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

By

MELISSA MARIE MORENO

A THESIS 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

To my friends and family that have been there for me and to Michelle Masferrer; for inspiring me to make choices that fuel my life for the better

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

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 regarding managing gathering, collecting and reporting data.

YSI   Yellow Springs Instrument (manufacturer) of devices that can measure water quality observations

Abstract of Thesis 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

August 2021

Chair: William E. Pine III
Major: Interdisciplinary Ecology

Advances in technology such as expanded remote sensing and animal tracking platforms have triggered the rapid expansion of data available for natural resource scientists. I customized a modern data workflow to promote reproducibility in our workflows and to reduce data collection errors. We incorporated 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 files.

Making these data and scripts available to multiple users is challenging. I have developed a repository structure using GitHub, so that files and data may be publicly available, which includes the following 1) evaluating our previous repository structure and workflow; 2) creating a new and consistent structure and workflow among all project repositories; and 3) establishing and maintaining a file naming convention that encompasses any file that could be in a repository.

I also assessed shoreline change- near Cedar Key, Florida using NAIP (National Agriculture Imagery Program) aerial images and ESRI's ArcMap© spatial data analysis

extension Digital Shoreline Analysis Systems on three different time periods: 1994-2007, 2010-2019, and 1994-2019. I determined that shoreline features had both eroded and accreted during this period with more locations experiencing erosion. Reasons for the observed patterns of shoreline change may include sea-level rise, storm frequency and magnitude, and the local patterns of sediment transport.

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 a single discrete location and a single point in time sample collections to continuous real-time observations at multiple locations (Martinelli et al., 2016). While many monitoring programs' scale and technological capacity have 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 essential 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 considerable restoration effort to reverse observed declines in crucial 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 frequently occur throughout the project. Several restoration programs in this funding program require basic adaptive management concepts 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 inform the restoration actions such as the type of substrate to use in an oyster restoration project or monitoring program design as the project is ongoing, and by 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 levels 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 numerous spatial sites. For both cases, there is a need to capture and process data to meet

14

standards and then complete routine analyses of these data to ensure they help inform 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 disclosing restoration actions from these analyses. For this project to efficiently operate in an adaptive management framework, we developed a system that captures data as it is collected, guides the data to analyses, documents data, 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 findings of 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 that 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 include 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 need 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

15

previously stated objectives is not new and is a central aspect of the adaptive

management process for natural resources as initially 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. To be

informative, 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 research, 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

fundamental 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

include (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) version control allows these changes to be undone if needed, (4) version control 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 is observations of water conductivity and temperature gathered hourly from autonomous sensors. These data types are measured and recorded by the sensor deployed in the water and are output in a standard format that can be interpreted for analysis directly by a computer. The second metric of project interest is oyster counts 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 entered into a computer as a standard data form before being 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 number of sampling trips needed to optimize oyster density estimates. We use open-source software and tools, widely available and familiar to many field biologists, such as program R (Lefcheck, 2016) and Microsoft Excel. This paper 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 data types 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 I predefined and abbreviate (e.g., Lone Cabbage= LC, Horseshoe Beach = HB). A number is added to the end to determine the site 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 site and reference the same site 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 sampling location's naming convention before the sampling event.

  Another aspect of my naming convention standards, which directly relates to data management, is how 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 clearly identify the date and site, especially when investigating sensor readings that 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

The LCR project collects 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 scheduled to ensure a continuous stream of data by reducing data errors due to biofouling or equipment loss. These "living data" have the highest frequency of occurrence (the greatest number of observations) and require strict data management protocols to launch and maintain the sensors and import 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. These low-tide events de-water oyster reefs allow teams of people to count and measure oysters to document the 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 cm, the data validation will reject this entry because there is a range of acceptable oyster heights that it will allow. The project research coordinator will then reconcile any mismatch or errors

in the dual data entry worksheets. This double data entry system was created to reduce the chance of data entry errors and human-introduced errors.

Water Quality Measured by Field-Crews

The LCR project also collects water quality measurements using a hand-held YSI (Yellow Springs Instrument) device to provide an additional check on our autonomous sensor observations during water quality service trips. 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.

While conceptually, each of these data types appears 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 discovered through regularly updating these data types may also address many data management challenges that researchers may confront.

<div align="center">

Establishing a Modern Data Workflow

</div>

Data to support research efforts assessing the effectiveness of the LCR restoration are collected in the field are transferred and stored in a relational database for QA/QC and analyses. Database development efforts for this project started before data collection through the development of database "blueprints" via whiteboard 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 the University of Florida Academic Research Consulting & Services (ARCS, http://arcs.uflib.ufl.edu/). A key database need that was identified in blueprinting of the database 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 I have developed for water quality management addresses goals and unique 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, I 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 are 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 datasheet.

3.A. Water quality sensor files are then uploaded into a secure University of Florida internal server. A trigger starts the Python import process into the MySQL relational database, which permanently stores raw files as an archive. The YSI measurements (recorded by hand in the filed) 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

To control access to critical database import-export features, I implemented a dedicated username and password controls to maintain access to the MySQL database to track access and database changes. The database stores water quality data in specific tables where the sensor serial number and location must be predefined before 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). 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 (as a data release) using version control in the project GitHub master data repository (https://github.com/LCRoysterproject). This repository includes an up-to-date master branch that 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 team member to ensure data integrity. Every pull request requires a detailed message

describing each change in the event an update to the master branch has to be investigated. Version control allows 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 will determine why the observations are labeled as duplicates. 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. Also, 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 where (Figure 1-5):

1. Datasheets are standardized before 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. 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 in 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 data sheets were designed to (1) detail the format of information to be recorded, (2) minimize errors, (3) allow for easy transcription from field observation to paper to enter 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. For 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. Doing this reduces the chances 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 predefine a minimum and maximum range expected of any given oyster height. 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, investigating the discrepancy using the original datasheets. 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 before 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 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 the funding agency. Because data workflow is standardized, each time new data are collected, these standard computer scripts can be run to inform ongoing research efforts. For example, to inform field sampling, we routinely use a type of power analysis to inform field efforts during winter oyster sampling (Moore and Pine in-press). Before the field sampling season, data from previous years are used to develop preliminary sampling guides regarding 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 update the sampling effort for that field season based on observed oyster density and variability from 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 it is only possible because of a robust functioning data workflow.

Discussion

Establishing a data management workflow is receiving increased recognition as a critical part of ecological research (Lowndes et al 2017). Thus, creating a data management workflow from the beginning of the research initiative makes data management a more straightforward endeavor to maintain than 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 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 and 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 various skills, using software a biologist may already be familiar with from other work elements, to employ a modern workflow.

The data workflow for the LCR project I describe may not be a universal template for every ecological project, but many challenges this design has overcome are likely common to other projects. The critical concept of creating a data management workflow before beginning the LCR project is one of the main reasons this workflow has persisted throughout the project. Developing the workflow necessitated vital elements of the project, such as the location of water quality sensors, to be determined. While initially, the time required to develop the data workflow was significant, 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 data workflow 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.

There are many advantages to using open-sourced tools (e.g., GitHub, R programming, and MySQL) in a data management workflow. First, these software are

free, widely available, and have an active online support environment. 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 that 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 its 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 ensure that they follow 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 problems may make the training process

more acceptable. It is essential to communicate effectively with team members to

guarantee they collect and maintain 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 cannot 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 would likely not be exceeded for decades of data collection given our current data storage needs. The use of GitHub repositories 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 that continually improves management policies and practices based on data outcomes (Pahl-Wostl, 2007). With increasing data types available to inform management decisions, 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 not to be widely implemented in practice (Walters, 2006; Weimer 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.

