure proper version control utility (Moreno et al., 2020, GitHub Workflow for the LCR Oyster Project).

Adding Water Quality Measurements to our Permanent MySQL Relational Database and Version Control

We use dedicated username and password controls to maintain access to the MySQL database to track access and database changes. We store water quality data in specific tables where the sensor serial number and location must be pre-defined prior to importing the sensor observations. These pre-definitions allow us to track which sensors are in which location at a specific time. The MySQL database relates to multiple data tables through foreign keys (e.g., specified MySQL columns) and in our project the tables are related through sensor serial number and site location (Figure 1-5). We use R scripts to pull these unedited observations and process additional QA/QC procedures. These processed observations and their accompanying scripts are then updated using version control in the project GitHub master data repository (https://github.com/LCRoysterproject). This repository includes an up to date master branch which is protected from any unintended or incorrect updates using GitHub repository restrictions. To submit any changes to the master branch (referred to as a pull request), it is mandatory to have the changes/edits reviewed by another member of our team to ensure data integrity. Every pull request requires a thorough message describing each change, in the event an update to the master branch has to be investigated. Version control allows for team members to view a previous iteration of the master data branch and go back to that iteration if needed (Perez-Riverol et al., 2016). This workflow protects the master branch from possibly merging accidental or incorrect changes, giving a layer of needed security to the data.

Automated Data Checks Through Python and R Scripts

Water quality observations are imported into our MySQL relational database through custom Python scripting. The Python import process provides QA/QC procedures such as flagging duplicate water quality observations. If observations are flagged through the Python import process a review takes place to find out why the observations are labeled as a duplicate. All unique observations are imported into our MySQL relational database, where they will be additionally reviewed via R programming scripts. The R scripts check for out-of-range measurements and additional scripts remove flatlined data observations which usually suggest biofouling on the sensor. Additionally, water quality visualizations help check for data integrity. The R scripts are not automated, but they do provide a way to provide quick and efficient checks on the data as they are migrated to the database.

Oyster Observation Workflow

Detailed information on the dual data entry system using a structured data packet is available in the data entry documentation for the Lone Cabbage project (Moreno et al., 2020, Data Packet Structure for the LCR Oyster Project). Several of these entry processes are similar to those in the water quality workflow and will only be briefly reviewed here where (Figure 1-5):

1. Datasheets are standardized prior to going in the field include pre-populated fields including the location and date to minimize error.

2. In the field, counts of oysters are recorded by team members on datasheets by hand.

3.A. In the lab data are entered using a dual entry system and data validation tools are used to ensure that the data entered are within range and standardized (e.g., site location, capitalization, appropriate oyster height range, etc.).

3.B. Standard R scripts are used to estimate oyster densities (e.g., population abundances) and power analyses are done using these data as they are entered to inform field sampling efforts within the field season.

3.C. Processed data, scripts, and documents are then stored unto GitHub. Standardized GitHub workflows are used during collaborative projects to ensure proper version control utility.

Datasheets, Data Entry and Validation of Oyster Data from the Field

We developed standardized datasheets for recording information by hand from field observations. These datasheets were designed to (1) clearly detail format of information to be recorded, (2) minimize errors, (3) allow for easy transcription from field observation, to paper, to entry into the computer. Observations of oyster counts and measurements from the field transects are entered into the data entry form built in Microsoft Excel. Within Excel, data validation checks are automatically applied. These data validation checks provide initial assessments that every new manually entered observation is restricted and limited to what is applicable for that column. As an example, restrictions include oyster height measurement ranges, site location names, and acceptable date ranges for surveys. For some types of data entry, such as site names, we use drop down menus such that the person entering the data must choose the name of the site from a predefined list, instead of typing the name. This reduces the changes of entering a name incorrectly but would still allow for an incorrect name to be selected. For other types of data entered, such as oyster heights, we pre-define a minimum and maximum range that is expected of any given oyster height such that if a value entered for a height falls outside the accepted range, then, a warning is issued and must be addressed. We also require that two people separately enter oyster observations, in two separate Microsoft Excel tabs. An additional Microsoft Excel tab compares the two entry sheets to determine whether the separate entered versions are identical. If the dual-entry versions are not the same a “check” notification will appear on the Excel cells (e.g., the cell column and row number) that do not match. The flagged cells will then be reconciled by a third team member, who will investigate the discrepancy using the original data sheets. The process of a dual-entry workflow is known to significantly reduce data entry errors (Barchard & Pace, 2011).

Adding Oyster Observations to a Central Storage and Version Control

Reconciled oyster observations are ultimately stored in our master data repository on GitHub and team members are required to follow the same workflow as defined for water quality data to upload this info to GitHub. The workflow ensures that every new type of oyster data updated are reviewed prior to merging with the protected `master` branch. It is also important to note that oyster measurements are not stored in MySQL since our MySQL relational database was created specifically for water quality observations. Oyster data are also stored in a University of Florida protected server with remote storage as a backup to the GitHub servers.

Regularly Updated Data and Adaptive Management

We have developed scripts of computer code that use these data to create summaries and routine analyses to provide information to (1) adaptively update sampling efforts and (2) meet reporting requirements of funding agency. Because of data workflow is standardized, each time new data are collected, then these standard computer scripts can be run to inform ongoing research efforts. As an example, to inform field sampling we routinely use a type of power analyses to inform field efforts during winter oyster sampling. Prior to the field sampling season, data from previous years are used to develop preliminary sampling guides in terms of location and number of transect samples to take. As field collections begin and new data are collected and entered into the computer, we re-run the power analyses to then update the sampling effort for that field season based on observed oyster density and variability from within season sampling. This allows us to allocate effort to locations where they provide the most information to meet project objectives. This increases overall project efficiency and maximizes learning, but is only possible because of a robust functioning data workflow.

Discussion

Establishing a data management workflow is receiving more attention in ecological efforts. Thus, creating a data management workflow from the beginning of the research initiative makes data management an easier endeavor to maintain than trying to reconcile and document the aspects of the study after a manuscript has been prepared (Archmiller et al., 2020) or a research project completed. Data and scripts without proper initial data management workflows can lead to an increased effort, likelihood of mistakes and increased time to to properly archive and clean data following data collection. While it is possible for post-reconciliation of data collected in theory, this rarely happens in practice (Nelson & Grubesic, 2018). Our data management addresses many of the challenges with “living data” such as reducing human introduced error, permanent data storage, and version control for text and scripts. Our workflow uses a mixture of familiar software (e.g., Microsoft Excel) and versatile software (e.g., R programming, GitHub, MySQL), which encompasses an array of skills needed from a biologist to employ a modern workflow.

Our data management workflow may not work for every ecological project, however many elements discussed in this paper should still be applicable. The key concept of creating a data management workflow prior to beginning the LCR project is one of the main reasons this workflow has persisted throughout the project, because in developing the workflow this necessitated key elements of the project, such as location of water quality sensors, to be determined. While initialy the time required to develop the datawork flow was signficiant, and in many ways equal to the time allocated in the early phases of the project to the ecological questions simultaneously being addressed, the time saved through using a datawork flow through having ready access to the data for project planning efforts and identifying problems with water quality sensors rapidly has likely exceeded the initial planning time required for the data workflow. The principles of securing and validating data should also be considered of high importance for monitoring efforts as well.

There are many advantages to using open-sourced tools (e.g., GitHub, R programming, and MySQL) in a data management workflow. Firs, these software are free, widely available, and have an active support environment of users online. Second, this workflow can be learned by biologists and others interested in using these data after basic instruction in R and data management principles from online training programs such as The Carpentries (<https://carpentries.org/>). Many universities also offer R programming courses which teach the basics of statistical analysis with R (e.g., WIS 4601, Quantitative Ecology, <https://wec.ifas.ufl.edu/undergraduate-students/undergraduate-course-listing/>) and similar data management techniques described in this paper (e.g., WIS 6934, <https://datacarpentry.org/semester-biology/> ). Using GitHub offers much desired flexibility in code development through “pull requests” (Rahman & Roy, 2014) and version control (Blischak et al., 2016). GitHub consistently updates their software features making it a reliable resource for many projects.

Some initial difficulties to our workflow may arise in teaching team members how to use the workflow and to ensure that they are following workflow processes. Because these types of workflows are not common in ecological studies, team members likely have no experience with this type of framework. However, team members likely do have experience with topics such as challenges in sharing documents among a team and presenting this workflow as a solution to these types of problems may make the traninig process more acceptable to the team. It is important to communicate effectively with team members to guarantee they are collecting and maintaining data within the workflow procedures. Another limitation to the current workflow structure is that the MySQL database can only handle only numeric and character data and is not able to store maps or other image types. However, for the application of the LCR project, this is not a constraint. A MySQL database does require SQL programming knowledge such that fundamental changes to the database structure require more advanced programming knowledge. However, as I have shown, interfacing with the MySQL database through a more common language such as R is readily doable as part of a team of applied researchers.Other possible limitations to the existing workflow include MySQL file size limitations, which given our current data storage needs would likely not be exceeded for decades of data collection. The use of GitHub respositories could in theory have limits of 1 GB and up to 100 MB for an individual file (<https://help.github.com/en/github/managing-large-files/what-is-my-disk-quota#file-and-repository-size-limitations>), which can make it difficult to store large files without compressing them. However, despite these limitations with MySQL and GitHub, their functionality greatly outweighs their restrictions.

Adaptive management is described as a process which continually improves management policies and practices based on data outcomes (Pahl-Wostl, 2007). With increasing types of data available to inform decisions related to management, the use of data workflow structures to explicitly improve the adaptive management life cycle and improve decision making is clear. However, adaptive management programs continue to not be widely implemented in practice (Walters 2006Weimer et al., 2007.). Funding programs such as GEBF and others from the consolidated Deepwater Horizon oil spill settlements explicitly call for adaptive management programs to be used to improve restoration decision making. This chapter has demonstrated how the data workflow can be incorporated into this type of program under an adaptive management umbrella.

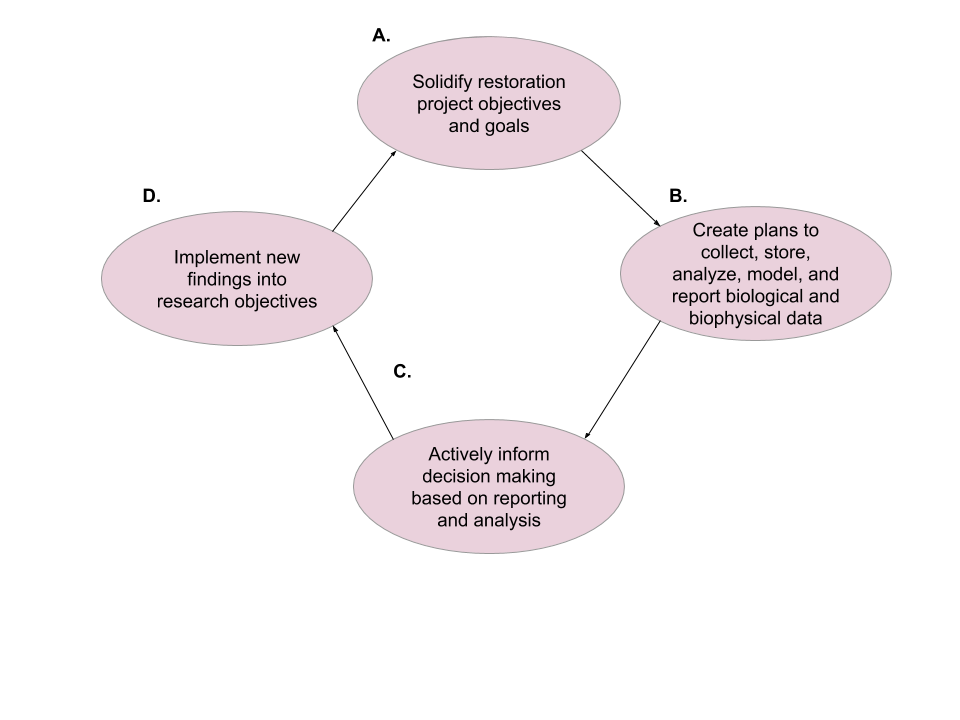


Figure 1-1. The adaptive management process for ecological restoration projects. A reliable data workflow contributes at each phase of the adaptive management process from planning to implementation.

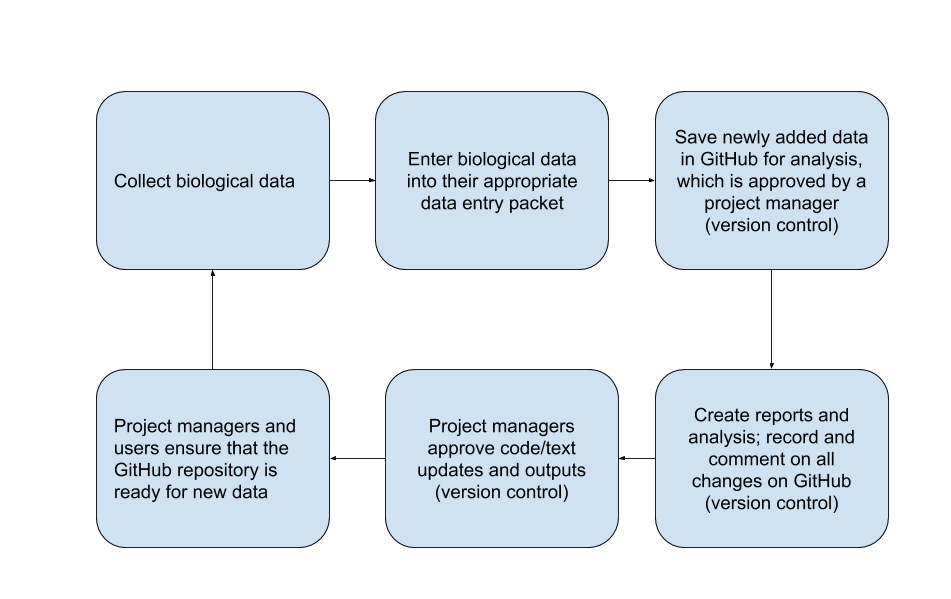


Figure 1-2. Generalized version control workflow for the LCR project, detailed workflow information can be found here (Moreno et al., 2020, GitHub Workflow for the LCR Oyster Project).