A. Solidify restoration project objectives and goals

D. Implement new findings into research objectives

B. Create plans to collect, store, analyze, model, and report biological and biophysical data

C. Actively inform decision making based on reporting and analysis
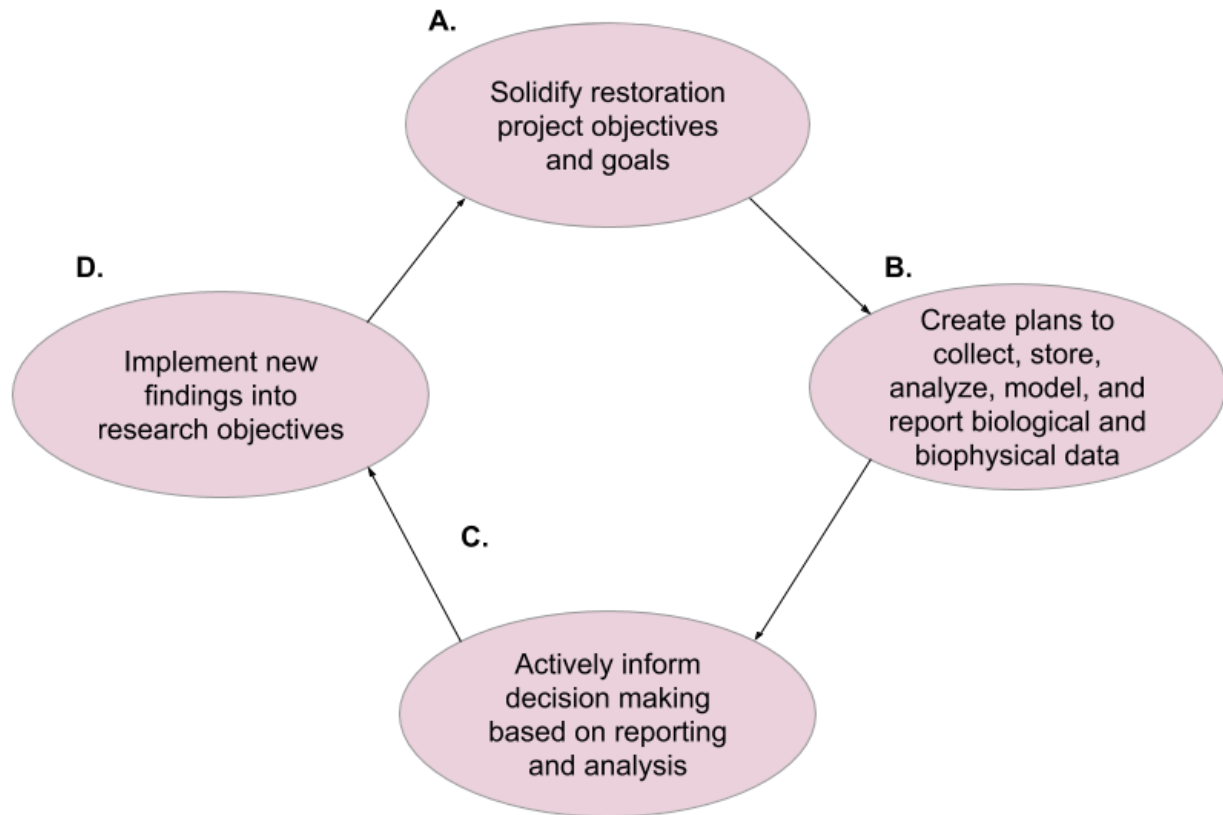
Figure 1-1. The adaptive management process for ecological restoration projects. A reliable data workflow contributes to each phase of the adaptive management process from planning to implementation.
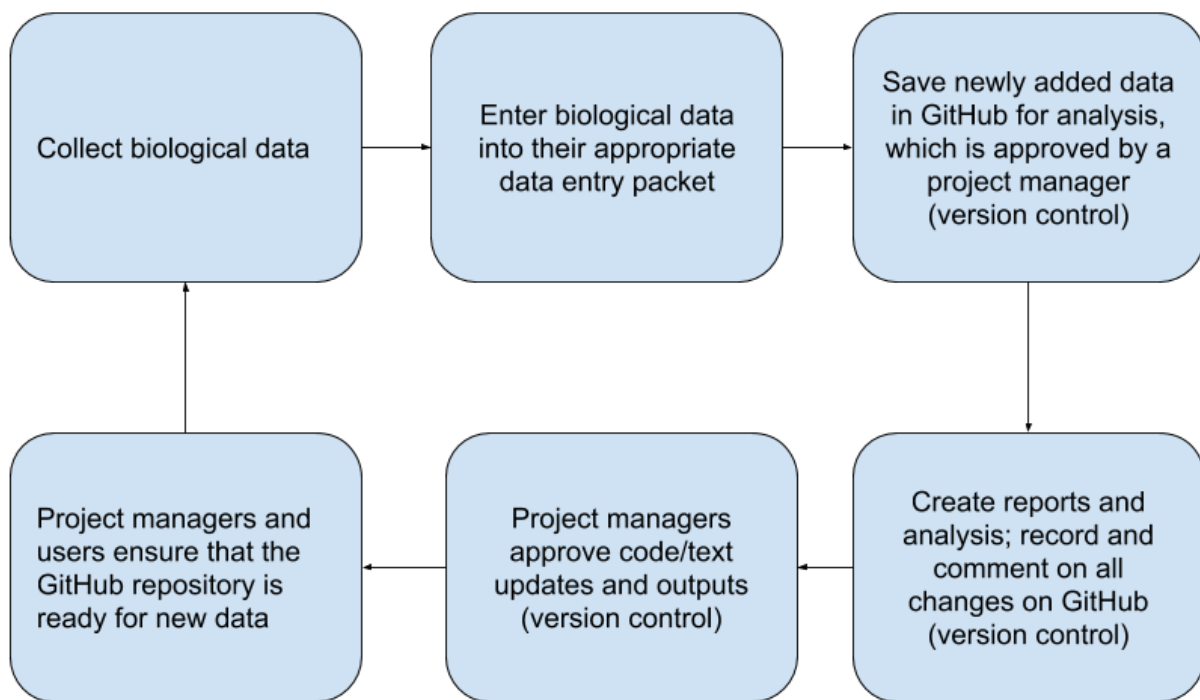
Figure 1-2. Generalized version control workflow for the LCR project; detailed workflow information can be found here (Moreno et al., 2020, GitHub Workflow for the LCR Oyster Project).
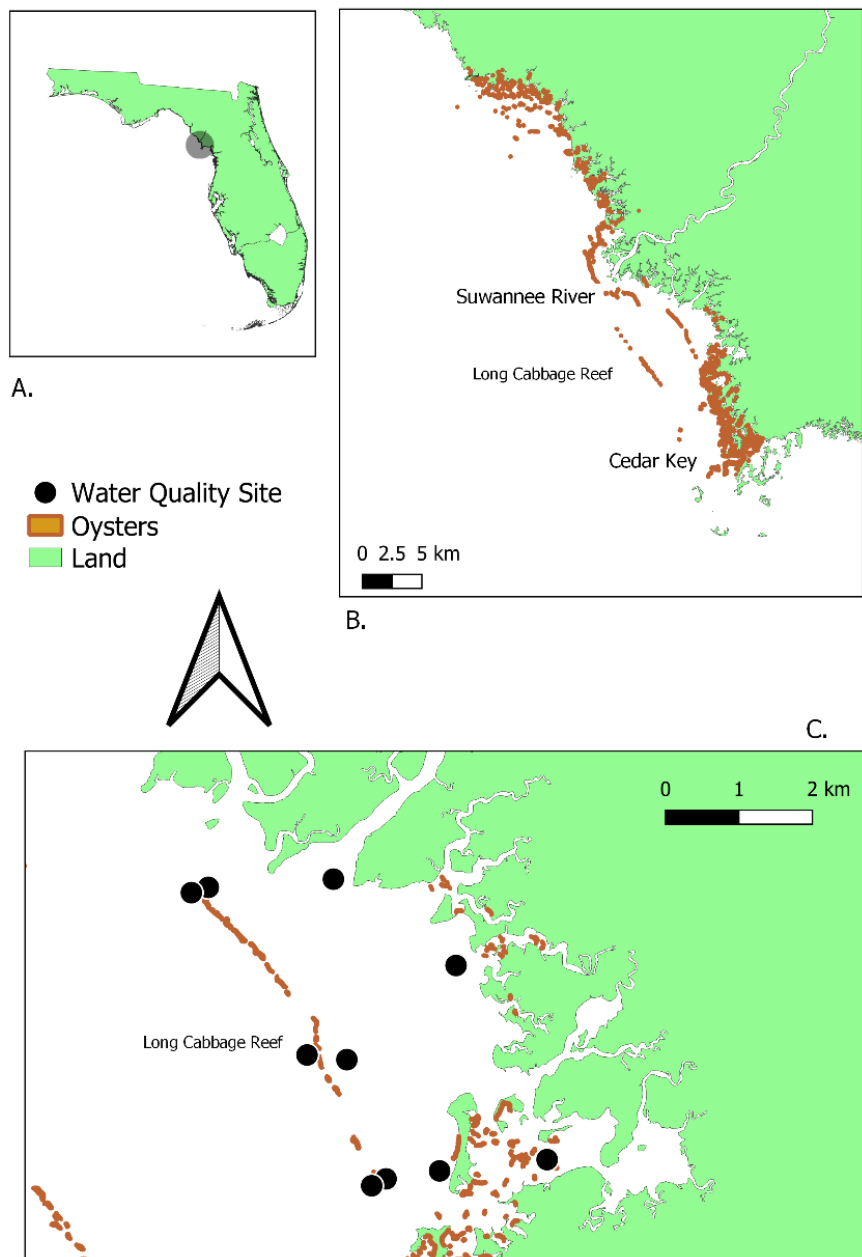
Figure 1-3. Water quality location map. A) Lone Cabbage Reef area; B) Florida coastline between identifying Suwannee Sound river mouth, Lone Cabbage Reef, and Cedar Key, FL; C) Lone Cabbage Reef with water quality sites identifies (black circles). The oyster shapefile used in this map is from a sampling effort in 2001.
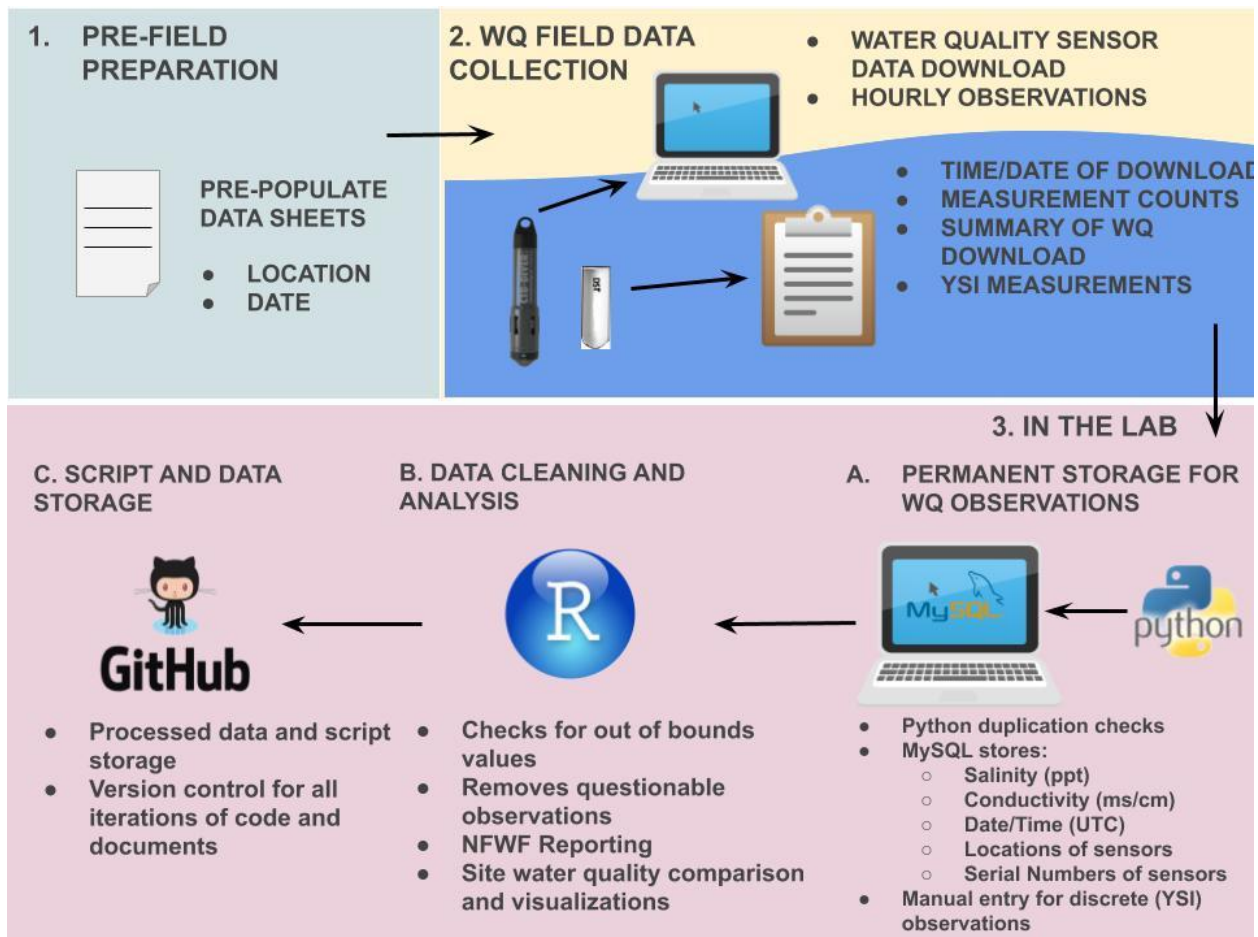
Figure 1-4. Schematic describing the data workflow for the Lone Cabbage Reef water quality observation network.
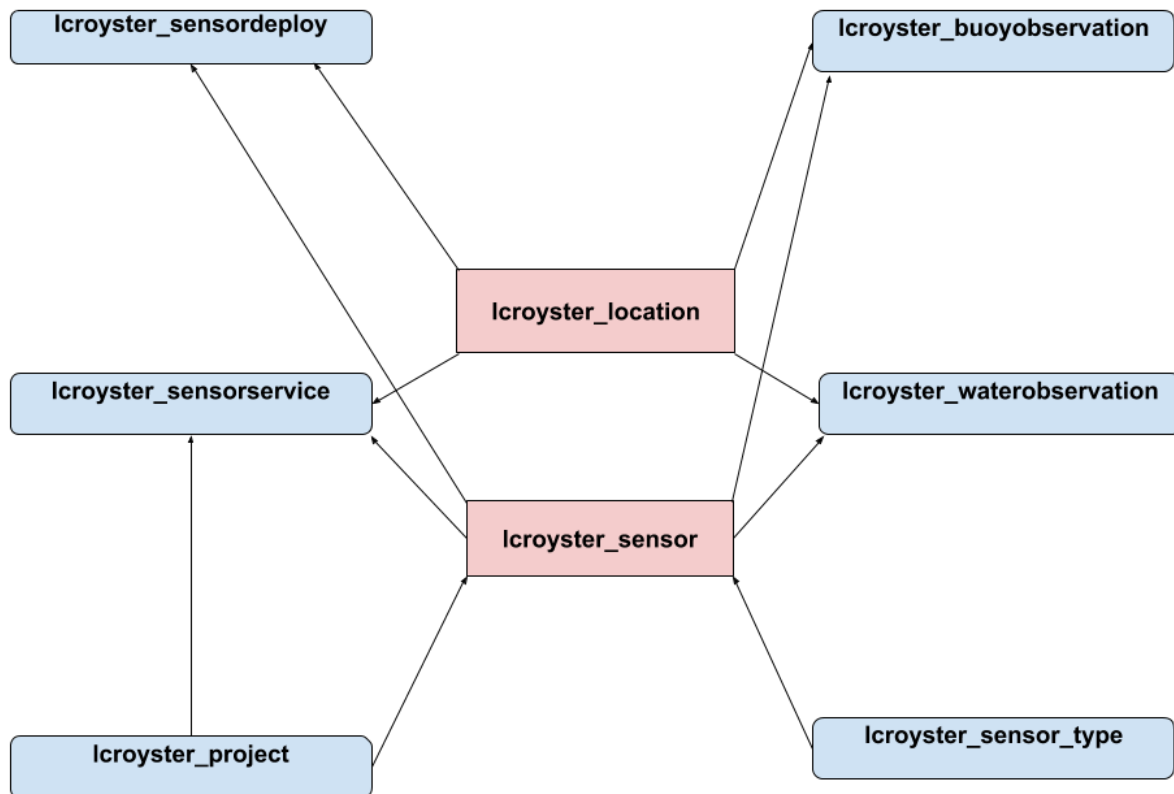
Figure 1-5. Diagram of how the tables in the Lone Cabbage Reef MySQL relational database are connected by a sensor's location (lcroyster_location) and serial number (lcroyster_sensor). More information on the description of the tables and the data they archive can be found at Moreno et al., 2020, MYSQL workflow for the LCR Oyster Project.
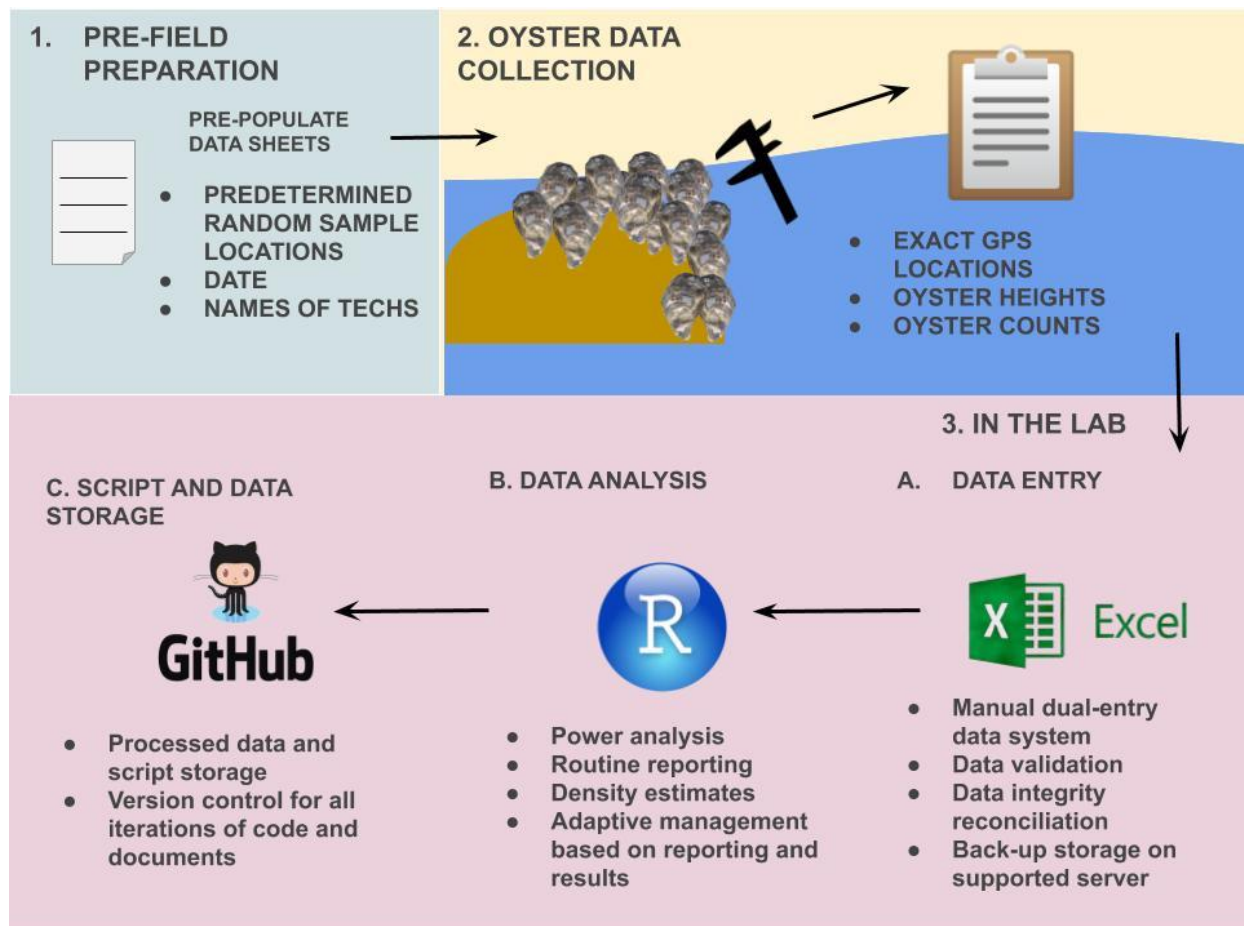
Figure 1-6. Schematic describing the data workflow for the Lone Cabbage Reef oyster samples from field collections.