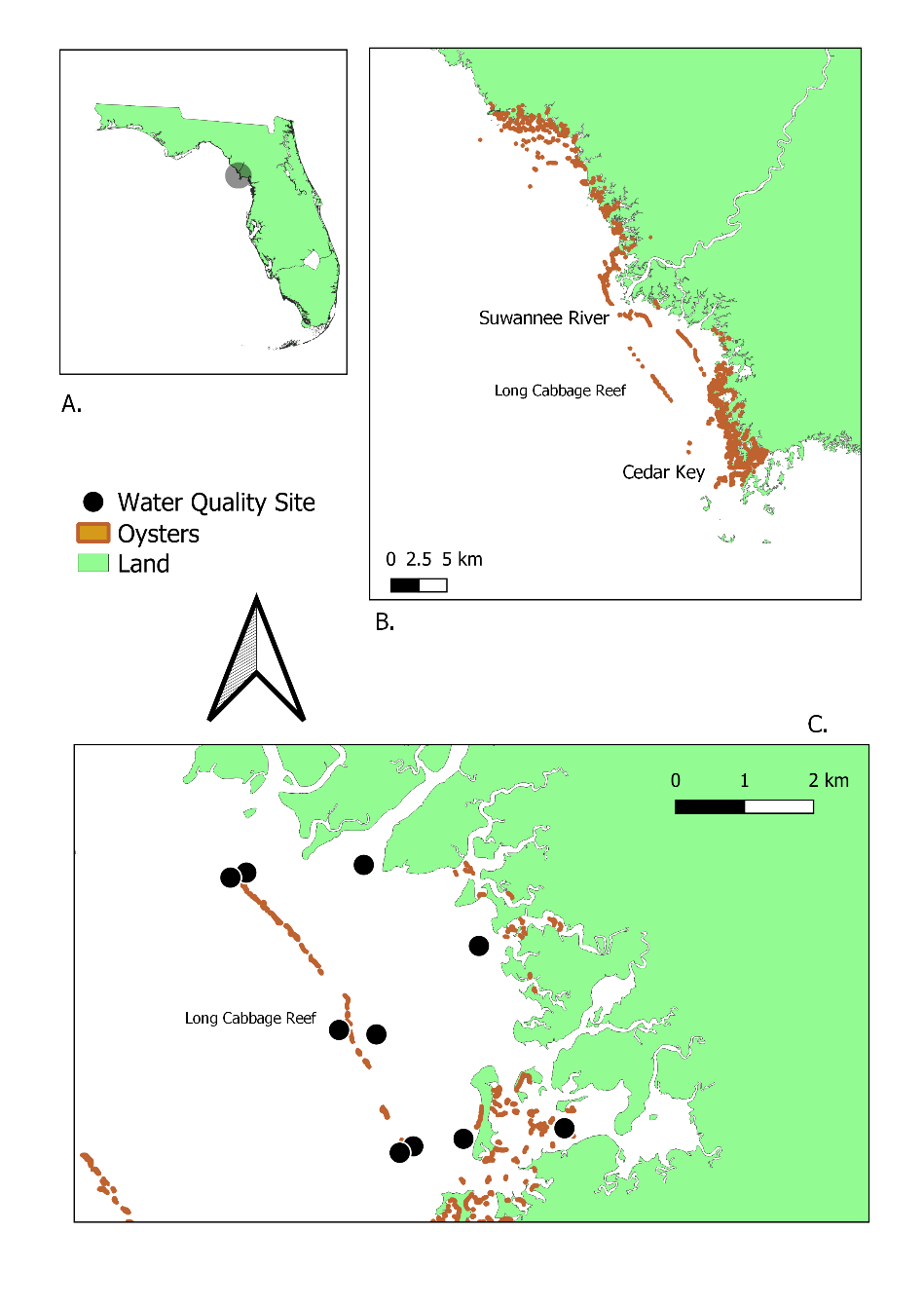


Figure 1-3. Water quality location map. A) Map of Florida identifying Lone Cabbage Reef area; B) Florida coastline between identifying Suwannee Sound river mouth, Lone Cabbage Reef, and Cedar Key, FL; C) Lone Cabbage Reef with water quality sites identifies (black circles). The oyster shapefile used in this map is a from a University of Florida sampling effort in 2001.



Figure 1-4. Schematic describing the data workflow for the Lone Cabbage Reef water quality observation network.

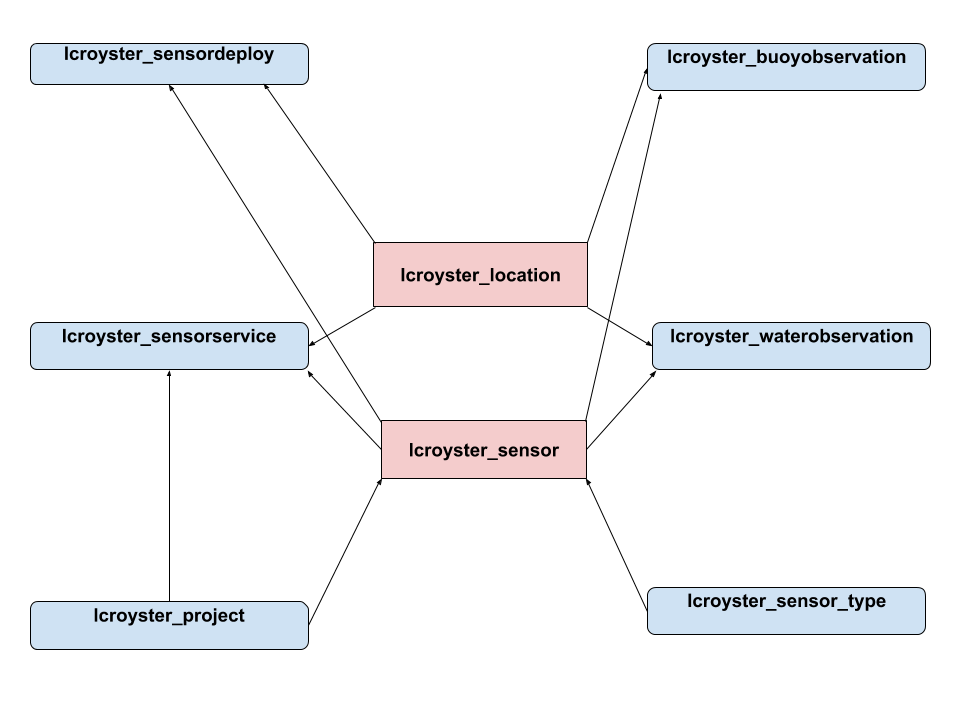
Figure 1-5. Diagram of how the tables in the Lone Cabbage Reef MySQL relational database are connected by a sensor’s location (lcroyster\_location) and serial number (lcroyster\_sensor). More information on the description of the tables and the data they archive can be found at Moreno et al., 2020, MYSQL workflow for the LCR Oyster Project



Figure 1-6. Schematic describing the data workflow for the Lone Cabbage Reef oyster samples from field collections.

CHAPTER 2

TAILORING GITHUB FOR ECOLOGY

Introduction

Traditional field ecology is currently experiencing a data revolution. Advances in technology ranging from satellite observations to autonomous sensors create new opportunities for ecologists to use data that are regularly updated (“living data” Yenni et al., 2018) to study ecosystems (). These expanded data resources often require scientists to work effectively and efficiently with these “living data” to maximize knowledge gained from these data sources. Yet despite new data sources commonly becoming available, researchers are rarely trained in best practices to capture, manage, and analyze these types of data (Lowndes et al., 2017). To understand these living data ecologists must adapt new tools as part of their analyses toolbox to manage and interpret this information to accomplish basic needs such as informing future field sampling plans or assessing relationships among different variables (Mislan et al., 2016). Developing a framework to efficiently write, share, and archive computer code used to analysis data in a reproducible framework (i.e., can be readily interpreted and done by others in the future) is a key aspect of “open” science platforms which encourages transparency in data availability and analyses (reference) (Alarid-Escudero et al., 2019). Several software platforms including GitHub are designed as collaborative tools to facilitate teams of people working on software coding projects. These tools were originally developed to share computer code across teams working more on information technology projects (ref). But as ecologists are increasingly working as part of collaborative teams and adopting analyses and data sharing frameworks that promote transparency and reproducibility, GitHub has become an increasingly common tool for these researchers as well (Gilroy & Kaplan, 2019).

The LCR oyster project is a large restoration project in Suwannee Sound, Florida (Chapter X) to restore an oyster reef to provide local and ecosystem level benefits. This project generates data from multiple sources including observations of oyster populations and measurements by field biologists and continuous autonomous water quality data via sensors. These data are updated at different frequencies and require specific attention to be processed. However, once they are processed there is a need to store these data so that they may be used among project team members and collaborators. I developed a GitHub version control system to track living data and code used to analyze these data in a framework that could be shared across multiple users. This paper describes how this system was developed including foundational concepts related to standardizing naming conventions, GitHub repository structures, and managing data availability to different LCR project repositories to increase reproducibility and transparency. This chapter uses the LCR project experience as a case history and demonstration project of developing a GitHub repository system for a collaborative project sharing common types of ecological data.

Terminology overview

* GitHub- an online version control software, free, and accessible to anyone with Internet ([www.github.com](http://www.github.com))
* living data- data that are collected and updated frequently to continuously analyzed (Yenni et al., 2018)
* version control- a system which allows the users to track iterative changes to code and text (Blischak et al., 2016)
* project repository- term used to identify one type of analysis that is conducted on an LCR project dataset
* README.md- a markdown file which includes information about folder and files contained in the repository
* user- any person using GitHub
* team member- specifically referring to an LCR project collaborator
* admin- specific members of the LCR project tasked to monitor project repository’s pull requests in GitHub
* pull request- a process in GitHub to submit changes and a message containing a description of the change, in order to track the version of the change, an additionally an admin can review and accept the changes
* ”master“ branch- the first branch created with a GitHub repository, in the contexts of the LCR project, it is the most up to date branch with the most user limitations for pull requests, this branch requires admin reviews
* branches- essentially copies of the GitHub repository master branch, in the context of the LCR project each collaborator specifically make edits to their own branch prior to a pull request to the `master` branch
* merge conflicts- when branches have competing commits during a pull request, needing to be resolved by an admin (<https://help.github.com/en/github/collaborating-with-issues-and-pull-requests/about-merge-conflicts>)
* commit message- a written text of why a pull request is being submitted

Github and Version and Control

Computer scripts and text within these scripts are the critical instructions documenting how a data analyses was done to generate results used to reach a conclusion or inform a decision. How this code is written can vary between person to person, and when teams of people are collaborating on coding projects, identifying how one person did or did not edit a section or code, or what changes were made to a code that may be caused the code to stop working, or improve the code, can be difficulty to readily identify (Blischak et al., 2016) . To keep track of these changes, version control systems which track changes to the computer code by assigning identifying version characteristics to each copy of the code before the code is consolidated, may be implemented. Version control systems can help track changes in various ways such as comparing different versions of the code and highlighting where keystroke differences exist, and then requiring the user to write comments in plain language describing each change made to the code by the user before allowing those changes to be incorporated into a new version of the code. The version control software will then create a unique version identifier for each iteration and allow the user to revert back to those changes if needed (Noble, 2009). This is especially useful when a user decides to share their code or text with other collaborators. When the user receives new comments from the code/text they have shared, they are able to see what has been changed in relation to the original document. The user will then be able to implement those changes through the version control system so these changes are tracked and transparent to other users.

In the LCR project, code collaboration is common. Before the development of the GitHub repository structure, code writing collaborations in our project ultimately ended in several files of the same script but each file was a slight iteration of each other (e.g., rscript\_1.R, rscript\_2.R). It became confusing which script was the most up to date, and which script should be used for which analysis. Using version control archives each iteration of each file (in case a previous version is needed), while only keeping the most up to date version of that file active in the repository. This makes it much easier to find the most up-to-date file for use across collaborations. We utilize GitHub to keep a record of which files have been changed, who has changed them, and why they were changed. Using a version control software has saved us timed in trying to determine when and why certain script changes were implemented.

Challenges Working in One Repository: Lesson Learned

As the LCR project started to collect a consistent stream of data, it became apparent that its GitHub structure as original designed was becoming increasingly more difficult to maintain and manage. One of the main complaints of users was that it was difficult to find scripts and their data source. Collaborators working in the main GitHub repository were not always following repository guidelines, however the guidelines at the time did not address many of our newfound needs such as how to account for multiple working projects using the same data. Without having proper guidelines, the main GitHub repository quickly began to grow and expand (Figure 2-1) in a way that was unmanagement. Our main GitHub repository started to store multiple projects inside of itself, leading to a confusion of what was in the repository was which data and scripts were used for each of the different projects. We soon came to realize that the GitHub repository structure we had employed was not effective in keeping our files or projects organized. This led to a major revision in GitHub repository structure.

Creating a New Github Repository Structure and Workflow: LCR Project v2.0

The main goal of the GitHub repository version 2.0 for the Lone Cabbage project to keep different sub-projects separate (i.e., water quality summaries for reports separate from water quality analyses) but to have common living data sources. We extracted the different projects inside our main GitHub repository and created individual project repositories. Each project repository follows the same guidelines for folder structure (<https://github.com/LCRoysterproject/repo_structure>) defined earlier, but instead of the entire project falling under one repository, separate repositories were created. These new project repositories also include descriptions in their README.md file about the folders and files inside of them. These README.md files are essential in maintaining transparency of what each script does and what their outputs are. Our project team members are required to make updates to README.md file’s as they create new files for scripts and text, ensuring clarity and transparency in the repository.

Why is this structure an improvement over the first version of the LCR project GitHub repository structure? First, these individualized project repositories are self-sustaining and only team members actively working on that project have access to them. Second, these repositories are independent from one another, and their scripts are not influenced by other project repositories’ scripts. Because some of these project repositories need to access the most up to date LCR project generated data (e.g., water quality, oyster measurements, oyster counts) it was imperative that our workflow included a way that the project team member could access these data. In the second version of the structure, all LCR generated data are processed and then stored in a master data repository. This data repository may contain relevant data which could be used among different project repositories (Figure 2-2). This master data repository also contains commonly used scripts and text for routine project reporting efforts (e.g., water quality MySQL extraction code, sampling power analysis). Every LCR team member has access to this master data repository. In the event in which a team member needs to update the data for their analysis, they are able to do so without limitations or approval. These data are also protected from being wrongfully edited or deleted by GitHub branch permissions. Only LCR project admins are allowed to make updates or changes to the master data repository, which adds an extra layer of security to this repository.

Controlling Access via Github Permissions and Branch Workflow

GitHub has settings which can limit who can edit or modify a repository’s branches (Perez-Riverol et al., 2016). This is useful for protecting some branches from permanent change, while at the same time allowing the files to be used by multiple team members. As an example, for each LCR project repository which has collaborators, we establish a protected `master` branch, and open collaborator branches. Collaborators are able to edit and modify their own branches however they please, but they are not able to update or modify the ”master‘ branch unless approved by a project admin via a pull request. Project admins are expected to review a pull request rigorously and work with the collaborator if there are any discrepancies in the pull request. Using a system that checks the work of collaborators has helped us reduce errors in code, text, and data and can be implemented across many ecological efforts.

Furthermore, GitHub has the functionality to make repositories public or private. Whether a repository is public or private is ultimately up to the administrators of that project repository. Public repositories are open and searchable to the public. Private repositories are only initially viewable to the creator of the repository. Additionally, in the LCR project we also limit the users who have access to any given repository. All users have access to the master data repository, however they do not have access to other collaborator repositories, unless an LCR admin grants them access. We allow some repositories to be public and protected, and actively worked on repositories to be private. GitHub allows project managers to change the status of any repository, to private or public, at any time. These types of repository functionalities can allow many ecologists to actively work on their research while protecting their data, analyses, and findings.

Naming Conventions for Repository, Files, and Folders

Proper file naming conventions help users understand the contents of the file without having to click on it. For scripts, naming conventions exists in which if a script file creates a function or a certain output the output file should also be named the same (<https://style.tidyverse.org/package-files.html#names-1>). In the LCR project we created a set of consistent set of guidelines for filenames (Table 2-1). The overall guidance to naming files is to keep the cases consistent, in our case lowercase, and to keep the structure of naming the file the same. We use an overall naming guidance of study, location, and project summary, in that order, to name files. These guidelines help each collaborator really think about: (A) why they are creating a new file; and (B) what does this file ultimately intend to do. We hope that these file naming guidelines it will help other ecologists in pursuit of standardizing their filenames and increase their file transparency.

Transparency

GitHub has options to increase transparency for an ecological project. Hosting a public project repository on Github can improve the probability of researchers and the public to find the repository, and possibly collaborate. Commit messages, through a pull request, are easily seen and located in GitHub, and allow for collaborators to understand any change submitted to the repository. Transparent repositories are unlikely to be “scooped” by another researcher that can claim the data and the analysis is theirs, through the continuous stream of commit messages leading to the final product. An additional benefit to a transparent repository is that many eyes will be available to evaluate code and text, which can increase the time is takes to debug a script issue. Generally, most scientists with interested in related research are more willing to collaborate with the original scientist than compete with them (Prlić & Procter, 2012).

Discussion

Ecology is an ever-growing field of science with rapidly changing data streams to help advance knowledge. However, without tools in place to organize, manage, and serve these data, the utility of these data streams can be limited. Living data are common in many ecological studies, and these data provide unique challenges in management and reproducibility. I have demonstrated how using a well-established GitHub repository for living data can lead to effective and easy data sharing (White et al., 2013), which increases the transparency of the effort of collecting and analyzing data leading to a greater impact (Piwowar et al., 2007) from the research.

We employed a repository structure workflow on the recommendations of the University of Florida Academic Research Consulting & Services (ARCS). The repository structure we developed for the LCR project solves many of the challenges we were facing while working with our own generated “living data” and multiple working projects. Separating the multiple working projects into their own repositories keeps the projects simple and organized. There is also little confusion on what is in the repository specifically since every repository is required to include a README.md file describing every folder and file. A README.md file also shapes the first impression of a repository and increases the searchability of repositories in GitHub (Prana et al., 2019), which may lead to greater transparency and collaborative efforts (Jones, 2013).