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 these data sources' knowledge. Despite new data sources commonly becoming available, researchers are rarely trained in best practices to capture, manage, and analyze these data types (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 write, share efficiently, and archive computer code used to analyze data in a reproducible framework (i.e., can be readily interpreted and done by others in the future) is a crucial aspect of "open" science platforms that encourages transparency in data availability and analyses (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 initially developed to share computer code across teams working more on information technology projects (Perez-Riverol et al., 2016). 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 an extensive restoration project in Suwannee Sound, Florida (Chapter 1) 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 to 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 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)

- Living data- data that are collected and updated frequently to continuously analyzed (Yenni et al., 2018)

- Version control- a system that allows the users to track iterative changes to code and text (Blischak et al., 2016)

- Project repository- a 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 way to submit track changes in a repository through a message, and an admin may review a record of the change

- 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 before 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 data analyses are completed within the LCR project to generate

results used to reach a conclusion or inform a decision. How this code is written can

vary from person to person, and when teams of people are collaborating on coding

projects, identifying how one person did or did not edit a section of code, or what

changes were made to a code that may be caused the code to stop working, or improve

the code, can be difficult to identify (Blischak et al., 2016) readily. To keep track of these

changes, version control systems that 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 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 can see what has been changed concerning the original document. The user will then implement those changes through the version control system, so they are tracked and transparent to other users.

In the LCR project, code collaboration is standard practice. Before developing the GitHub repository structure, code writing collaborations in our project ultimately ended in several files of the same script. Still, 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 (if 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 record which files have been changed, who has changed them, and why they were changed. Using version control software saved us time in determining when and why specific 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 initially designed, was becoming increasingly more challenging 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 did not address many of our newfound needs, such as accounting for multiple working projects using the same data. Without proper guidelines, the main GitHub repository quickly began to grow and expand (Figure 2-1) in an unmanaged way. Our main GitHub repository started to store multiple projects inside of itself, leading to confusion in the repository in which data and scripts were used for each of the different projects. We soon realized that the GitHub repository structure we had employed was not effective in keeping our files or projects organized. This led to a significant revision in the 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 was to keep different sub-projects separate (i.e., water quality summaries for reports separate from water quality analyses) but to have a shared "living data" source. 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 the transparency of what each script does and its outputs. Our project team members must update README.md files 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-sufficient, and only team members actively working on that project have access to them. Second, these repositories are independent, and other project repositories' scripts do not influence their 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), our workflow needed to include 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. If a team member needs to update the data for their analysis, they can 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 that 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. For each LCR project repository with collaborators, we establish a protected `master` branch and open collaborator branches. Collaborators can edit and modify their branches however they please. Still, they cannot 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 collaborators' work 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 any repository, private or public status at any time. These 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 exist in which if a script file creates a function or a specific 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 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 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 will help other ecologists pursue 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 finding the repository and possibly beginning collaborations. Through a pull request, commit messages are easily seen and located in GitHub and allow collaborators to understand any change submitted to the repository. Transparent repositories are unlikely to be "scooped" by another researcher who can claim the data. 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, increasing the time it takes to debug a script issue. Generally, most scientists 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 more significant impact (Piwowar et al., 2007) from the research.

Advantages of our approach include that GitHub is accessible to anyone with the Internet. Many training programs 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 of using GitHub is that if a repository is accidentally deleted, a user has 90 days to retrieve it. A benefit in using our described approach is that it can be applied to any ecological effort with 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, a small group of ecologists can maintain this type of workflow with basic GitHub workflow training (Yenni et al., 2018).

Some disadvantages to our approach include that ecological projects, conducted in smaller teams, typically do not have the funding to hire a full-time data manager (Hampton et al., 2013), making it a timely effort to curate an efficient data management workflow. Data management planning is also typically 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 to their GitHub data management workflow. It is also important to note that even though GitHub has been discussed in detail 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 to organize their project through concise workflows, standardize naming conventions, and well-document README.md files. We hope 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 good future collaborations. The investment in creating a comprehensive data management workflow in GitHub will help decrease the time it 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.

Main Project Repository
- contains a specific data type
- all projects located in main repository
- all scripts and texts using this data type included in this repository

Singular LCR Generated Processed Data Type

Pull request to update data into project

Project A
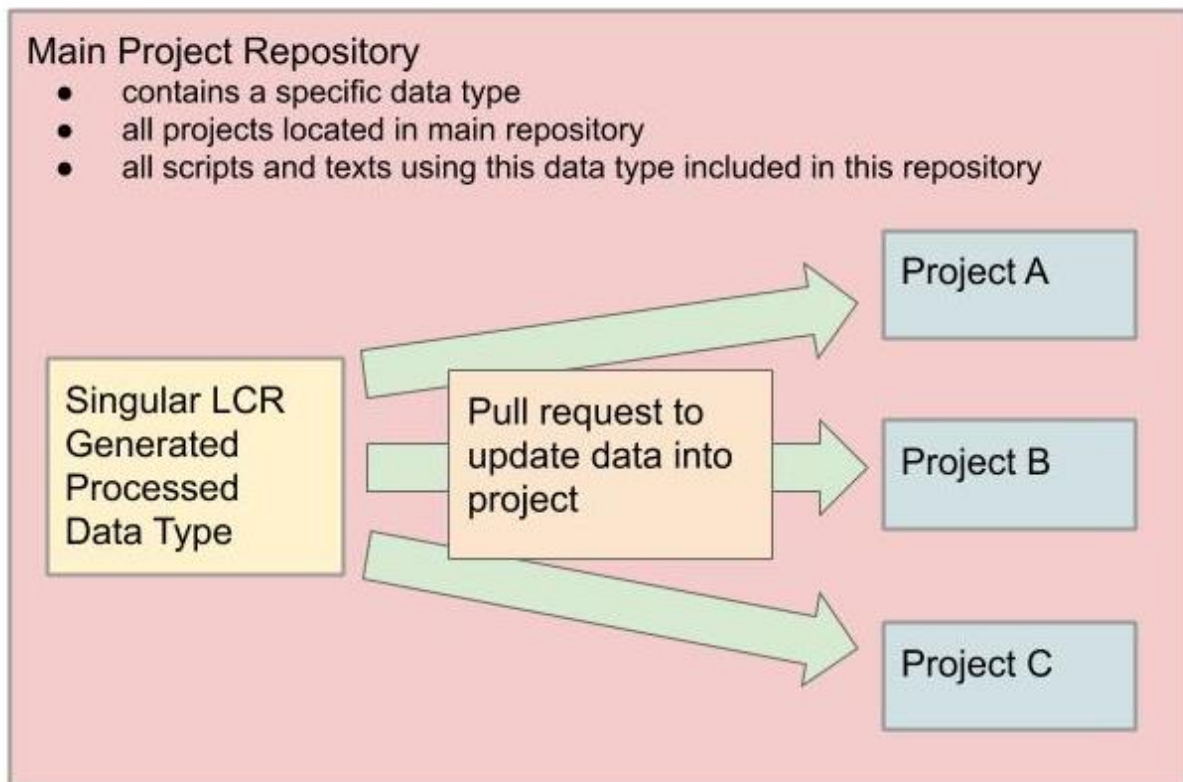
Project B

Project C

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

A.

Processed Data
- water quality
- oyster counts
- oyster measurements

B.

Master Data repository
- all project generated data
- most up to date data
- iterations of previous datasets
- widely used scripts

C.

Update project repository with the newly added data from the Master Data repository through direct download

D.