Advantages of our approach include that GitHub and free and accessible to anyone with internet. There are also many training programs which can teach a user how to utilize GitHub efficiently for their project (e.g., <https://guides.github.com/activities/hello-world/>, lab.github.com). Another advantage in using GitHub is that if a repository is accidently deleted, a user has 90 days to retrieve the repository. A benefit in using our described approach is that it can be applied to any ecological effort which has a consistent stream of data by allowing a master data repository to be accessible to all team members while still protecting the repository from adverse or unintentional changes. Additionally, this type of workflow can easily be maintained by a small group of ecologists with basic GitHub workflow training (Yenni et al., 2018).

Some disadvantages to our approach include that ecological projects, conducted in smaller teams, normally do not have the funding to hire full-time data manager (Hampton et al., 2013), making it a timely effort to curate an efficient data management workflow, initially. Data management planning is also typically and generally underutilized and underappreciated in ecological project designs (Michener & Jones, 2012). It is up to the project team member to manage the data, take the initiative, and adhere their GitHub data management workflow. It is also important to note that even though GitHub has been discussed in detailed as an effective way of project organization and transparency, it might not be a one stop solution for reproducibility in science (Ram, 2013).

The approach we have described in this paper is meant to be a guideline for ecological efforts who desire to make their project organized through concise workflows, standardized naming conventions, and well documented README.md files. Our hope is that this paper can serve as a mechanism for designing version control software, such as GitHub, to meet the needs of an ecological project with a continuous stream of data and multiple working analyses and projects. Increasing transparency through managing a well-documented repository and through README.md files, may also lead to useful future collaborations. The investment in creating a thorough data management workflow in GitHub will help decrease the time is takes to effectively reproduce analysis by reducing the time it takes to locate files and their outputs, which will allow ecologists more time to analyze and interpret their data and less time trying to manage it.

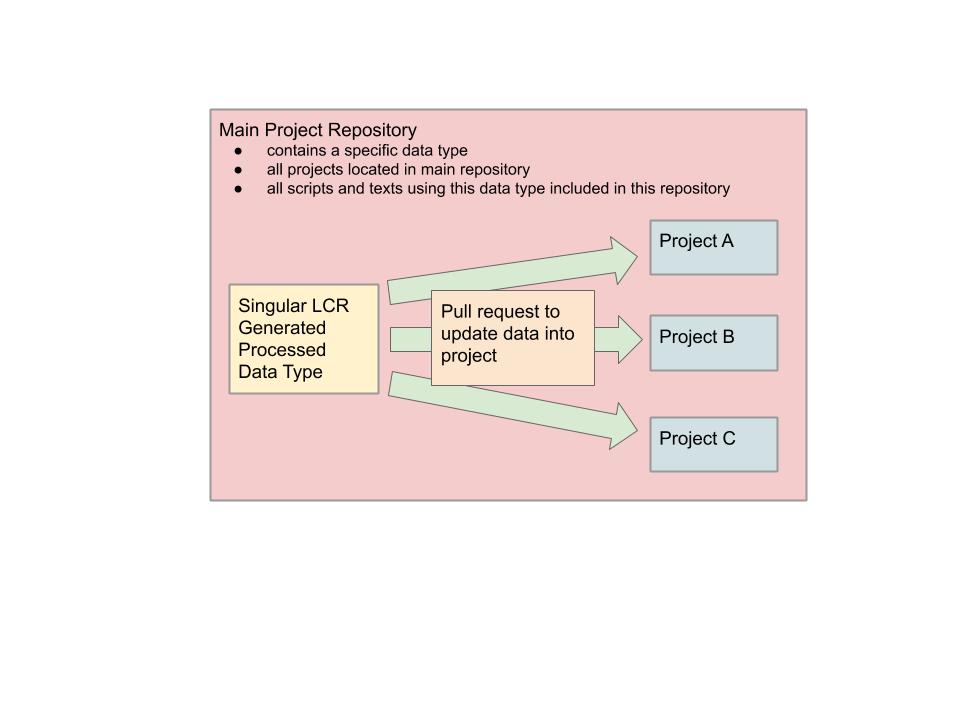


Figure 2-1. Visualization of our main project repository structure and various projects in the same repository. The box of the visual encompasses all of the projects, code, and text belonging to a single data collection type. Multiple projects were located in the single repository usually discernable by separated folders. Confusion arose when projects used scripts and data from other projects without proper documentation.

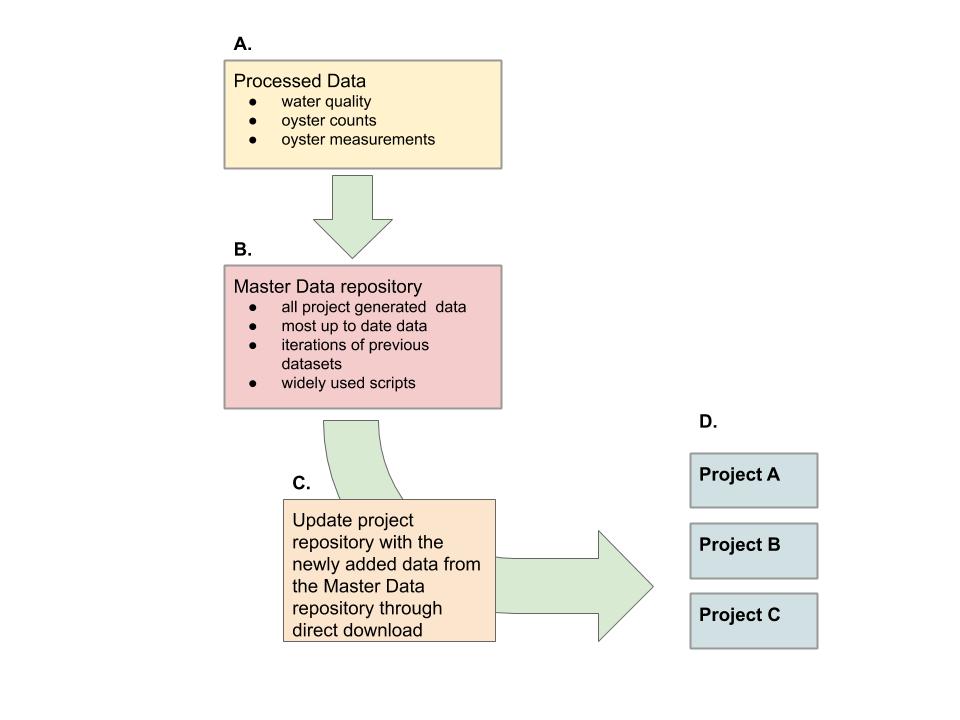


Figure 2-2. New LCR project workflow which describes how project repositories update to work with newly added data. A. Data are processed and cleaned via MySQL or scripts, B. Data are updated to the master data repository, reviewed, and approved by LCR project admins, C. Project repositories update their repository data by downloading the data directly from the master data repository, D) Project repositories conducting individual analysis on LCR datasets with newly downloaded data, ready for reporting and analysis.

Table 2-1. Table of naming conventions for file types, example, and description of the example

|  |  |  |  |
| --- | --- | --- | --- |
| File Type | Naming Convention | Example | Definition |
| Project Repository | study\_location\_projectsummary | bird\_bb\_monitoring | Big Bend camera and survey bird monitoring project |
| Scripts | lowercase, no uppercase (snake case) nor all caps, all names with separate words need to include a underscore ( \_ ) and no spaces, no dates in the names unless it helps with the descriptions of the content, script file names should be descriptive and concise. Scripts that have a single output should be named in a similar fashion to its filetype output. | discharge\_1941\_2018\_quantile.R | R script which reports quantiles from river discharge from 1941 to 2018 |
| Figures | study\_location\_type\_summary.filetype | oys\_lco8a\_map\_transect.tiff | oyster transect on reef element LCO8A map in a tiff image |
| Tables | study\_location\_summary.filetype | wq\_lcr\_inshore\_vs\_offshore.csv | LCR water quality inshore and offshore comparison |
| Data | every dataset file is required to be in lowercase, no uppercase (camelCase) nor all caps, all names with separate words need to include a underscore ( \_ ) with no spaces, no dates in the names unless it helps with the descriptions of the content | discharge\_1941\_2018.csv | River discharge data from 1941 to 2018 in a text file |

CHAPTER 3

CASE STUDY: DEER ISLAND, FLORIDA TIME PERIOD SHORELINE ANALYSIS USING DSAS

Introduction and Background

Shoreline changes can occur due to multiple factors including SLR (sea-level rise), anthropogenic human activity and hurricane intensity (Yu et al., 2011). The combination of these processes can influence erosion and accretion. Shoreline changes may affect a shoreline’s resilience to storm surges including flooding and species diversity implications (Desantis et al., 2007). It was observed by USGS (United States Geological Survey) that shoreline changes along the Gulf of Mexico, specifically in Florida, were relatively steady between the 1800s and 1990s (Morton et al., 2004). Since then, the Gulf of Mexico coastline, with its low relief geomorphology particularly along the west coast of Florida, has been noted to be vulnerable to coastal erosion (Geselbracht et al., 2011).

Climate Change and SLR

More recently, climate change induced SLR and its impacts on coastal zones is of growing interest. The Earth’s climate is warming due to an accumulation of greenhouse gases in the atmosphere, largely in part due to anthropogenic fossil fuel burning and deforestation. Warming climate change causes thermal expansion of sea water, and land ice to melt into the ocean, initiating SLR (Cazenave & Cozannet, 2014). Sea-level rise is considered to be a likely candidate for widespread global erosion. Erosion occurs when SLR drifts the high-water line (line on the shore where the water usually reaches at high water) landward in relation to the slope of the coastal area. Erosion on sandy beaches involves the relocating of sand from inshore to offshore. Storm events temporarily increase the local sea-level of the sandy beach, and ultimately storm waves are able to reach higher elevations on the beach, causing erosion. This exchange implies that sea water levels have a direct relationship with sandy beach erosion (Zhang et al., 2004).

Characteristics of Sandy Shorelines and Sedimentation

Sandy shorelines are characterized by active environments and unstable substrata, which consists of sand, mixed sand, quartz, and/or silica. The unstable nature of sandy shores make a harsh ecosystem for biota and may incorporate a significant range of physical environment conditions and ecosystem functioning. These shorelines accumulate sediment accretion by wave deposited particles. Particles originate from inland erosion and may be transported along rivers (Brown & McLachlan, 2002). Sediment to sandy shores may also be added by marine biogenic sources such as pieces of marine skeletons, sponge spicules, and shell fragments (McLachlan, 1990). Threats to sandy shorelines include disruption of sand transport, storms, SLR, and human activities. Our study site is in a siliciclastic, sand-starved and low-wave-energy system dominated by marshes that open towards the sea (Hine et al., 1988). Shoreline profiles may drastically change over time due to these low energy wave systems (Jackson et al., 2002).

Suwannee River Sedimentation and Discharge

The Suwannee River is the second largest river in Florida spanning 396 kilometers long and is considered to be a significant point source of sedimentation, approximately 11 kilometers north of our study area. The Suwannee River is a partially spring-fed system which also drains the coastal plain of Georgia and provides a restricted point source input of siliciclastic sediment, creating a small 20-kilometer delta. The surrounding coastal regions of the Suwannee River are otherwise known to be sediment starved, but a great significant sedimentology event has been shown that the Suwannee River has reworked ancestral fluvial sands and serves as a source for sandier marsh sediments (Wright et al., 2005).The Suwannee River normally has high discharge peaks between February and April and low discharge peaks between August and October (Purtlebaugh & Allen, 2010). The median annual river discharge measured of the Suwannee River at the USGS Wilcox site (latitude 29.58, longitude -82.93) is 2418 m^3/s with a minimum discharge of 998 m^3/s and a maximum discharge of 4971 m^3/s (USGS, 2021).

Human Development and Impacts

Three Florida counties encompass the region of our study site, Dixie, Levy, and Taylor (Figure 3-1, A and B). These counties which are projected to increase in human population by 2045 as depicted in (Figure 3-1, C). These Florida counties are recorded to have lowest population densities along the Florida coastline (Geselbracht et al., 2011) . In the future it could be likely that businesses and people will want to develop housing along this shoreline. Human development on coastlines may accelerate coastal erosion by creating a fixed position of the shoreline and stabilizing inlets (Finkl & Charlier, 2003). Increased human developments may also negatively impact coastal species diversity. Species biodiversity is threatened by the increase of urbanization and environmental coastal degradation (Finkl & Charlier, 2003) . Czech’s et al. (2000) documents urbanization as the highest cause for species endangerment. For example, the shorebird Piping Plover (*Charadrius melodus*) is known to forage and nest in areas of low human population (Thomas et al., 2002), implying that shoreline areas with higher human densities would not be an ideal habitat for this species. Species biodiversity, both vegetative and animal, could be at risk due to an increase of urbanization along coastlines (McKinney, 2006) and accelerated shoreline erosion.

Big Bend Habitats for Species Richness

The Northeastern Gulf of Mexico region of Florida is ranked as an area of high importance for conserving and protecting habitats for at least 30 species of shorebirds. Within those thirty species, four threatened species are considered to be of “extremely high priority” for protection, and include the American Oystercatcher, Red Knot, Snowy Plover, and Piping Plover (Withers, 2002). The coastlines in the Big Bed region (Figure 3-2) are described as having low wave energy (described as waves falling well below the high-water line of a shore), which can be ideal for migrating shorebirds because low wave energy on shorelines can facilitate the accumulation of vegetative litter (Nordstrom et al., 2006). These shorebirds use the primarily cordgrass marsh shorelines habitats of the Big Bend for foraging, mating, and shelter. Shorebirds in the Big Bend have been documented to have the least abundance and species richness, in a study comparing Gulf of Mexico regions shorebird use of coastal habitats (Withers, 2002).

Major Hurricanes Near the Study Area

It was reported that intense storm clusters hit the Gulf of Mexico between 1994 and 2015 (Sankar et al., 2018). During this time frame the Storm of the Century, hit the west coast of Florida on March 1993, with wind speeds of > 15 m/s for 16 hours recorded at the Crystal River Power Plant roughly 100 km south from our study area (Goodbred & Hine, 1993a). The Storm of the Century caused devasting damage to the Waccasassa Bay (approximately 30 kilometers south of our study site), 3-meter water storm surges, a storm deposit which reached 12 cm on the levees and up to 2 cm on the marsh surface (Goodbred & Hine, 1993a). Hurricane Irma, similarly hit the west coast of Florida on September 2017, was recorded as a Category 3 hurricane with the highest recorded wind gusts of 64 m/s where it made landfall near Marco Island, Florida and continued northward along the west coast of Florida, ultimately causing excessive rain and storm surges near Cedar Key, Florida. The highest recorded rainfall along the path of Hurricane Irma was recorded at a peak of 55 cm in Fort Pierce, Florida. The highest recorded storm surge was recorded along the near the Matanzas Inlet at 2.3 m (east coast of north Florida). The combination of heavy rainfall and storm surges overflowed at least 32 rivers and creeks which led to many creeks and rivers overflooding near St. John’s river (Pinelli et al., 2018). Hurricane Betsy was a storm that impacted the Big Bend region in 1965 and the damage caused from this hurricane influenced a change in the areas’ low wave energy ultimately contributing to permanent shoreline loss which amounted to 12 meters (Warnke et al., 1966).

Reason for Effort

This shoreline analysis study was endeavored to identify possible factors that may be influencing shoreline loss. Since the study site is uninhabited, and tourism is not prevalent these were not considered as probable factors. The analysis was split into three time frames in order to locate an area of shoreline change where an identifiable factor may have triggered shoreline erosion or accretion. Two out of the three time frames spilt up the available imagery into equal years, however there are not an equal amount of imagery available covering each 12.5 year period (imagery spans a total of 25 years). The last time frame includes all imagery to calculate how much total shoreline was loss or gained from the years 1994 to 2019.

Shoreline loss as also need near our study site has also been captured recently. In the mid-1960s the US Army Corps of Engineers constructed spoil islands as part of the cross Florida barge canal project. These spoil islands consist of a straight line of islands perpendicular to the west Florida coast. Coastal changes have severely eroded or inundated these spoil islands, thus reducing habitat for animals (Vitale, 2019). Derrick Key is an example of a spoil island that was clearly visible in aerial photographs in 1982 and now the island is completely submerged (in 2016 photography). Major shoreline differences are noticeably observed in the 34 years, time between the imagery, for this specific spoil island. Large scale efforts to analyze shoreline changes in Florida have been studied in the past (Yu et al., 2011; Sassaman et al., 2017; Houston, 2015; Li & Gong, 2016) however it is interesting to note the possible effects of SLR on a smaller or regional scale, which might highlight processes which might be affecting larger- scale ecosystems and habitats.

Materials and Methods

Study Area

Our study area is located on the west-central Florida coastline in the Suwannee Sound region of the Big Bend (Figure 3-3). The selected shoreline is a small barrier island called Deer Island. Deer Island is a privately owned uninhabited island approximately 13 kilometers north of the main villages of Cedar Key, Florida. Historically, Native Americans intermittently inhabited Deer Island for thousands of years. Early Florida settlers were reported to live and camp on the island as well. The 1800 Florida census registered only 4 people to have identified this island as their home. There is a cabin near the south of the island depicted on a 1951 USGS Cedar Key Quadrangle map (USGS, 1955). This island is specifically located in the Big Bend Aquatic Seagrass Preserve and connects with the Lower Suwannee National Wildlife Refuge (http://www.beachrealtyfla.com/DeerIsland.htm). Deer Island is approximately ~0.364 km2 of total area, which ~0.182 km2 (50%) of the are lie below the mean high tide mark and are considered sovereign, ~0.101 km2 (27.74%) are upland acres, and ~0.081 km2 (22.25%) are wetlands with elevations as high as 4.3 meters. The island is densely forested with large pines, cedars, palms, oaks, palmettos and many more plant species (“Deer Island Florida, United States.”, 2021). The shoreline attributes reported on Deer Island is ~1.3 +/- kilometers of Gulf of Mexico white sand beach and ~1.3 +/- kilometers of waterfront facing the mainland (“A Deer Island A Sportsman's Paradise.”, 2021).

Imagery Selection Process

Locating relatively cloud-free imagery for a specific location in Florida can be an exhaustive effort. Since our study location is unpopulated and contains no popular historic landmarks, historic aerial images are not frequently taken. To reduce the effort on locating usable imagery, Google Earth Pro was utilized. Google Earth Pro does not capture any of its own imagery, it does however locate and use imagery, in its finder view, that is comparatively cloud-free and with the highest resolution. Google Earth Pro was able to give minimal metadata of the imagery such as which agency captured the imagery and the date of the image, when using the time slider feature. Then USGS’s Earth Explorer (<https://earthexplorer.usgs.gov/>) was used to further locate the actual imagery and collect its metadata. Upon inspection it was determined that NAIP (National Agriculture Imagery Program) was the agency that acquired the most frequent and most detailed aerial imagery of our study site. The specifications for NAIP aerial imagery require 1-meter ground sample distance with a horizontal accuracy that matches within six meters of photo-identifiable ground control points. These points are then used during imagery inspection. Contractually, NAIP makes attempts to comply with the specification that no more than 10% cloud cover be allowed in each aerial imagery tile. Aerial imagery are available as digital ortho quarter quad tiles (DOQQs) geotiffs, and which also correspond to the USGS topographic quadrangles (<https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/>). It was also important to select imagery that were fairly at the same time of the year, similar river discharge and precipitation levels. All imagery chosen are between the months of October through January. Table 3-1 includes all metadata associated with the imagery used in this analysis. Furthermore, observed weather and median river discharge were collected, including the observed weather for the day of imagery collection and median river discharge measured.

National Agriculture Imagery Program employed sensor types which had three band imagery categorized as RGB (red, green, blue), up until 2007. After 2007, four band color infrared imagery were collected and categorized as CIR/CNIR (red, green, blue, and infrared). Four band imagery is multispectral, which means the sensors can collect information from several parts of the electromagnetic spectrum. Table 3-2 includes the sensor type associated with each image. Our November 2007 image is the first image in our series which uses color infrared (CIR/CNIR). The advantage of using CIR/CNIR imagery us that it allows the user to view the imagery in a false color for NDVI (Normalized Difference Vegetation Index) analysis. In this analysis our shoreline does need to identify vegetation on the sandy shoreline, since there are clear and distinct separation of sand and vegetation.

Most GIS software can only display three bands at one time. Vegetation can be seen on Deer Island but it is not necessary for our DSAS analysis because the island vegetation is distinct and not integrated into the sandy shoreline. Additionally, the DSAS user manual does not have any recommendations for using or not using true color image composites. To specify a natural color display the settings in the GIS software should be band 1 set to red, band 2 set to green, and band 3 set to blue. True color in this instance is useful because the white sandy shoreline of Deer Island is in stark contrast with the dark ocean water and dark vegetation.

Digital Shoreline Analysis System (DSAS)

The DSAS is a GIS-based system created and maintained by USGS (United States Geological Survey). For this analysis the DSAS ArcMap© extension was used. The DSAS extension casts transects along the baselines (starting point for transects) and measures the gaps between the shoreline positions during defined years. Baselines are constructed by the user, and in this analysis was created using the Buffer tool in ArcMap©. These shoreline positions provide the basic data needed to calculate their shifts. One of each type of change metric (“Digital Shoreline Analysis System (DSAS) Version 5.0 User Guide.”, 2021) was used in this analysis, a LRR (Linear Regression Rate) for statistical analysis and the Net Shoreline Movement (NSM) calculation for the distance measurement. A LRR can be ascertained by fitting a least-squares regression line to every shoreline point in a transect. The regression line is positioned so that the sum of the squared residuals is at its most minimal value. The linear regression rate is the slope of the regression line. The NSM is the distance between the oldest shoreline portion to the youngest shoreline position for each transect, calculated in meters. The LRR statistic was used because all the data provided can be used regardless of accuracy, and the calculations are based on accepted statistical notions. In contract NSM statistics only require the baseline position and the last shoreline position to make its calculations. The justification for using NSM statistics is to know the total measurement distance of erosion and/or accretion, which may have high biological significance since this fine scale of analysis has not been conducted in this study area.

The DSAS analysis generates transects which are perpendicular to the reference user created baseline (Figure 3-4). The analysis explains that an intersection point is a cross between the casted transect and the shoreline boundary position for each specified year. The DSAS analysis then uses the distance, in meters, to conduct various calculations, which were previously described. Using the distance between transects, the DSAS can also generate forecasted transects for10- and/or 20-year projections.

The DSAS calculations require an operational workflow to gather and create the necessary components. The components needed are shoreline baselines, additional shorelines of interest (varying in different time periods), DSAS transects (which are cast some the baseline and intersect the additional shorelines positions), measurement distances, measurement points, and shoreline uncertainty. All objects used in the DSAS are stored in an ArcMap© Personal Geodatabase, as per USGS requirements for this analysis. The DSAS operational workflow includes the following steps: (1) Set default parameters and fields to created shoreline and baseline layers, transects, shoreline calculations, metadata and file output locations; (2) Cast transects and select their maximum search distance, transect spacing, and smoothing distance; (3) Calculate change statistics such as confidence intervals, shoreline intersection threshold, rate of output display, and summary report; (4) Create data visualization for LRR and NSM; and (5) Shoreline forecasting for a 10 and/or 20 year forecast.

DSAS Parameters and Selections

Selected NAIP Geotiff aerial imagery were in the Universal Transverse Mercator (UTM) coordinate system, Zone 17 North and in the 1983 North American Datum (NAD83) (Table 3-1). Using ESRI’s ArcCatalog© and ArcMap©, separate shapefiles for each aerial image’s shoreline was create, traced, and digitized. The scale used to digitize was 1:3,000, and a MWL was discerned by looking at the whitest and brightest part of the shoreline that was not influenced by the dark ocean color. Shorelines were then merged into a new single shapefile using the ArcMap© Merge tool. The ArcMap© Buffer tool was used to create a new shapefile that contained a 100-meter buffer around each shoreline in the new merged shorelines shapefile. A section of the buffer was selected to act as the baseline for transect casting for the DSAS calculations. The baseline selected can be found on the east side of Deer Island and is entirely inland. Both a baseline shapefile and merged shoreline shapefile are required for DSAS calculations (Figure 3-5, Inputs).

The DSAS parameters set for this analysis were a 20-meter transect spacing, a 2000-meter search distance for shorelines, and a smoothing distance of 500 meters. A 20-meter transect spacing was the minimum transect spacing allowed for the size of the study site. A 2000-meter search distance looked for shorelines 2000 meters way from the baseline. A smoothing distance is a user- specified smoothing value which can facilitate and orthogonal transect intersect by creating a baseline (which is not displayed in the final product). The intention of the smoothing distance is to prevent transects from intersecting with one another when there is a curve in the baseline. The larger the smoothing distance results in a longer reference line and produce more uniform transect orientations, which is recommended for smaller shorelines. The default setting for 90% confidence interval to calculate LRR and NSM rates remain unchanged.

Results

The calculations for the shoreline analysis are displayed in a colorblind colorramp. The LRR coloramp displays rates of change in m/yr. The NSM coloramp displays the distance of measurements in meters. The DSAS calculations follows the standard that a negative rate implies erosion and a positive rate implies accretion. The interpretation of the results go as follows.

Shoreline Analysis for Years 1994-2007

The Figure 3-6 displays the erosional transects calculated in the colors pink through black and accretional transects starting from orange to yellow colors. The LRR rates in Figure 3-7, A demonstrate more negative LRR rates calculated. The LRR rates range from the highest erosional; rate of -5.0 to -3.0 (m/yr) and the highest accretional rate of 3.0 to 4.0 (m/yr). The most frequent LRR rate range is -2.0 to -1.0 (m/yr) accounting for 30.5% of all transects calculated. The least frequent LRR rate range is the accretional rates of after 2.0 m/yr representing 0% of all transects. For the calculated NSM (Figure 3-6, B), the highest erosion distance measurements range from to -70.5 to -35.9 meters and the maximum accretion distance measurements range from 4.3 to 6.4 meters. The most frequent NSM measurement range is -10.0 to 2.2 meters accounting for 29.3% of all transects calculated (Figure 3-7, B). The least frequent NSM measurement range is the accretional distance between 4.3 to 6.4 meters accounting for 1.2% of all transects calculated (Figure 3-7, B).

Shoreline Analysis for Years 2010-2019

The Figure 3-8 displays the erosional transects calculated in the colors pink through black and accretional transects starting from orange to yellow colors. The LRR rates in Figure 3-9, A demonstrate more negative LRR rates calculated. The LRR rates range from the highest erosional; rate range of -4.0 to -3.0 (m/yr) and the highest accretional rate range of 1.0 to 2.0 (m/yr). The most frequent LRR rate range is -2.0 to -1.0 (m/yr) and -0.5 to 0.5 where both ranges equal 28% each for all transects calculated (Figure 3-9, A). The least frequent LRR rate range is the accretional rates of after 2.0 m/yr representing 0% of all transects. For the calculated NSM (Figure 3-9, B), the highest erosion distance measurements range from to -41.8.5 to -20.1 meters and the maximum accretion distance measurements range from 8.7.3 to 9.9 meters. The most frequent NSM distance is -6.7 to 2.9 meters accounting for 25.6% of all transects calculated (Figure 3-9, B). The least frequent NSM distance is the accretional measurement range between 8.7 to 9.9 meters accounting for 4.9% of all transects calculated (Figure 3-9, B).

Shoreline Analysis for Years 1994-2019

The results in Figure 3-10 include all the shorelines from Figures 3-8 and 3-9 for the LRR and NSM calculations. The erosion LRR rates (Figure 3-11, A) in this analysis range from the highest erosional rate of -4.0 to -3.0 (m/yr) to the highest accretional rate range from 3.0 to 4.0 (m/yr). The most frequent LRR rate range is -2.0 to -1.0 (m/yr) accounting for 39% of all transects calculated (n=82). The least frequent LRR rate range are the accretion rates greater than 1.0 (m/yr) accounting for 0% of all transects calculated. For the NSM calculations (Figure 3-11, B), the highest erosion distance measurements range from -91.8 to -68.5 meters accounting for all transects calculated. The NSM maximum accretion distance measurements range from 10.5 to 11 meters account for 1.2% of all transects calculated. The most frequent NSM distance measurement range is -45.7 to -22.9 meters accounting for 25.6% of all transects calculated. The largest erosion measurement distance is seen at the north end of Deer Island, while the middle has some areas of accretion and less erosion. The south end of Deer Island has some acute peaks of erosion, however not as high as the north end.

Table 3-3 concludes the DSAS NSM results calculated 81.70% of all transects for this date range were a negative distance in comparison to the baseline for years 1994 to 2019. This only leaves 18.29% left of positive distance for the calculated transects. The average distance calculated for each transect is -29.1 meters (taking into account positive and negative transects). The average of all negative distances is -36.83 meters. The maximum negative distance measured is -91.71 meters while the maximum positive distance calculated is 10.91 meters.

Table 3-4 concludes the DSAS LRR calculations for years 1994 to 2019. The average LRR rate calculated is -0.95 m/yr. The percent of all transects that are erosional is 76.83% (n= 63). The maximum value of erosion calculated is -3.32 m/yr while the maximum value for accretion is 0.62 m/yr. The average of all erosional rates is -1.33 m/yr. The percent of transects that have statistically significant erosion is 69.51% while the percent of all transects that have statistically significant accretion is 10.98%.

Shoreline Analysis for 10 and 20-Year Prediction

The 10-year prediction (Figure 3-11, A) demonstrates a uniformity of erosion particularity in the south and center of Deer Island. The north end of Deer Island has an area south of the shoreline bulge that is projected to be eroded by the 10-year prediction. The center of Deer Island has a predicted area of accretion, but the majority of the 10-year projection is predicting that the west shoreline of Deer Island may erode. The 20-year prediction (Figure 3-11, B) is very similar to the 10-year prediction model, but with more drastic erosion in the north and south end. A summary of the 10- and 20- year shoreline predictions suggest that erosion may occur on the north and south ends of the western shoreline of Deer Island with some accretion located around the center of the island.

Imagery Observation

Observing the NAIP aerial photographs, in the same scale (1:3,000) and the same projection (NAD 1983 UTM, Zone 17N) also displays results. Figure 3-13, A consists of the 1994 aerial imagery and was taken on a day with a max wind speed of 19.31 KPH and a median river discharge of 274.9 (m^3/s) (Table 3-1). Figure 3-13, B consists of the 2019 aerial imagery aerial imagery and was taken on a day with a max wind speed of 11.27 KPH and a median river discharge of 146.9 (m^3/s) (Table 3-1). In the 2019 imagery reveals that the south western creek has a larger open mouth to the sea compared to the 1994 imagery. The north eastern bulge that looks round and full in 1994 imagery looks to be pointy and thin in 2019 imagery. The DSAS shoreline results are a more accurate representation of the erosion and accretion on Deer Island, but even looking at aerial evidence it does appear that there are most instances of erosion than accretion.

Discussion

Results in this analysis suggest that more shoreline erosion has occurred along the Deer Island shoreline than accretion. The DSAS results depict consistent erosional transects along all periods analyzed. The negative NSM range for the time period 1994- 2007 (Figure 3-7, B) is almost twice as negative as the time period 2010-2019 (Figure 3-9, B). The time period of 2010-2019 also has a greater percentage of NSM accretion (Figure 3-9, B) compared to the 1994- 2007 time period (Figure 3-7, B). Comparatively, LRR m/yr rates for the time periods display overall erosional results with some differences. Time period 1994-2007 (Figure 3-7, A) displays an LRR rate range of -0.3 < to > -2.0 that is two times the percentage of transects in comparison to that rate range in time period 2010-2019 (Figure 3-9, A). It is interesting to think about how and why this these rates and distance measurements are not the same between the time periods analyzed. A year prior to the first imagery,1993, in the time series the Storm of the Century hit the Big Bend region. There is evidence during this storm event that sandy coasts were susceptible to shoreline erosion (Goodbred & Hine, 1993). Despite the Storm of the Century happening prior to our shoreline analysis, abrupt shoreline changes due to an intense weather event coupled with SLR may have triggered an unbalance of natural erosion and accretion rates on Deer Island during the 1994- 2007-time frame, especially considering storm clusters encompassed this time frame.