Project A
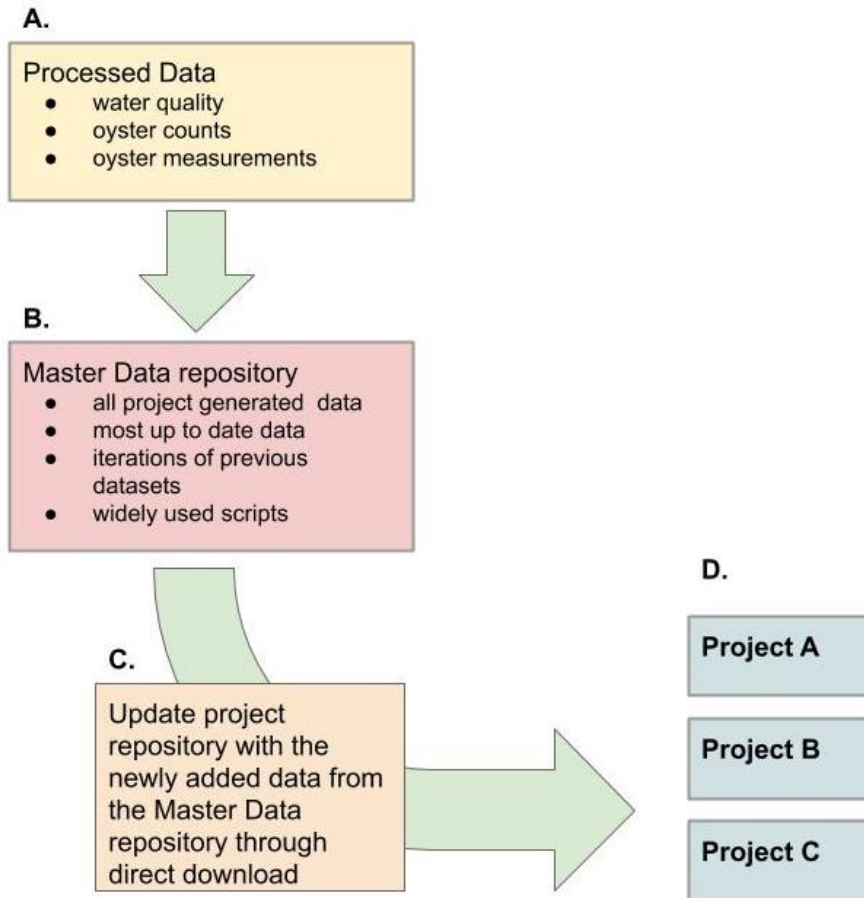
Project B

Project C

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 an 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 similarly to their file type 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 an 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
A DIGITAL SHORELINE ANALYSIS SYSTEM (DSAS) APPLIED ON SANDY
SHORELINE CHANGES IN DEER ISLAND, FL

Introduction

Coastal shoreline changes can occur due to multiple factors, including sea-level

rise (SLR), anthropogenic human activity such as development, and imbalances in

sediment transport which may be chronic or in the aftermath of storms (Yu et al., 2011).

The combination of these processes can influence erosional events and accretion of

shoreline areas. When shorelines change, the resilience of these geographic features

may be compromised, resulting in cascading effects to species, their habitats, and

ecosystems (Desantis et al., 2007). In Florida, the highest rates of shoreline migration

are usually concentrated around tidal inlets (Adams et al., 2016, Morton et al., 2004).

The west coast of Florida is characterized by low wave energy and low relief

geomorphology making it particularly vulnerable to factors driving coastal erosion

(erosion is defined as the process of depleting away of material from a shoreline due to

an imbalance in the supply and export of material resulting in the retreat and loss of

land and that may be potentially recovered through natural or man made processes)

(Morton et al., 2004; Geselbracht et al., 2011; Marchand et al., 2021). Florida has been

known to mitigate long-term erosion through strategies such as beach nourishment and

hard engineering structures (Morton et al., 2004). However, these long-term "solutions"

are usually reserved for areas of high tourism in Florida (Elko et al., 2021).

Climate Change and SLR

In recent decades, the rate of SLR in many regions of the world has increased,

most likely due to changing climate (Cahoon and Gunternspergen 2010, Mimura 2013).

This acceleration in SLRs and the observed impacts on coastal environments, including

urban areas (Habel, 2020) and natural regions (Williams et al., 1999), is of increasing concern to property owners, municipalities, and natural resource managers (Kopp et al., 2019). The impacts of SLR globally are predicted to be profound to human communities and the natural environment (Brown et al., 2013; Curtis and Schneider, 2011). Shoreline migration occurs when SLR shifts the high-water line (location on the shore where the water usually reaches high water) landward concerning the slope of the coastal area (Zhang et al., 2004).

In the Gulf of Mexico region, including the west coast of Florida, sand-dominated shorelines are common. Sandy shorelines are characterized by active environments and unstable substrates, consisting of sand, quartz, and/or silica (Brown & McLachlan, 2002). The unstable nature of unstable sand shorelines can create a harsh environment for biota, necessitating unique adaptations to this volatile environment (Brown & McLachlan, 2002). Sand shorelines accumulate sediment accretion by wave deposited particles of sand, mixed sand, quartz, or silica. These particles originate from a combination of inland erosion of material, which is transported and deposited within coastal environments at river mouths (Brown & McLachlan, 2002) and marine biogenic sources, including marine skeletons, sponge spicules, and shell fragments (McLachlan, 1990). Stable shorelines exist at an equilibrium of new material being deposited at the same rate material is eroded. However, when this equilibrium is disrupted or altered due to sediment transport or erosion, shoreline habitats can change rapidly. This exchange implies that seawater levels directly correlate with sandy beach erosion (Zhang et al., 2004).

Shoreline loss has been observed in the Big Bend and adjacent areas of Florida in recent decades. Seavey et al., (2011) documented the loss of about 66% of the offshore (most westward) intertidal oyster bars between Corrigan's Reef and Horseshoe Beach. Vitale et al., (2019) documented how following the construction in the mid-1960s, of spoil islands as part of the Cross Florida Barge Canal by the US Army Corps of Engineers wildlife including endangered shorebirds began to use these spoil islands as habitat. However, in recent decades these spoil islands have experienced significant erosional events (Vitale et al., 2021) and related loss of habitat available to these wildlife. Similar patterns have been observed for natural islands in the region (Vitale et al., 2021), including Derrick Key near Cedar Key, Florida. Derrick Key was a well-known local landmark to commercial fishing families in the Cedar Key regions for decades and photographs from 1982 demonstrate several prominent geographic features including vegetation. However, by 2016 the island was submerged (Vitale et al., 2021) Large scale efforts to analyze shoreline changes in Florida have been completed in recent years (Yu et al., 2011; Sassaman et al., 2017; Li & Gong, 2016) and in general these studies document A) shoreline changes in the west coast of Florida are commonly eroding, B), storms, sea-level rise, and wave action possibly contribute to shoreline migration, C) and beach nourishment may provide a way to combat erosional event effects temporarily. However, studies that focus on possible effects of shoreline change and SLR on small or regional scales within Florida, particularly at the scale of habitat patches used by wildlife such as nesting shorebirds or oyster reefs, are not as common.

Study Site

  The Suwannee River estuary, Suwannee Sound, is a siliciclastic, sand-starved, and low-wave-energy system dominated by marshes that open toward the sea (Hine et al., 1988). Shoreline profiles in these systems can change over time due to low wave energy (Jackson et al., 2002), and storm events have also been observed to cause rapid change at specific locations south of the Suwannee Sound (Goodbred and Hine 1995).

  The Suwannee River is one of the largest rivers in Florida, spanning 370 km in length, from southern Georgia to the west-central Florida coastline, and is considered a significant point source of sedimentation in the Suwannee Sound (Wright et al., 2005). The Suwannee River is a partially spring-fed system that drains the coastal plain of Georgia and provides a restricted point source input of siliciclastic sediment, creating a small 20-kilometer delta (Wright et al., 2005). The surrounding coastal regions of the Suwannee River are otherwise known to be sediment starved. An ancient significant sedimentology event has 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 has typically high discharge peaks between February and April and low discharge peaks between August and October (Purtlebaugh & Allen, 2010). The mean annual river discharge measured of the Suwannee River at the USGS Wilcox site (latitude 29.58, longitude -82.93) is 324.40 m^3/s with a minimum discharge of 30.29 m^3/s and a maximum discharge of 1554.59 m^3/s, calculated from years 1950 to 2021 (USGS, 2021). The Suwannee River flow can back up into the fresh spring run during periods of high spring flow and can produce negative velocities and discharge.

Suwannee Sound is encompassed by three rural Florida counties, Dixie, Levy, and Taylor. Currently, these counties are among the lowest human population density for coastal counties in Florida, but their populations are predicted to increase in future decades (Bureau of Economic and Business Research, 2021). Human development on coastlines may accelerate coastal erosion by creating a fixed position of the shoreline and stabilizing inlets, alters sediment transport (Finkl & Charlier, 2003). Increased human developments may also negatively impact coastal species diversity. Species biodiversity is threatened by the increase of urbanization and coastal environmental degradation (Finkl & Charlier, 2003). Czech et al. (2000) documented urbanization as the highest cause for species endangerment. For example, bird species, including Piping Plover (*Charadrius melodus*) and American oystercatcher (*Haematopus palliates*), are known to forage and roost in coastal areas of low human population, including Suwanee Sound (Thomas et al., 2002). Species biodiversity, both vegetative and animal, could be at risk due to increased urbanization along coastlines (McKinney, 2006) and accelerated shoreline migration.