Despite analyzing such brief time periods on a small shoreline, many changes have occurred. Note that in 1994 (Figure 3-13, A), a bulge like feature (on the northeast side) looks to be mostly eroded by 2019 (Figure 3-13, B). Even erosion of small features such as that bulge can make an impact on the available habitats for animals. Many species depend on shorelines for food, nesting, and shelter (O'Connell et al., 2005). Shorebirds rely on shorelines for feeding habitats during migration in the winter months. Habitat loss, due to erosion, limits the availability of food and resources for these shorebirds, possibly resulting in increased competition. This increased competition may exclude individuals from a foraging site, increase mortality rates for these excluded shorebirds, and ultimately lead to limitations in numbers (Galbraith et al., 2005). The Big Bend region of Florida is already experiencing low shorebird species richness and population abundance, implying that an area already struggling with species biodiversity, despite the lack of human impact, will have negative shorebird impacts because of consistently eroding shorelines. During a high erosion storm event, many sandy-shore animals may also be washed up to shore, stranded up shore, or left to die due to exposure. Sandy- shore creatures naturally are able to survive storm events due to their defense mechanisms but are not always able to survive in the event of significant shoreline erosion (Brown & McLachlan, 2002). Whether shorelines erode slowly, but constantly, or in a storm event, extreme shoreline erosion negatively impacts animal species.

It is interesting to note, that although the overall shoreline experienced erosion, there is evidence accretion might have occurred in the middle of the shoreline during the entire time period analyzed (Table 3-3). During this time period the maximum positive distance gained was 10.91 meters occurring around the central shoreline on transect 44 (Table 3-3). Accretion for our study site may come from intense meteorological events since there is a scant supply of sand being dispersed by the Suwannee River (Goodbred et al., 1998). However, it is unclear how much accretion can occur with the perpetuation of sea- level rise consistently stressing the sandy shoreline substrate. Sea-level rise has the second greatest effect on shoreline change on the east coast of Florida but has very similar effects on the west coast as well. There is a possibility for Florida to provide beach nourishment to areas where erosion is evident. Currently there is no schedule to provide beach nourishment to our study site.

During this analysis, the main source of error arises with the missing imagery years 2007- 2008 and 2011- 2012. If those missing years were available for analysis, it would provide a closer interpretation of the true erosion differences between the two 12 to 13-year time periods. Since our study site is uninhabited, remote and not a tourist destination, it is not surprising to see that NAIP is not contracted to fly over this area every year. Another source for possible errors is the individual digitization of each shoreline due to user errors. Since the available imagery was used to digitize the years’ shoreline, the digitization of each shoreline may differ from user to user. In this study only one person digitized each shoreline to reduce this error. The resolution of each image was at least 1-meter resolution, which may be considered “high” resolution in comparison to 30-meter resolution from Landsat 7 and 8 (Fisher et al., 2018), where Landsat imagery may also be used for analysis. The higher the resolution is, the more likely the digitized shorelines are accurate.

The prediction models are based on a linear regression rate calculated by DSAS, termed Kalman filter (Kalman, 1960). The Kalman filter conducts an analysis to minimize the error between the observed and modeled shoreline position to develop the forecast where the rate and uncertainties are considered (Long & Plant, 2012). Our prediction models project that more shoreline erosion is expected (Figure 3-12). This can concerning since currently our study site is not impacted by human development, however that may change in the future since the human population is expected to increase along the surrounding counties (Figure 3-1). Continuing the practice of excluding development is wise given the instability and high erosion rates on the island. The prediction models may be used as a reliable source of information for land management directors who seek to protect uninhabited shorelines along the Big Bend.

This study has revealed brief historical trends of coastal evolution along an undeveloped sandy shoreline. The shoreline statistics revealed greater meters of erosion during the first-time frame 1994-2007, possibly due to a major hurricane impact. Storm and storm clusters may significantly impact barrier island morphology in this area with direct storm wind and tide impacts along with changes to wave energy. Changes in the frequency of storms may contribute to numerous changes to the low energy waves described in our study site. Long term sea-level rise and sediment supply are considered major factors that stimulate shoreline erosion and/or accretion (Sankar et al., 2018). This research has demonstrated that sandy shorelines in this area may be susceptible to more erosion than accretion due to the factors mentioned, which may ultimately lead to multidecadal shoreline loss.

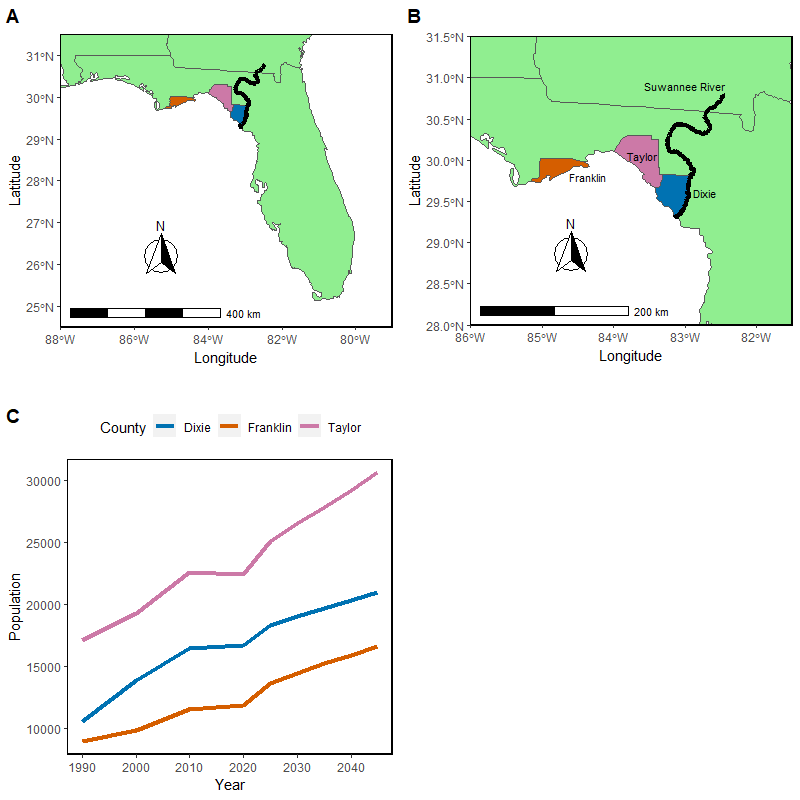


Figure 3-1. A) Map of Florida with Dixie, Franklin and Taylor counties identified along side the Suwannee River; B) Zoomed in map of study area with Dixie, Franklin and Taylor counties identified along side the Suwannee River; C) Generated figure based on census and projection human population data of Dixie, Franklin and Taylor counties 1990-2045 (Bureau of Economic and Business Research, 2021)



Figure 3-2. Island degradation of Derrick Key in the Cedar Keys, Florida from 1982 (left) to 2016 (right), (Vitale, 2019).



Figure 3-3. Location of Deer Island, Florida. A) Map of the entire state of Florida; B) Zoomed in study site; C) Zoomed into map scale of 0.03 to Deer Island with a scale bar in kilometers. Shoreline shapefile downloaded at my.fwc.com, (Original source map scale 1: 2,000,000 Scale, and digitized in 2017).

Table 3-1. Table of metadata for each aerial image used in this analysis including date, median river discharge, observed weather, and additional imagery metadata. River discharge information is calculated by data from <https://tidesandcurrents.noaa.gov/> at Cedar Key, Florida Station 8727520, and observed weather provided by [www.wunderground.com](http://www.wunderground.com). Imagery metadata are provided by USGS Earth Explorer, <https://earthexplorer.usgs.gov/>.

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Median River Discharge (m^3/s)  Station ID= 02323500 | Observed weather | Metadata (USGS Earth Explorer) |
| January 20, 1994 | Value= 274.9 | Avg Temp (C)- 3.41 Precipitation (cm)- 0.00  Max Wind Speed (KPH)- 19.31 | Entity ID: DI00000000018672 (found in DOQ)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 3  Sensor Type RGB |
| December 30, 1998 | Value= 180.37 | Avg Temp (C)- 9.30 Precipitation (cm)- 0.00  Max Wind Speed (KPH)- 25.75 | Entity ID: DI00000001164809 (found in DOQ)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 3  Sensor Type RGB |
| November 02, 2007 | Value= 66.5 | Avg Temp (C)- 19.31  Precipitation (cm)- 0.00  Max Wind Speed (KPH)- 22.53 | Entity ID: N\_2908356\_NW\_17\_1\_20071102 (found in NAIP)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 3  Sensor Type CIR |
| September 19, 2010 | Value= 120.0 | Avg Temp (C)- 25.32 Precipitation (cm)- 0.00  Max Wind Speed (KPH)-24.14 | Entity ID: M\_2908356\_NW\_17\_1\_20100919 (found in NAIP)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 4  Sensor Type CNIR |

Table 3-1. Continued

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Median River Discharge (m^3/s)  Station ID= 02323500 | Observed weather | Metadata (USGS Earth Explorer) |
| October 13, 2013 | Value= 232.2 | Avg Temp (C)- 22.13 Precipitation (cm)- 0.00  Max Wind Speed (MPH)- 10 | Entity ID: M\_2908356\_NW\_17\_1\_20131013 (found in NAIP)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 4  Sensor Type CNIR |
| November 12, 2015 | Value= 171.9 | Avg Temp (C)- 19.27 Precipitation (cm)- 0.00  Max Wind Speed (KPH)- 16.09 | Entity ID: M\_2908356\_NW\_17\_1\_20151112 (found in NAIP)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 4  Sensor Type CNIR |
| October 26, 2017 | Value= 226.3 | Avg Temp (C)- 12.60 Precipitation (cm)- 0.00  Max Wind Speed (KPH)- 14.48 | Entity ID: M\_2908356\_NW\_17\_1\_20171026 (found in NAIP)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 1.0  Units METER  Number of Bands 4  Sensor Type CNIR |
| November 10, 2019 | Value = 146.9 | Avg Temp (C)- 14.12 Precipitation (cm)- 0.00  Max Wind Speed (KPH)- 11.27 | Entity ID: M\_2908356\_NW\_17\_060\_20191110 (found in NAIP)  Map Projection UTM  Projection Zone 17N  Datum NAD83  Resolution 0.60  Units METER  Number of Bands 4  Sensor Type CNIR |

Table 3-2. National Agriculture Imagery Program (NAIP) aerial imagery band wavelength ranges in units (µm) (<https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/APFO/support-documents/pdfs/fourband_infosheet_2017.pdf>)

|  |  |  |
| --- | --- | --- |
| Sensor Type | Color and wavelength (µm) | Band and color channel to display true color |
| RGB | Blue 400–500  Green 500–600  Red 600–700 | 1 – Red channel  2 – Green channel  3 – Blue channel |
| CIR/ CNIR | Blue 400–500  Green 500–600  Red 600–700  Near Infrared 800–900 | 1 – Red channel  2 – Green channel  3 – Blue channel  4 – Near Infrared (not show on screen display) |



Figure 3-4. Example of DSAS transect casting (<https://www.usgs.gov/centers/whcmsc/science/digital-shoreline-analysis-system-dsas?qt-science_center_objects=0#qt-science_center_objects>)



Figure 3-5. DSAS components and operational workflow (“Digital Shoreline Analysis System (DSAS) Version 5.0 User Guide.”, 2021).

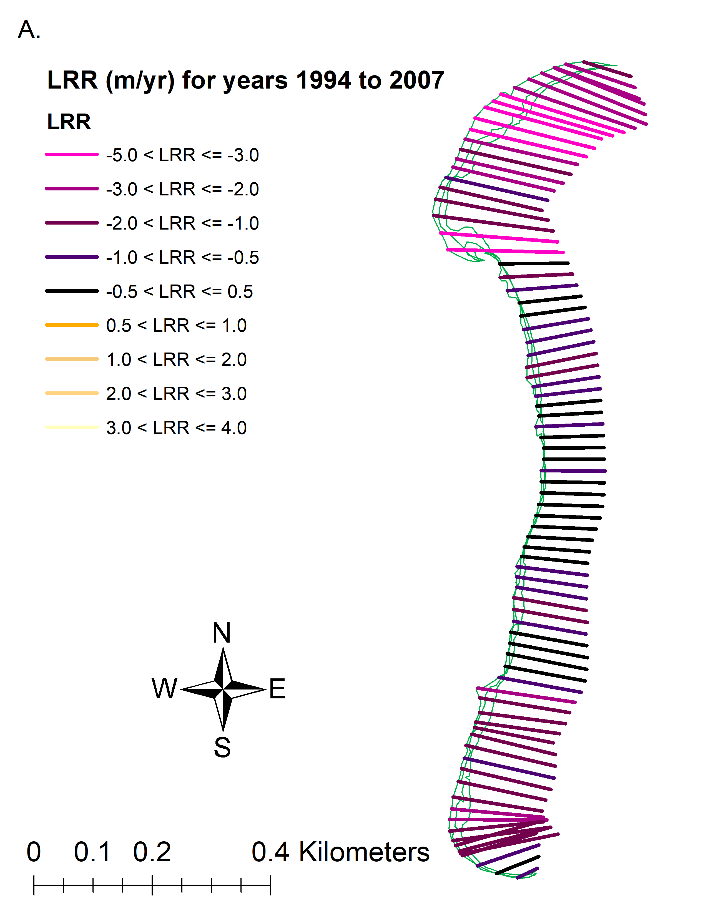
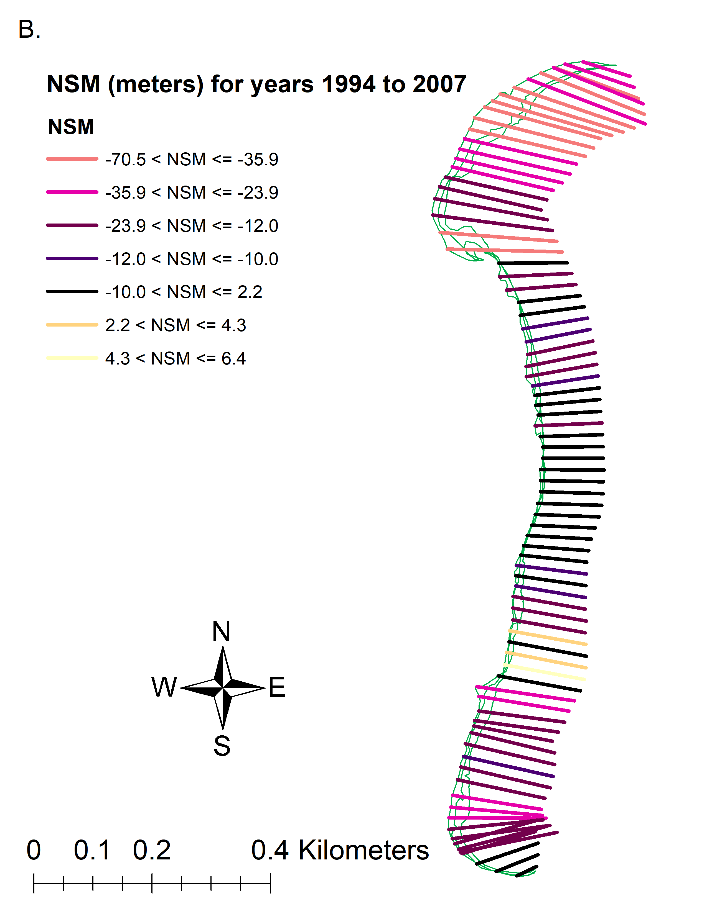


Figure 3-6. Results of DSAS calculations for years 1994 to 2007. A) Linear Regression Rates model (in m/yr) displaying transects which colors correspond to Figure 3-7, A. The transects (n= 82) display where along the shoreline erosion and accretion has been detected. B) Net Shoreline model (in meters) displaying transects which colors correspond to Figure 3-7, B. The transects display where along the shoreline erosion and accretion has been detected. The shorelines (green) include digitized shorelines for years 1994 to 2007 in this figure.

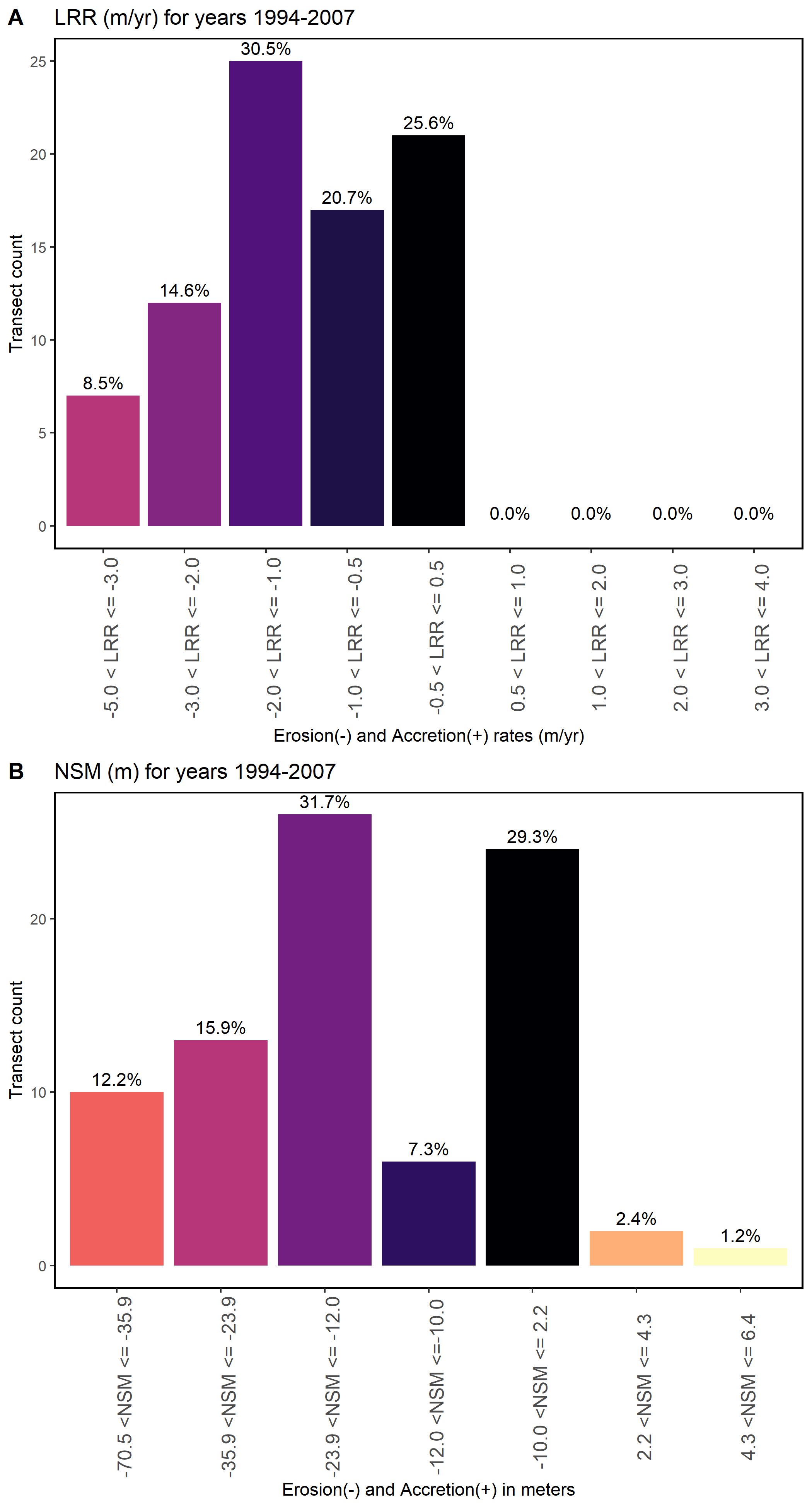


Figure 3-7. Figure of the DSAS statistics for years 1994 to 2007, A) LRR results for the transects counted of each of the bins calculated for erosion and accretion in meters, the negative x-values are erosion in meters and the positive x-values are accretion, B) NSM results for the transects counted of each of the bins calculated for erosion and accretion in meters, the negative x-values are erosion in meters and the positive x-values are accretion. The total amount of transects calculated by DSAS is 82.

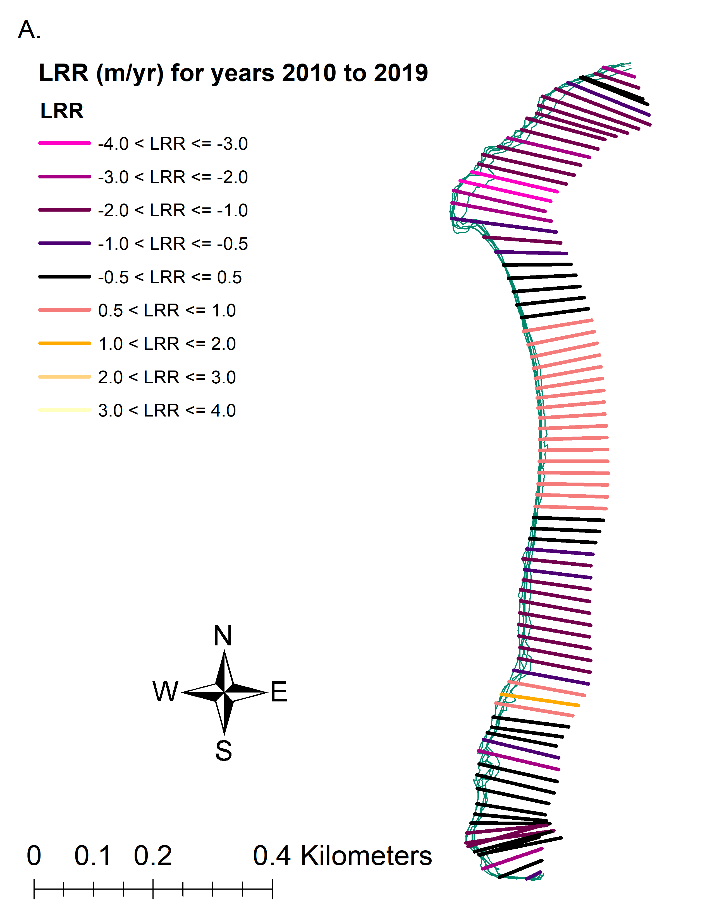
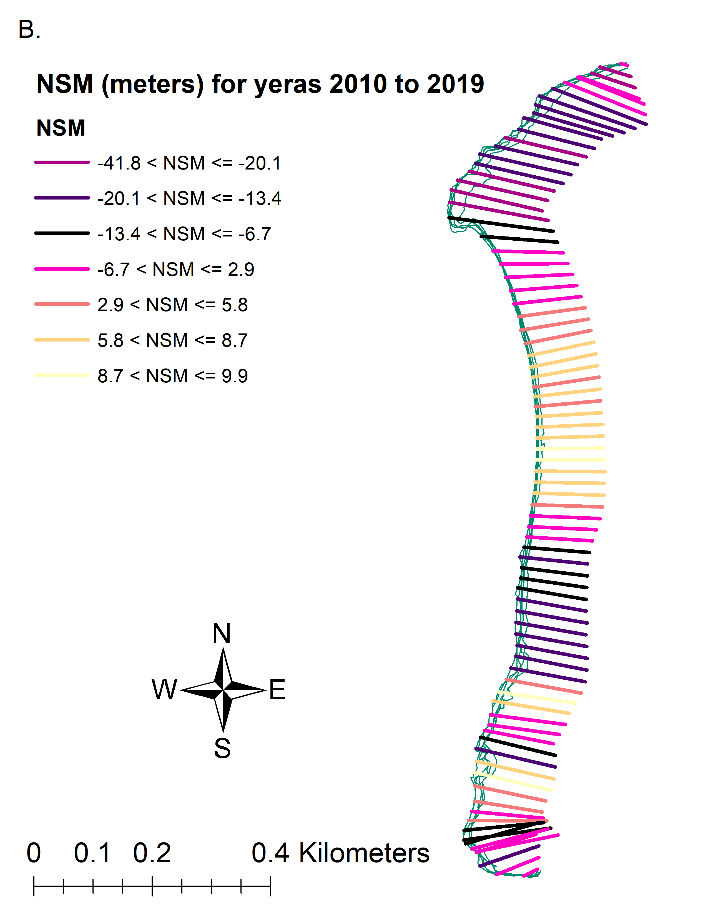


Figure 3-8. Results of DSAS calculations for years 2010 to 2019. A) Linear Regression Rates model (in m/yr) displaying transects which colors correspond to Figure 3-9, A. The transects (n= 82) display where along the shoreline erosion and accretion has been detected. B) Net Shoreline model (in meters) displaying transects which colors correspond to Figure 3-9, B. The transects display where along the shoreline erosion and accretion has been detected. The shorelines (green) include digitized shorelines for years 2010 to 2019 in this figure.

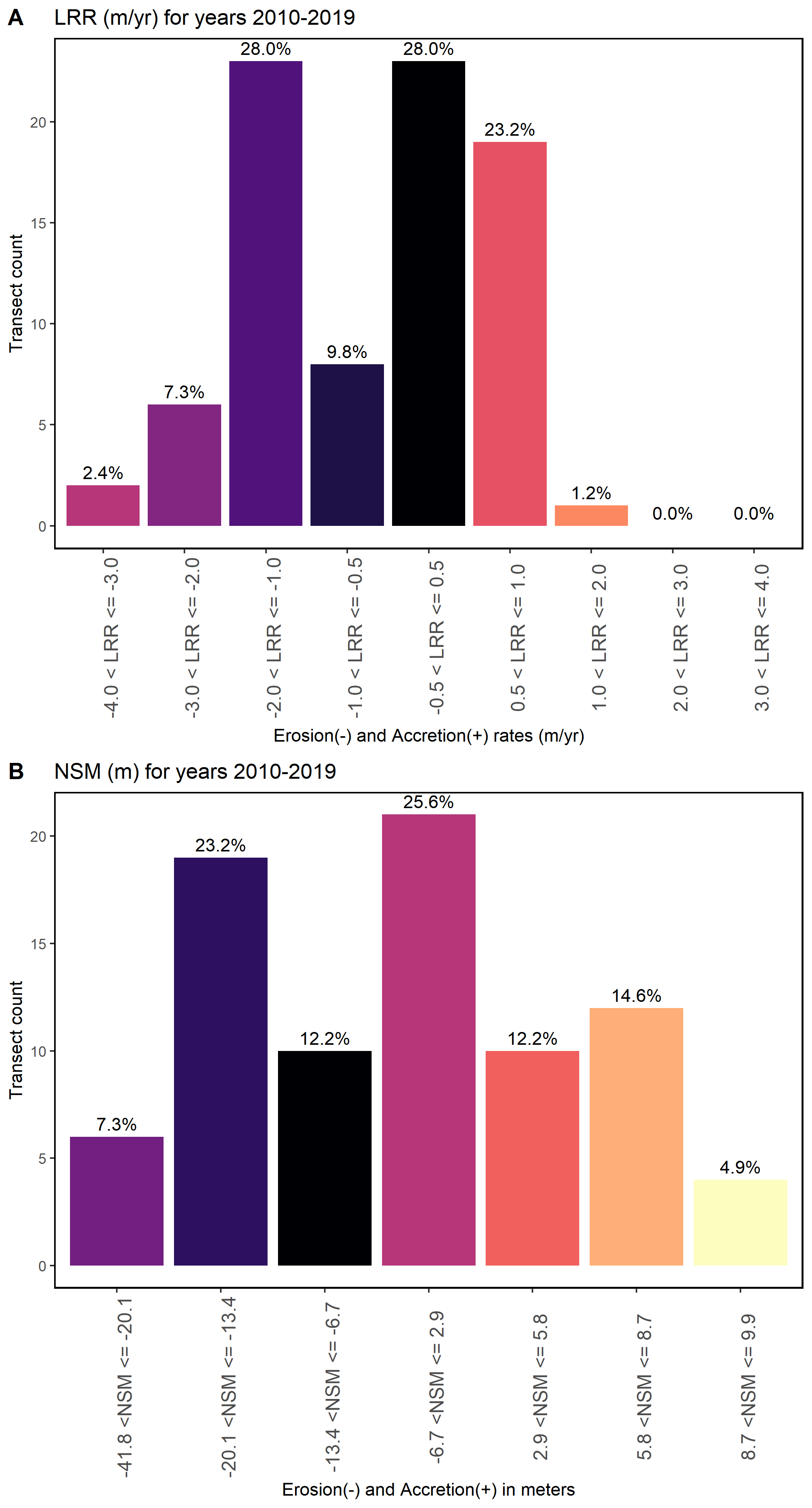


Figure 3-9. Figure of the DSAS statistics for years 2010 to 2019, A) LRR results for the transects counted of each of the bins calculated for erosion and accretion in meters, the negative x-values are erosion in meters and the positive x-values are accretion, B) NSM results for the transects counted of each of the bins calculated for erosion and accretion in meters, the negative x-values are erosion in meters and the positive x-values are accretion. The total amount of transects calculated by DSAS is 82.

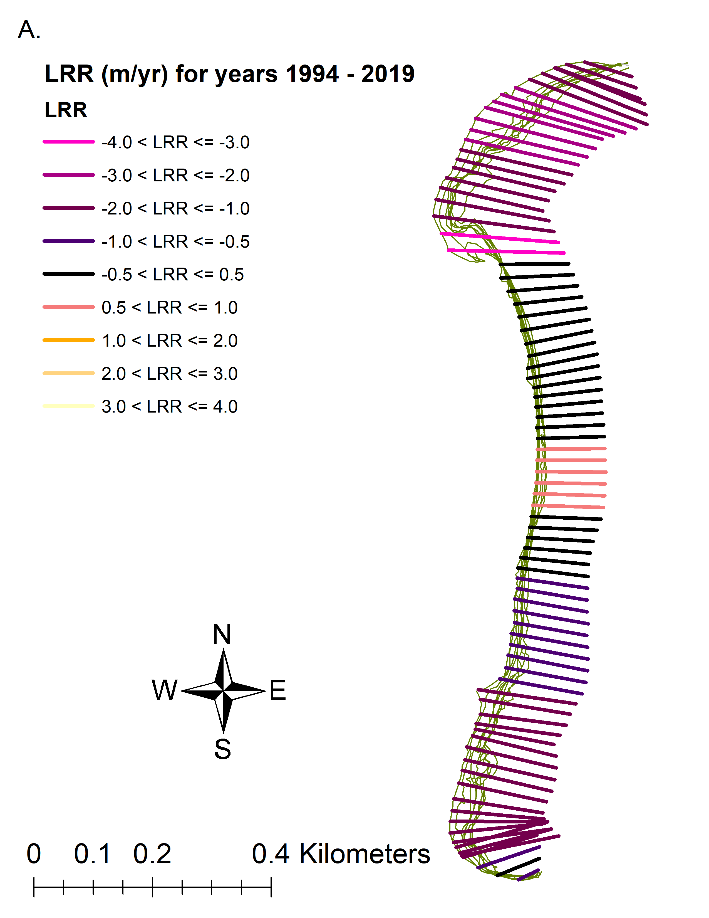
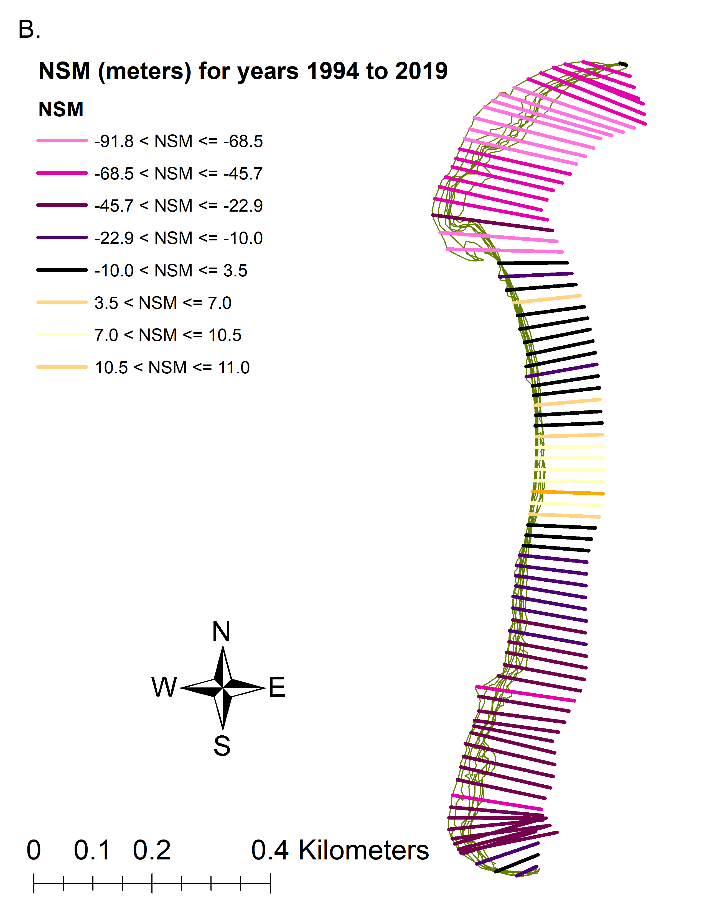


Figure 3-10. Results of DSAS calculations for years 1994 to 2019. A) Linear Regression Rates model (in m/yr) displaying transects which colors correspond to Figure 3-11, A. The transects (n= 82) display where along the shoreline erosion and accretion has been detected. B) Net Shoreline model (in meters) displaying transects which colors correspond to Figure 3-11, B. The transects display where along the shoreline erosion and accretion has been detected. The shorelines (green) include all digitized shorelines in this figure.