Recent Storm Events in Suwannee Sound Region

Suwannee Sound is considered a low-energy environment because the nominal wave height is below the nominal high water line. High energy events, including tropical and winter storms, can increase wave and wind action in the region. Recent significant storms in the area include the "storm of the century" in March 1993, with wind speeds of > 15 m/s for 16 hours recorded at the Crystal River Power Plant roughly 100 km south of Suwannee Sound (Goodbred & Hine, 1993a). This weather event caused extensive damage to Waccasassa Bay (approximately 30 kilometers south of Suwanee Sound), including 3-meter water storm surges and storm-driven sediment deposits of up to 12

cm on coastal shoreline features up to 2 cm the marsh surface (Goodbred & Hine, 1993a). In 2017, Hurricane Irma, a category 3 hurricane, made landfall near Marco Island and moved north along the Florida coastline, causing excessive rain and coastal flooding in the Suwannee Sound region. In 2016, Hurricane Hermine caused major flooding in Cedar Key, Florida, including the highest observed storm surge of >2.0-m (Berg, 2017). In 2017, Hurricane Irma, a category 3 hurricane, made landfall near Marco Island and moved north along the Florida coastline, causing excessive rain and coastal flooding in the Suwannee Sound region. Tropical Storm Eta also made landfall in Cedar Key, as a minor storm in 2020 (Lyons, 2020).

Within Suwannee Sound (Figure 3-2), prominent coastal features include numerous tidal creeks, intertidal and subtidal oyster bars, and small islands. One prominent island, Deer Island, is a privately owned uninhabited island approximately 13 kilometers north of Cedar Key, Florida. Historically, Native Americans intermittently inhabited Deer Island for thousands of years (USGS, 1955), and long-time Cedar Key residents report early Florida settlers lived and camped on the island as well. The 1800 Florida census registered four people who identified this island as their home, and a small cabin is identified on a 1951 USGS Cedar Key Quadrangle map (USGS, 1955) of the region. At present, this island is 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 1,300 meters long from north to south and approximately 250 meters at its widest point (Monés et al., 2012). Because of its historical and cultural significance to the region, Deer Island is commonly used by local residents as a geographic reference point for

navigation and a recreation area. These same residents have also reported that Deer

Island has changed in recent decades both in shape and area.

Objectives

I examined the Deer Island area to assess whether shoreline change observed

by local residents in recent decades was evident from available imagery of this region.

Because Suwannee Sound is a region of low human population density and the

immediate shoreline areas surrounding Suwannee Sound are state or federally

protected lands, including the Lower Suwannee National Wildlife Refuge, shoreline

change in this area is less likely to be influenced by factors that have contributed to

shoreline change elsewhere such as development. Any observed change in shoreline

feature are more likely to come from other factors including SLR and storm events.

<div align="center">

Materials and Methods

</div>

I cataloged available imagery of shoreline features for the region of Suwannee

Sound. After compiling images from the National Agriculture Imagery Program (NAIP), a

25 year period of 1994 to 2019 was identified as having suitable images for use. This

time series was divided into three time frames to locate an area of shoreline change

where an identifiable factor may have triggered shoreline migration or accretion. These

time frames were chosen to encompass a general period of 12.5 years each. Two out of

the three time frames divided the available imagery into equal years; however, there is

no equal amount of imagery available covering each 12.5 years period (e.g., 1994-

2007, and 2010-2019). I also combined all imagery to calculate how the shoreline

changes from 1994 to 2019.

Imagery Selection Process

Locating relatively cloud-free imagery for a specific location in Florida can be difficult. Since our study location is unpopulated and contains no popular historic landmarks, historical aerial images are not frequently taken. To reduce the effort on locating suitable cloud-free imagery I used Google Earth Pro as a tool to screen images. Google Earth Pro does not capture any original imagery; however, it locates and uses imagery from multiple sources within its finder view that meet specific criteria including location, resolution, and absence of clouds. After an initial screening of available imagery using Google Earth Pro, I took this group of available images and then linked their information to the metadata for each image via the USGS Earth Explorer (https://earthexplorer.usgs.gov/) modules. This allowed for the full metadata for each image to be collected including the time the image was collected, which was crucial to relate to data standards, such as tidal height, I adopted for image standardization. The most common source of images include the NAIP (National Agriculture Imagery Program). The data standards for this program require a 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 attempts to comply with the specification that no more than 10% cloud cover be allowed in each aerial imagery tile. Aerial imagery is available as digital ortho quarter quad tiles (DOQQs) GeoTIFFs, 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 essential to select imagery close to the same time of the year, similar river discharge, and precipitation levels. All imagery chosen are between October through January. Table 3-1 includes all metadata

associated with the imagery used in this analysis. Furthermore, observed weather and median river discharge are described, including the observed weather for the day of imagery collection and median river discharge measured. All selected images are available within the electronic repository for this thesis (https://github.com/melimore86/dsas_analysis).

National Agriculture Imagery Program employed sensor types with three-band imagery categorized as RGB (red, green, blue) until 2007. After 2007, four-band color infrared imagery was collected and categorized as CIR/CNIR (red, green, blue, and infrared). The imagery was uploaded into ArcMap© for analysis. Four band imagery is multispectral, which means the sensors can collect information from several parts of the electromagnetic spectrum. To specify a natural color display, I set ArcMap© to read band 1 as red, band 2 as green, and band 3 as blue. Table 3-2 includes the sensor type associated with each image. The first available suitable image, November 2007 did use color infrared (CIR/CNIR), which could allow false color for NDVI (Normalized Difference Vegetation Index) analysis. However, this was not necessary for any image because the vegetation is distinct from, and not integrated into, the sandy shoreline. The mean high water line was not necessarily easy to determine in all imagery, and so the shoreline to ocean line was chosen to digitize. There is a clear and distinct separation between sand and vegetation. Additionally, the DSAS user manual does not have any recommendations for using true-color image composites.

Due to the unknown time of day the aerial imagery was captured the tidal height information is a source of uncertainty in this analysis. The tidal height information would provide information if the imagery was in a low or high tide event at the time the imagery

was taken. The DSAS user manual considers natural influences over the shoreline position such as wind, waves, and tidal height as uncertainty that is accounted for in DSAS calculation default settings unless otherwise adjusted.

Digital Shoreline Analysis System (DSAS)

The DSAS is a GIS-based system created and maintained by United States Geological Survey (USGS). 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. The user constructs baselines, and this analysis was created using the Buffer tool in ArcMap©. These shoreline positions provide the necessary data needed to calculate their shifts. The DSAS analysis generates transects perpendicular to the reference user-created baseline (Figure 3-3). 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 generate forecasted transects for 10- 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, other shorelines of interest (varying in different 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 recommended by the DSAS User Manual. The DSAS operational workflow includes the following steps: (1) Set default parameters and fields to the 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) Beta Shoreline Forecasting for a 10 and/or 20-year prediction. One of each type of change metric ("Digital Shoreline Analysis System (DSAS) Version 5.0 User Guide.", 2021) was used in this analysis, LRR (Linear Regression Rate) for statistical analysis to calculate the rate of shoreline change, and Net Shoreline Movement (NSM) calculation for the distance measurement.

Linear Regression Rate (LRR)

A LRR of shoreline change can be found by fitting a least-squares regression line to all points for every shoreline in a transect. The parameters for the regression were found by minimizing the sum of squares between observed and expected points and the key regression parameter of interest was the slop which equals the rate of change estimated by the LRR. The LRR calculation can be used regardless of accuracy or trends ("Digital Shoreline Analysis System (DSAS) Version 5.0 User Guide."). This method may be susceptible to outlier effects and tends to underestimate the rate of change compared to other statistics used by DSAS (e.g., End Point Rate (EPR)).

Net Shoreline Movement (NSM)

In contrast to LRR, NSM calculations only require baseline position shoreline and the last shoreline position to (two shorelines) interpret shoreline shape change. The NSM measures the distance between the oldest shoreline (e.g., 2019) and the youngest shoreline (e.g., 1994) for each casted transect measured in meters. The NSM estimate provides information on the magnitude (area) and direction of shoreline change

(whether erosion or accretion is occurring) and by default a measure of uncertainty associated with the NSM estimates (90% confidence intervals).

Beta Shoreline Forecasting

Version 5.0 of DSAS (v5.0) has a tool in development that takes observed rates of shoreline change and predicts future shoreline change, assuming a similar rate of change as observed, for periods 10 or 20 years into the future. These predictions are made using Kalman filter (Kalman, 1960) to join shoreline positions with model-derived positions to predict a future shoreline position (Long and Plant 2012). The Kalman filter methodology is initialized with the linear regression rate calculated by DSAS. Using the linear regression I used the DSAS extension to estimate the shoreline position and rate of every 10th of a year (e.g, year 10 and year 20). The DSAS extension also estimates the positional uncertainty at each time step. It also estimates the positional uncertainty at each time step. Critically, this methodology assumes that the linear rate of regression from the shoreline positions analyzed will be a reasonable approximation for the position of future shorelines.

DSAS Parameters and Selections

Selected NAIP Geotiff aerial imagery were in 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 were created, traced, and digitized. The scale used to digitize was 1:3,000, and an MWL was discerned by looking at the whitest and brightest part of the shoreline that was not influenced by the dark ocean color. Shorelines from each image were then merged into a new single shapefile using the ArcMap© Merge tool. The baseline selected can be found on the east side of Deer Island. Both a

baseline shapefile and a merged shapefile sequential images over time are required for

DSAS calculations (Figure 3-4, Inputs).