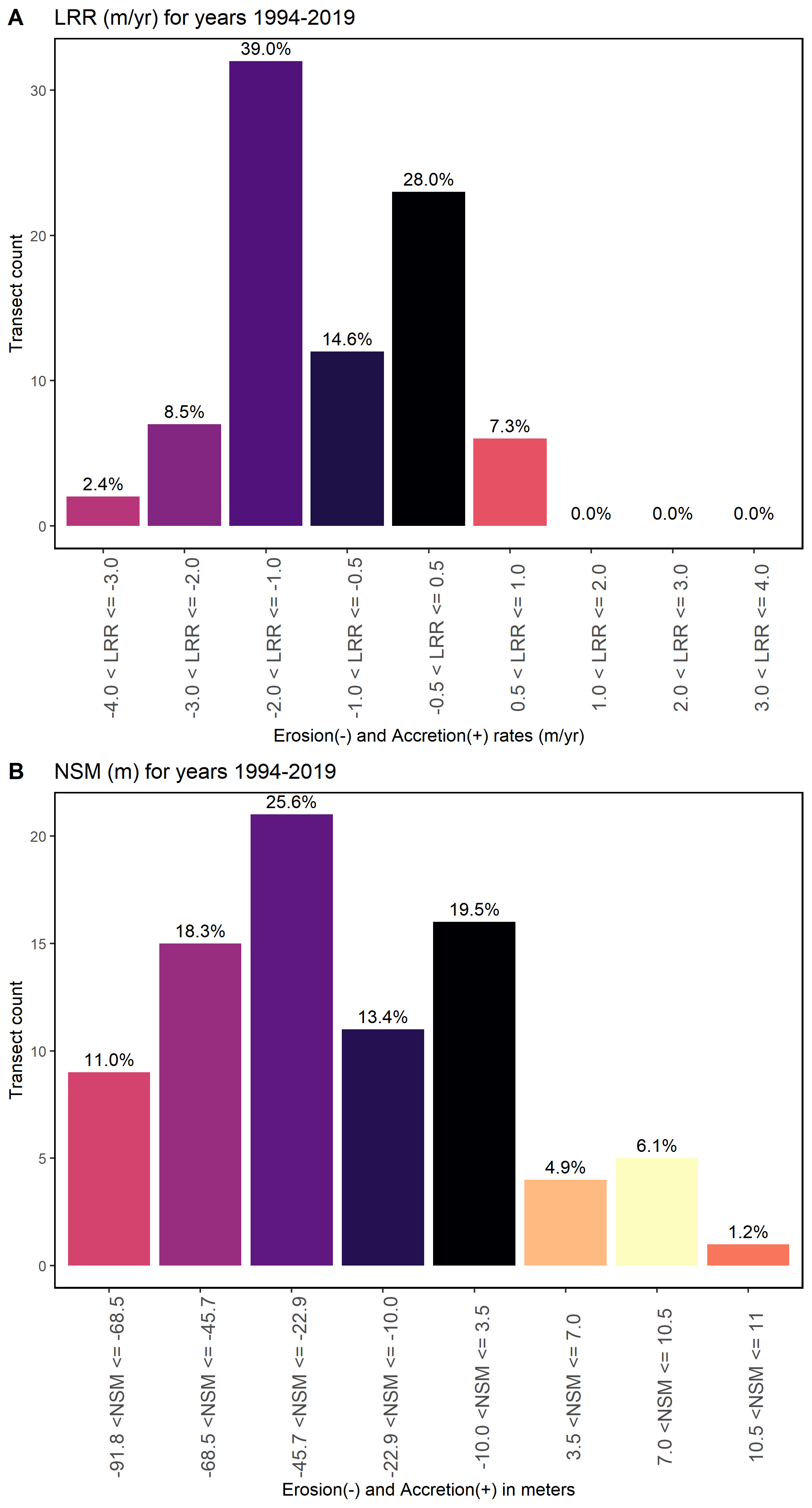


Figure 3-11. Figure of the DSAS statistics for years 1994 to 2019, A) LRR results for the transects counted of each of the bins calculated for erosion and accretion in meters, the negative x-values are erosion in meters and the positive x-values are accretion, B) NSM results for the transects counted of each of the bins calculated for erosion and accretion in meters, the negative x-values are erosion in meters and the positive x-values are accretion. The total amount of transects calculated by DSAS is 82.

Table 3-3. Summary statistics calculated by DSAS, Distance: NSM (Net Shoreline Movement)

|  |  |
| --- | --- |
| Summary Statistic | Value |
| Total number of transects (counts) | 82 |
| Average distance (meters) | -29.1 |
| Number of transects with negative distance (counts) | 67 |
| Percent of all transects that have a negative distance | 81.70% |
| Maximum negative distance (meters) | -91.71 |
| Maximum negative distance (transect ID #) | 12 |
| Average of all negative distances (meters) | -36.83 |
| Number of transects with positive distance | 15 |
| Percent of all transects that have a positive distance | 18.29% |
| Maximum positive distance (meters) | 10.91 |
| Maximum positive distance (transect ID #) | 44 |

Table 3-4. Summary statistics calculated by DSAS, RATE: LRR (Linear Regression Rate)

|  |  |
| --- | --- |
| Summary Statistic | Value |
| Total number of transects (counts) | 82 |
| Average rate (m/yr) | -0.95 |
| Average of the confidence intervals associated with rates | 0.49 |
| Reduced n (number of independent transects) | 900.00% |
| Uncertainty of the average rate using reduced n | 0.17 |
| Average rate with reduced n uncertainty (m/yr) | -0.95 +/- 0.17 |
| Number of erosional transects (counts) | 63 |
| Percent of all transects that are erosional | 76.83% |
| Percent of all transects that have statistically significant erosion | 69.51% |
| Maximum value erosion (m/yr) | -3.32 |
| Maximum value erosion (transect ID #) | 22 |
| Average of all erosional rates (m/yr) | -1.33 |
| Number of accretional transects (counts) | 19 |
| Percent of all transects that are accretional | 23.17% |
| Percent of all transects that have statistically significant accretion | 10.98% |
| Maximum value accretion (m/yr) | 0.62 |
| Maximum value accretion (transect ID #) | 44 |
| Average of all accretional rates (m/yr) | 0.31 |

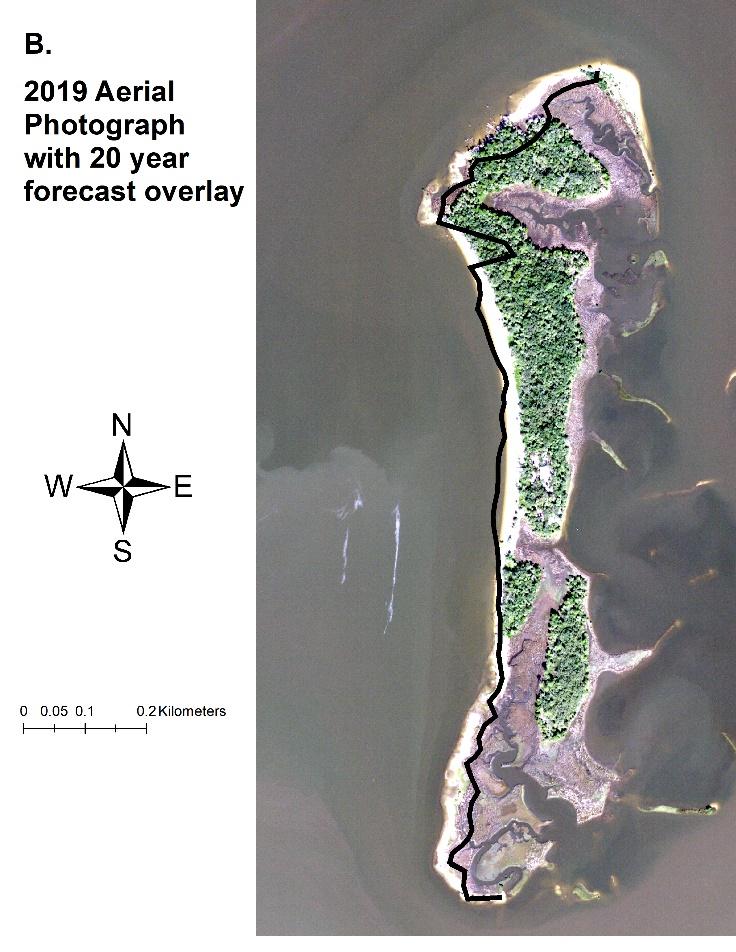
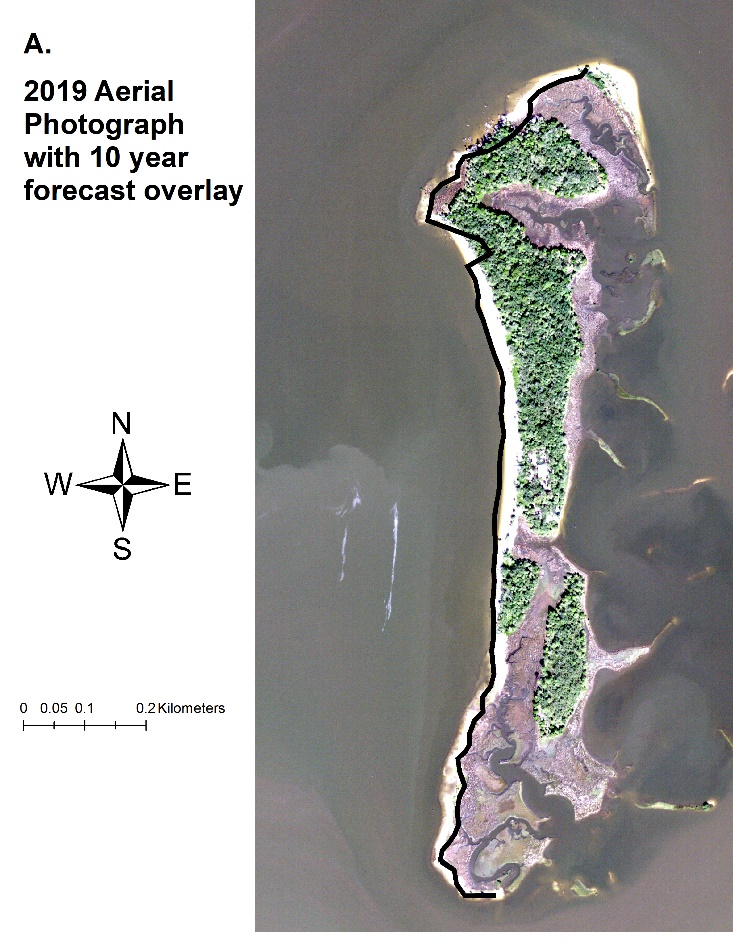
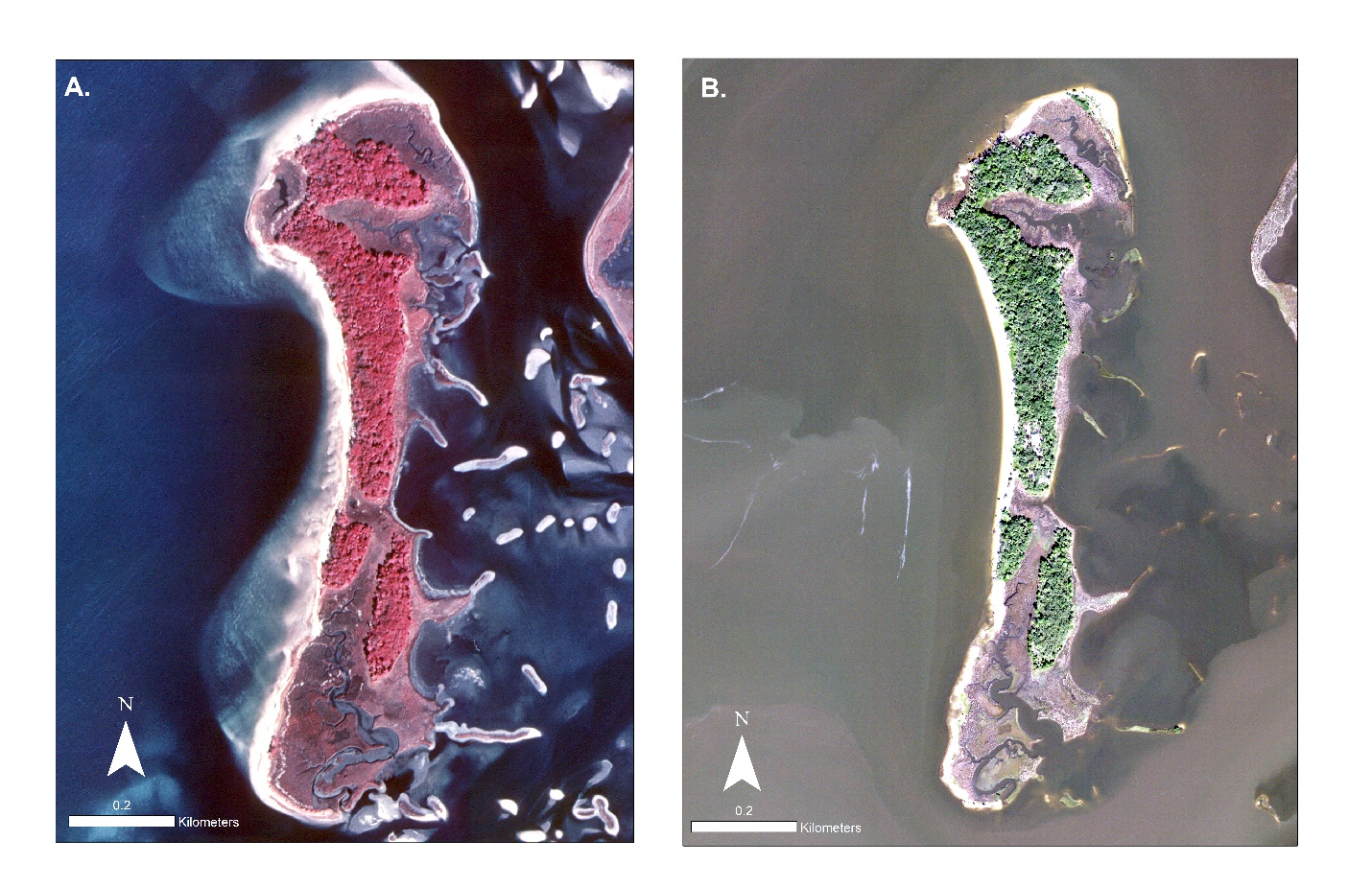


Figure 3-12. DSAS shoreline prediction forecast. A) Shoreline forecast for a 10-year prediction (thick black line) overlayed aerial imagery (2019) to display the predicted shoreline loss in comparison to the latest imagery selected. B) Shoreline forecast for a 20-year prediction (thick black line) overlayed aerial imagery (2019) to display the predicted shoreline loss in comparison to the latest imagery selected. Shorelines are located on the west side of each panel.