After I selected and pre-processed NAIP imagery, the DSAS analyses were

completed using the following parameter specifications:

(1) 100 m buffer around the merged shoreline shapefile

(2) the east side of the buffer was used as the baseline (transects from the

baseline were cast from east to west)

(3) transects were spaced at 20-m intervals; the minimum transect spacing

allowed by DSAS based on the small size of the study site

(4) 2000-m search for suitable shorelines was done adjacent to the transect;

search distance looked for shorelines 2000 meters away from the baseline

(5) a smoothing distance of 500-m was specified; a smoothing distance is a user-

specified smoothing value that can facilitate an orthogonal transect intersect by

preventing transects from intersecting with one another when there is a curve in the

baseline, and the larger the smoothing distance, the more likely to produce uniform

transect orientations, which is recommended for smaller shorelines

Results

Key shoreline features of Deer Island have changed based on an assessment of

imagery from 1994 to 2019. This change is apparent both from a visual examination of

the images, as well as analyses using the DSAS tools in ArcMap©. The rate of change

estimated by the DSAS tools is about -0.95 meters per year, which when used as a

baseline for forecasting future change suggests that within 100 years the shoreline of

Deer Island will erode approximately 95 meters. The calculations for the shoreline

analysis are displayed in a colorblind color ramp. The LRR color ramp displays rates of

change in meters/year (m/yr), and the NSM color ramp displays the distance of measurements in meters. The DSAS calculations follow the standard that a negative rate implies erosion and a positive rate implies accretion. Results for all study periods conclude a pattern of consistent erosional events and some events of accretion.

Imagery Observation

By overlaying NAIP aerial photographs, in the same scale (1:3,000) and projection (NAD 1983 UTM, Zone 17N) and similar environmental conditions (Table 3-1) from 1994 and 2019 demonstrates changes in Deer Island shoreline features which suggest shoreline migration over this time period (Figure 3-5 A, B). For example, contrasting the 1994 and 2019 imagery suggests that the southwestern inlet has a larger open mouth to the sea in 2019 than the 1994 imagery. This suggests that this feature has expanded over time, possibly due to erosion. The northeastern shoreline of Deer Island also has an outcropping that looks round and full in 1994 imagery but has changed shape to be pointy and thin in 2019 imagery. This change also suggests shoreline migration. Overall from visual inspection of imagery it does appear that larger portions of the Deer Island shoreline have experienced erosional events than accretion. Tidal heights are unknown during the time the aerial photograph was taken since the time metadata is not offered.

Shoreline Analysis for Years 1994-2007

The LRR analyses of these same images for the years 1994-2007 found years and transects of both erosion and accretion (Figure 3-6, A). Overall, erosion was estimated to have occurred in 74.4% of transects, and accretion was estimated to have occurred in 25.6% of transects measured between 1994-2007. The most common values among transects were erosional rates of about -2.0 to -1.0 m/yr (about 30% of all

transects calculated had erosion rates this high), with some transects within years estimated to have lost between -5.0 to -3.0 m/yr (Figure 3-7, A). For transects within a year where accretion was observed, accretion rates of -0.5 to 0.5 m/yr were estimated.

The NSM results also demonstrate years of erosion and accretion (Figure 3-6, B). The highest erosion distance measurements range from -70.5 to -35.9 meters, and the maximum accretion distance measurements range from 4.3 to 6.4 meters (Figure 3-7, B). 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).

Overall, both models suggest that between 1994 and 2007, Deer Island changed in the area due to erosional processes, with the highest erosion rates near the north and south end of the shoreline. Both models also suggest that the most frequently calculated rate/distance is the rates/distance with intermediate erosion.

Shoreline Analysis for Years 2010-2019

The LRR analyses for the years 2010-2019 found years and transects of both erosion and accretion (Figure 3-8, A). The most common values among transects were erosional rates of about -0.5 to 0.5 m/yr (about 28% of all transects calculated), with some transects within years estimated to have lost between -4.0 to -3.0 m/yr (Figure 3-8, A). For transects within a year where accretion was observed, accretion rates of 1.0 to 2.0 m/yr were estimated (1.2% of all transects calculated). Overall, erosion was estimated to have occurred in 47.6% of transects, and accretion was estimated to have occurred in 52.4% of transects measured between 2010-2019.

The NSM results also demonstrate years of erosion and accretion (Figure 3-8, B). The highest erosion distance measurements range from -41.8 to -20.1 meters, and the maximum accretion distance measurements range from 8.7 to 9.9 meters (Figure 3-9, B). The most frequent NSM measurement range is -6.7 to 2.9 meters accounting for 25.6% of all transects calculated (Figure 3-9, B). The least frequent NSM measurement range is the accretional distance between 5.8 to 8.7 meters accounting for 4.9% of all transects calculated (Figure 3-9, B).

Overall, both methods suggest that between 2010 and 2019, Deer Island mostly changed around the area due to erosional processes with some accretional instances. Both methods also suggest that the most frequently calculated rate/distance is the rates/distance that reflect mild erosion and mild accretion.

Shoreline Analysis for Years 1994-2019

The erosion LRR rates (Figure 3-10, 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) (Figure 3-11, A). The least frequent LRR rate range is the accretion rates greater than 1.0 (m/yr), accounting for 0% of all transects calculated. For the NSM calculations (Figure 3-10, 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, accounting for 1.2% of all transects (Figure 3-11, B). The most frequent NSM distance measurement range is -45.7 to -22.9 meters accounting for 25.6% of all transects calculated (Figure 3-11, B). The largest erosion measurement distance is seen at the north end of Deer Island, while the middle has some areas of accretion and less

64

erosion. However, the south end of Deer Island has some sharp peaks of erosion, however not as high as the north end.

Table 3-3 summarizes the DSAS NSM results where 81.70% of all transects for this time period were calculated as a negative distance identifying shoreline loss. Only 18.29% of the transect resulted in a positive distance. The average distance calculated for each transect is -29.1 meters. The average of all negative distances (all transects where the loss was observed) 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 summarises the DSAS LRR calculations for the 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 with statistically significant erosion is 69.51%, while the percent of all transects with statistically significant accretion is 10.98%.

Beta Shoreline Forecasting Analysis for 10 and 20-Year Prediction

The 10-year prediction (Figure 3-12, A) demonstrates the potential of uniformity of erosion, particularly in the center and south end 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 20-year prediction (Figure 3-12, B) is very similar to the 10-year prediction model, but with more drastic erosion in the north and south. 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.

Discussion

I found that Deer Island shoreline features have eroded over two periods of time, both 1994-2007 and 2010-2019. My results also suggest that the number of transects examined for change on Deer Island had a higher number of transects demonstrating erosional events in 1994-2007 than in 2010-2019. These differences may be a function of environmental factors, including storms, that likely influence change in shoreline features. As an example, the first imagery (1994) is taken one year after a severe winter storm (March 1993) that is known to have caused considerable changes in shoreline features elsewhere in the Big Bend region (Goodbred & Hine, 1993). It is possible that the 1993 storm contributed to several years of erosional events along Deer Island due to destabilizing shoreline features from the storm.

The observed overall net loss of Deer Island shoreline could result in loss of habitat used by wildlife resources, including shorebirds. Vitale et al. (2020) documented erosional events and shoreline migration of islands used as nesting habitats by American Oystercatchers, including islands in the Cedar Keys region. Because Oystercatchers demonstrate high site fidelity and long-lives, Vitale et al. (2020) suggest that the loss of these islands may create a type of ecological trap where birds return to these islands only to have the nests destroyed due to eroded shoreline areas and increased vulnerability to inundation. Many other 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 erosional events, 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).

It is interesting to note that although the overall shoreline experienced erosional events, evidence of accretion may have occurred in the center of the Deer Island shoreline over all time periods (Table 3-3). Using available data, the maximum estimated positive distance gained was 10.91 meters occurring around the central shoreline on transect 44 (Table 3-3). Accretion for our study site may also be a result of extreme meteorological events which could cause sand within the system to be redistributed as sand supply from the Suwannee River (Goodbred et al., 1998) is limited. However, a critical factor is that while accretion in the sand can occur, SLR is continuing to increase in this region. As SLR rise in this region is continuing in this region, and possibly accelerating, (https://tidesandcurrents.noaa.gov/sltrends/sltrends_station.shtml?id=8727520) the interactions between erosion, accretion, and sea-level rise are complicated for this region in predicting how shoreline features will change. Changes in the shoreline features may also contribute to ghost forests (Ury et al., 2021); evidence of vibrant of once vibrant vegetation but now only remnants of that vegetation remain. Ghost forests may provide some shelter for birds and animals, but cannot replace the abundant resources a lively forest can provide.

For these analyses, a source of uncertainty may arise due to the missing imagery years of 2007- 2008 and 2011- 2012. If those missing years were available for analysis, it may provide important information to assist with interpreting the true shoreline migration differences between these two time periods. A source of uncertainty in the

analysis arises from the lack of metadata tidal height information provided by USGS

Earth Explorer, which could be account for in the Shoreline Settings in ArcMap©. 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 uncertainty is in the individual digitization of each shoreline. Since the available

imagery was used for digitizing 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 compared 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. It is also

typical for shorelines be to measured at the mean high water line of the shoreline,

however the not every year of the imagery provides a clear mean high water line to

provide a full analysis.