Figure 3-13. NAIP aerial imagery used for analysis. A) Imagery from 1994 (scale 1:3,000); B) Imagery from 2019 (scale 1:3,000)

CHAPTER 4

Synthesis

Data management workflows are imperative to every ecological project. Ecological projects may be prone to mismanaged data and files due to the nature that every effort is different because of their hypothesis, instruments and data collectors.

-improves efficiencies

-reproducibility

-transparency

LIST OF REFERENCES

“A Deer Island A Sportsman's Paradise.” [Https://images1.Loopnet.com/d2/Z4L1-alqEsAlhPT\_YJ25N8OMkXU3L\_mAPAZYXiq2OVg/Document.pdf. (Accessed](Https://images1.Loopnet.com/d2/Z4L1-alqEsAlhPT_YJ25N8OMkXU3L_mAPAZYXiq2OVg/Document.pdf.%20(Accessed) January 2021)

Archmiller, A. A., Johnson, A. D., Nolan, J., Edwards, M., Elliott, L. H., Ferguson, J. M., Iannarilli, F., Vélez, J., Vitense, K., Johnson, D. H., & Fieberg, J. (2020). Computational Reproducibility in The Wildlife Society’s Flagship Journals. Journal of Wildlife Management, 84(5), 1012–1017. https://doi.org/10.1002/jwmg.21855

Barchard, K. A., & Pace, L. A. (2011). Preventing human error: The impact of data entry methods on data accuracy and statistical results. Computers in Human Behavior, 27(5), 1834–1839. https://doi.org/10.1016/j.chb.2011.04.004

Barone, L., Williams, J., & Micklos, D. (2017). Unmet Needs for Analyzing Biological Big Data: A Survey of 704 NSF Principal Investigators. https://doi.org/10.1101/108555

Blischak, J. D., Davenport, E. R., & Wilson, G. (2016). A Quick Introduction to Version Control with Git and GitHub. PLoS Computational Biology, 12(1). https://doi.org/10.1371/journal.pcbi.1004668

Brown, A. C., & McLachlan, A. (2002). Sandy shore ecosystems and the threats facing them: Some predictions for the year 2025. Environmental Conservation, 29(1), 62–77. <https://doi.org/10.1017/S037689290200005X>

Bureau of Economic and Business Research, Population Studies Program, [www.bebr.ufl.edu/population](http://www.bebr.ufl.edu/population). (Accessed January 28, 2021)

Cazenave, A., & Cozannet, G. le. (2014). Sea level rise and its coastal impacts. Earth’s Future, 2(2), 15–34. https://doi.org/10.1002/2013ef000188

Czech, B., Krausman, P. R., & Devers, P. K. (2000). Economic associations among causes of species endangerment in the United States: associations among causes of species endangerment in the United States reflect the integration of economic sectors, supporting the theory and evidence that economic growth proceeds at the competitive exclusion of nonhuman species in the aggregate. BioScience, 50(7), 593-601.

Desantis, L. R. G., Bhotika, S., Williams, K., & Putz, F. E. (2007). Sea-level rise and drought interactions accelerate forest decline on the Gulf Coast of Florida, USA. Global Change Biology, 13(11), 2349–2360. <https://doi.org/10.1111/j.1365-2486.2007.01440.x>

“Deer Island Florida, United States.” Deer Island - Florida, United States - Private Islands for Sale, [www.privateislandsonline.com/united-states/florida/deer-island](http://www.privateislandsonline.com/united-states/florida/deer-island). (Accessed on January 2021)

“Digital Shoreline Analysis System (DSAS) Version 5.0 User Guide.” <Https://Pubs.usgs.gov/of/2018/1179/ofr20181179.Pdf>. (Accessed January 2021)

Finkl, C. W., & Charlier, R. H. (2003). Sustainability of Subtropical Coastal Zones in Southeastern Florida: Challenges for Urbanized Coastal Environments Threatened by Development, Pollution, Water Supply, and Storm Hazards. Journal of Coastal Research, 19(4), 934–943.

Fisher, J. R. B., Acosta, E. A., Dennedy-Frank, P. J., Kroeger, T., & Boucher, T. M. (2018). Impact of satellite imagery spatial resolution on land use classification accuracy and modeled water quality. Remote Sensing in Ecology and Conservation, 4(2), 137–149. https://doi.org/10.1002/rse2.61

Galbraith, H., Jones, R., Park, R., & Herrod-Julius. (2005). Global Climate Change and Sea Level Rise: Potential Losses of Intertidal Habitat for Shorebirds. https://doi.org/10.1675/1524

Geselbracht, L., Freeman, K., Kelly, E., Gordon, D. R., & Putz, F. E. (2011). Retrospective and prospective model simulations of sea level rise impacts on Gulf of Mexico coastal marshes and forests in Waccasassa Bay, Florida. Climatic Change, 107(1), 35–57. https://doi.org/10.1007/s10584-011-0084-y

Gilroy, S. P., & Kaplan, B. A. (2019). Furthering open science in behavior analysis: An introduction and tutorial for using GitHub in research. Perspectives on behavior science, 42(3), 565-581.

Goodbred, S. L., Hine, A. C., & Stumpf, R. (1993). Sediment distribution patterns and the development of the marsh system rimming a shallow-water shelf embayment: Waccasassa Bay, Levy Co., FL. Geological Society of America, Abstracts with Programs;(United States), 25(CONF-9304188--).

Goodbred, S. L., Wright, E. E., & Hine, A. C. (1998). Sea-Level Change And Storm-Surge Deposition In A Late Holocene Florida Salt Marsh. In JOURNAL OF SEDIMENTARY RESEARCH (Vol. 68, Issue 2). http://pubs.geoscienceworld.org/sepm/jsedres/article-pdf/68/2/240/2812177/240.pdf

Hine, A. C., Belknap, D. F., Hutton, J. G., Osking, E. B., & Evans, M. W. (1988). Recent geological history and modern sedimentary processes along an incipient, low-energy, epicontinental-sea coastline; Northwest Florida. Journal of Sedimentary Research, 58(4), 567-579.

Holling, C. S. (1978). Adaptive environmental assessment and management. John Wiley & Sons.

Houston, J. R. (2015). Shoreline Response to Sea-Level Rise on the Southwest Coast of Florida. Journal of Coastal Research, 314, 777–789. <https://doi.org/10.2112/jcoastres-d-14-00161.1>

Jackson, N. L., Nordstrom, K. F., Eliot, I., & Masselink, G. (2002). ‘Low energy’sandy beaches in marine and estuarine environments: a review. Geomorphology, 48(1-3), 147-162.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems.

Lefcheck, J. S. (2016). PiecewiseSEM: Piecewise Structural Equation Modelling In R For Ecology, Evolution, And Systematics. Methods in Ecology and Evolution, 7(5), 573–579. https://doi.org/10.1111/2041-210X.12512

Li, W., & Gong, P. (2016). Continuous monitoring of coastline dynamics in western Florida with a 30-year time series of Landsat imagery. Remote Sensing of Environment, 179, 196–209. https://doi.org/10.1016/j.rse.2016.03.031

Long, J. W., & Plant, N. G. (2012). Extended Kalman Filter framework for forecasting shoreline evolution. Geophysical Research Letters, 39(13). https://doi.org/10.1029/2012GL052180

Lowndes, J. S. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O’Hara, C. C., Jiang, N., & Halpern, B. S. (2017). Our path to better science in less time using open data science tools. Nature Ecology and Evolution, 1(6). https://doi.org/10.1038/s41559-017-0160

Martinelli, M., Guicciardi, S., Penna, P., Belardinelli, A., Croci, C., Domenichetti, F., Santojanni, A., & Sparnocchia, S. (2016). Evaluation of the oceanographic measurement accuracy of different commercial sensors to be used on fishing gears. Ocean Engineering, 111, 22–33. https://doi.org/10.1016/j.oceaneng.2015.10.037

McLachlan, A. (1990). Sandy beach ecosystems, [in:] Ecology of sandy shores.

McKinney, M. L. (2006). Urbanization as a major cause of biotic homogenization. Biological Conservation, 127(3), 247–260. <https://doi.org/10.1016/j.biocon.2005.09.005>

Moreno, M. Pine, W.E. Aufmuth, J. Maxwell, D. & Smith, P. (2020). GitHub Workflow for the LCR Oyster Project. Zenodo. <http://doi.org/10.5281/zenodo.4319177>

Moreno, M. Aufmuth, J. Maxwell, D. Phillips, R. & Pine, W.E (2020). MYSQL workflow for the LCR Oyster Project. Zenodo. http://doi.org/10.5281/zenodo.4319191

Moreno, M. Coleman, T. Aufmuth, J. Maxwell, D. & Pine, W.E (2020). Data Packet Structure for the LCR Oyster Project Zenodo. http://doi.org/10.5281/zenodo.4319175

Morton, R. A., Miller, T. L., & Moore, L. J. (2004). National Assessment of Shoreline Change: Part 1 Historical Shoreline Changes and Associated Coastal Land Loss Along The U.S. Gulf Of Mexico.

Nelson, J. R., & Grubesic, T. H. (2018). The implications of oil exploration off the Gulf Coast of Florida. Journal of Marine Science and Engineering, 6(2). https://doi.org/10.3390/jmse6020030

Nie, M. A., & Schultz, C. A. (2012). Decision-Making Triggers in Adaptive Management. Conservation Biology, 26(6), 1137–1144. https://doi.org/10.1111/j.1523-1739.2012.01915.x

Nordstrom, K. F., Jackson, N. L., Smith, D. R., & Weber, R. G. (2006). Transport of horseshoe crab eggs by waves and swash on an estuarine beach: Implications for foraging shorebirds. Estuarine, Coastal and Shelf Science, 70(3), 438–448. https://doi.org/10.1016/j.ecss.2006.06.027

O'Connell, M. T., Franze, C. D., & Spalding, E. A. (2005). Special Issue No. 44. Saving America’s Wetland: Strategies for Restoration of Louisiana’s Coastal Wetlands and Barrier Islands. In Poirrier Source: Journal of Coastal Research.

Pahl-Wostl, C. (2007). Transitions towards adaptive management of water facing climate and global change. Water Resources Management, 21(1), 49–62. https://doi.org/10.1007/s11269-006-9040-4

Perez-Riverol, Y., Gatto, L., Wang, R., Sachsenberg, T., Uszkoreit, J., Leprevost, F. da V., Fufezan, C., Ternent, T., Eglen, S. J., Katz, D. S., Pollard, T. J., Konovalov, A., Flight, R. M., Blin, K., & Vizcaíno, J. A. (2016). Ten Simple Rules for Taking Advantage of Git and GitHub. In PLoS Computational Biology (Vol. 12, Issue 7). Public Library of Science. https://doi.org/10.1371/journal.pcbi.1004947

Pinelli, J.-P., Roueche, D., Kijewski-Correa, T., Plaz, F., Prevatt, D., Zisis, I., Elawady, A., Haan, F., Pei, S., Gurley, K., Rasouli, A., Refan, M., Rhode-Barbarigos, L., & Moravej, M. (2018). Overview of Damage Observed in Regional Construction during the Passage of Hurricane Irma over the State of Florida.

Purtlebaugh, C. H., & Allen, M. S. (2010). Relative Abundance, Growth, and Mortality of Five Age-0 Estuarine Fishes in Relation to Discharge of the Suwannee River, Florida. Transactions of the American Fisheries Society, 139(4), 1233–1246. https://doi.org/10.1577/t09-180.1

Rahman, M. M., & Roy, C. K. (2014). An insight into the pull requests of GitHub. 11th Working Conference on Mining Software Repositories, MSR 2014 - Proceedings, 364–367. https://doi.org/10.1145/2597073.2597121

Ram, K. (2013). Git can facilitate greater reproducibility and increased transparency in science (Vol. 8). http://www.scfbm.org/content/8/1/7

Sankar, R. D., Donoghue, J. F., & Kish, S. A. (2018). Mapping Shoreline Variability of Two Barrier Island Segments Along the Florida Coast. Estuaries and Coasts, 41(8), 2191–2211. https://doi.org/10.1007/s12237-018-0426-3

Sassaman, K. E., Wallis, N. J., McFadden, P. S., Mahar, G. J., Jenkins, J. A., Donop, M. C., Monés, M. P., Palmiotto, A., Boucher, A., Goodwin, J. M., & Oliveira, C. I. (2017). Keeping Pace with Rising Sea: The First 6 Years of the Lower Suwannee Archaeological Survey, Gulf Coastal Florida. Journal of Island and Coastal Archaeology, 12(2), 173–199. https://doi.org/10.1080/15564894.2016.1163758

Thomas, K., Kvitek, R. G., & Bretz, C. (2002). Effects of human activity on the foraging behavior of sanderlings Calidris alba. Biological Conservation, 109(1), 67–71. https://doi.org/10.1016/S0006-3207(02)00137-4

Tompkins, E. L., & Adger, W. N. (2004). Does Adaptive Management of Natural Resources Enhance Resilience to Climate Change? And Society, 9(2). https://doi.org/10.2307/26267677

U.S. Geological Survey, 1955, USGS 1:24000-scale Quadrangle for Cedar Key, FL 1955: U.S. Geological Survey. <https://www.sciencebase.gov/catalog/item/5a8a3ffbe4b00f54eb3ec75e>. (Accessed June 27, 2020)

U.S. Geological Survey,2021,02323500 Suwannee River Near Wilcox, Fla., <https://waterdata.usgs.gov/nwis/inventory/?site_no=02323500>. (Accessed January 28, 2021)

Vitale, N. E. (2019). Habitat Change, Predators, And Disturbance: Factors Influencing Productivity of American Oystercatchers (*Haematopus Palliatus*) Nesting in Florida’s Big Bend.

Walters, C. J. (1986). Adaptive management of renewable resources. Macmillan Publishers Ltd.

Warnke, D. A., Goldsmith, V., Grose, P., & Holt, J. J. (1966). Drastic beach changes in a low‐energy environment caused by Hurricane Betsy. Journal of Geophysical Research, 71(8), 2013-2016.

Weimer, T., Williams, B. K., Szaro, R. C., Shapiro, C. D., Adamcik, R., Boatman, M., Bransom, S., Casterson, J., Fay, J., Florence, S., Growitz, D., Hermans, C., Johnson, F. A., Kendall, J., Kubly, D., Mayer, M., Moyer, S., Pattison, M., Peterson, R., … Rodriguez, V. (2007). Adaptive Management the U.S. Department of the Interior Technical Guide Lead Authors Other Contributors Book Design.

Withers, K. (2002). Shorebird use of coastal wetland and barrier island habitat in the Gulf of Mexico. The Scientific World Journal, 2, 514–536. https://doi.org/10.1100/tsw.2002.112

Wright, E. E., Hine, A. C., Goodbred, S. L., & Locker, S. D. (2005). The Effect of Sea-Level and Climate Change on the Development of a Mixed Siliciclastic-Carbonate, Deltaic Coastline: Suwannee River, Florida, U.S.A. Journal of Sedimentary Research, 75(4), 621–635. https://doi.org/10.2110/jsr.2005.051

Yenni, G. M., Christensen, E. M., Bledsoe, E. K., Supp, S. R., Diaz, R. M., White, E. P., & Morgan Ernest, S. K. (2018). Developing a modern data workflow for living data. https://doi.org/10.1101/344804

Yu, K., Hu, C., Muller-Karger, F. E., Lu, D., & Soto, I. (2011). Shoreline changes in west-central Florida between 1987 and 2008 from Landsat observations. International Journal of Remote Sensing, 32(23), 8299–8313. https://doi.org/10.1080/01431161.2010.535045

Zedler, J. B. (2017). What’s New in Adaptive Management and Restoration of Coasts and Estuaries? Estuaries and Coasts, 40(1). https://doi.org/10.1007/s12237-016-0162-5

Zhang, K., Douglas, B. C., & Leatherman, S. P. (2004). Global Warming And Coastal Erosion. www.ipcc.ch.

BIOGRAPHICAL SKETCH

A biographical sketch is required of all candidates. The biographical sketch should be in narrative form. Third person, past tense, it typically includes the educational background of the candidate. The author should have replaced this paragraph with their own.