The prediction models are based on a linear regression rate calculated by DSAS

using a 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 migration is expected (Figure 3-11). This

model predicts that shoreline migration is expected to continue in future years based on

observed erosional event patterns. This prediction assumes that whatever mechanisms

driving the observed shoreline loss (SLR, erosion from storm events) are likely to

continue in the future. Under various climate change scenarios, storm events are

predicted to increase in severity and possibly frequency (Knutson et al., 2020), altering the rate of erosion.

Sea level rise may be the dominant feature driving changes in shoreline features along Deer Island. While erosion and accretion are indeed occurring, sea level has been monitored in Cedar Key for more than 100 years (NOAA station 8727520, https://tidesandcurrents.noaa.gov/sltrends/sltrends_station.shtml?id=8727520) and the long-term sea-level rise is about 2.23 mm/yr. This rate has shown an increase since 2010 (https://tidesandcurrents.noaa.gov/sltrends/sltrends_station.shtml?id=8727520), and an increasing rate of SLR would likely lead to an increased rate of shoreline inundation and loss.

This study has revealed brief historical trends of coastal evolution along an undeveloped sandy shoreline. The shoreline statistics revealed greater meters of shoreline migration during the first time frame 1994-2007, possibly due to a significant hurricane impact. Storm and storm clusters may significantly impact barrier island morphology in this area with direct stormwind 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 migration and accretion (Sankar et al., 2018). This research has demonstrated that sandy shorelines in this area may be susceptible to more erosional events rather than accretion due to the factors mentioned, which may ultimately lead to multidecadal shoreline loss.
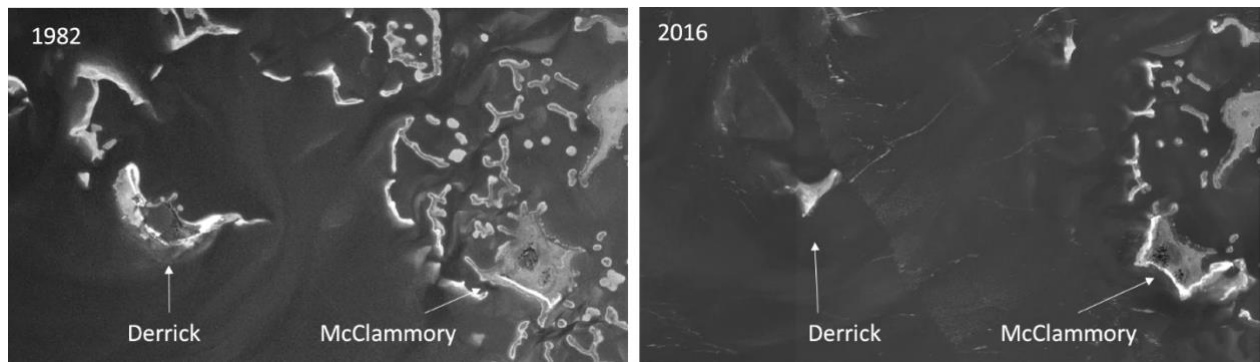
Figure 3-1. Island degradation of Derrick Key in the Cedar Keys, Florida from 1982
          (left) to 2016 (right), (Vitale, 2019).
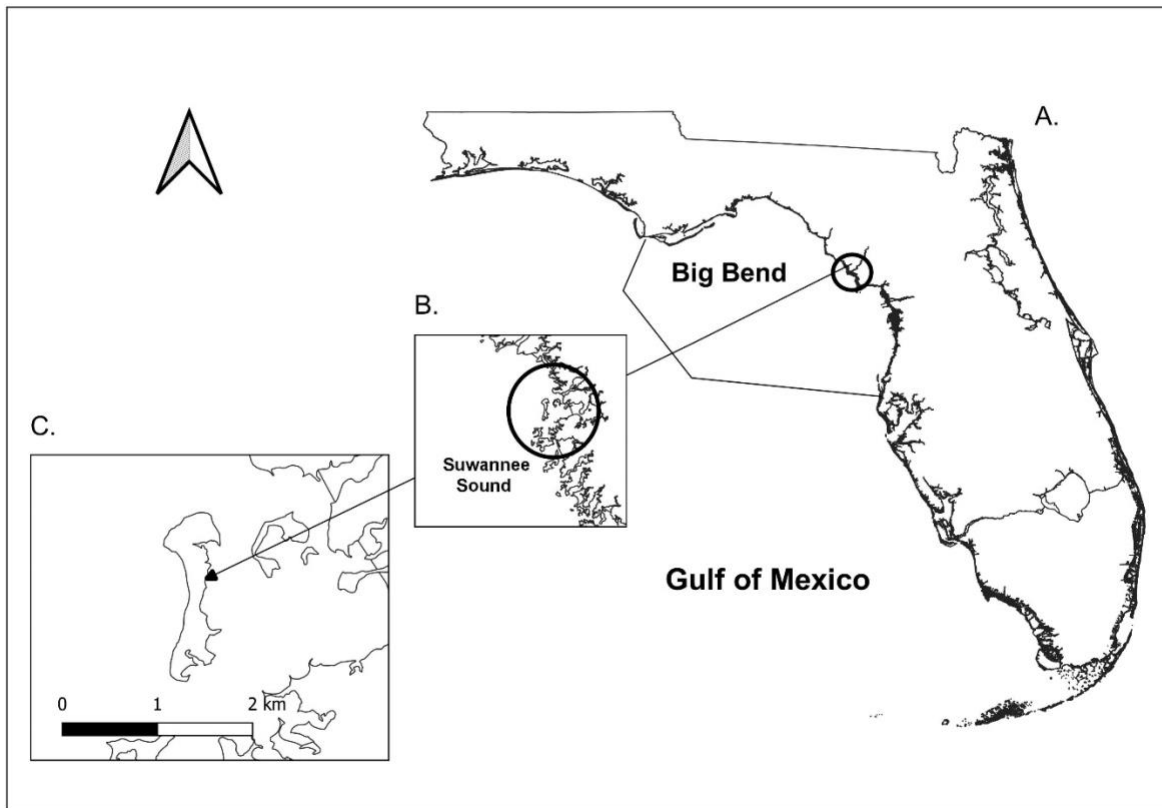
Figure 3-2. Location of Deer Island, Florida. A) Map of the entire state of Florida; B)
          Zoomed in study site; C) Zoomed in area of Deer Island with a scale bar in
          kilometers. Shoreline shapefile downloaded from National Geophysical Data
          Center, 2017. Shoreline in 1:2 million scale.
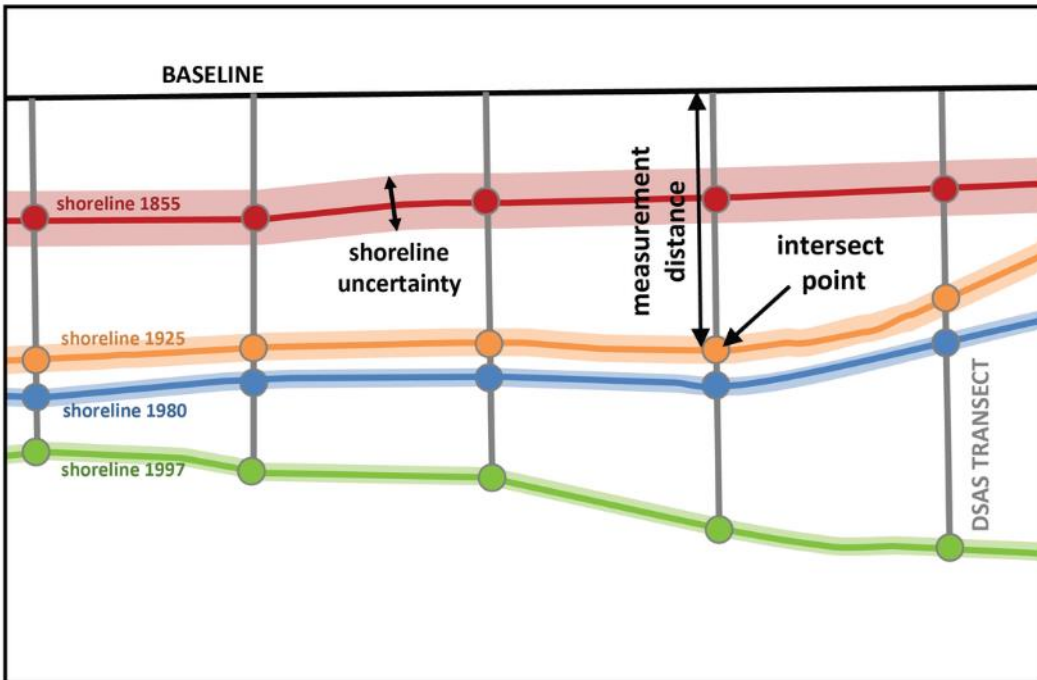
Figure 3-3. Example of DSAS transect casting
(https://www.usgs.gov/centers/whcmsc/science/digital-shoreline-analysis-system-dsas?qt-science_center_objects=0#qt-science_center_objects)
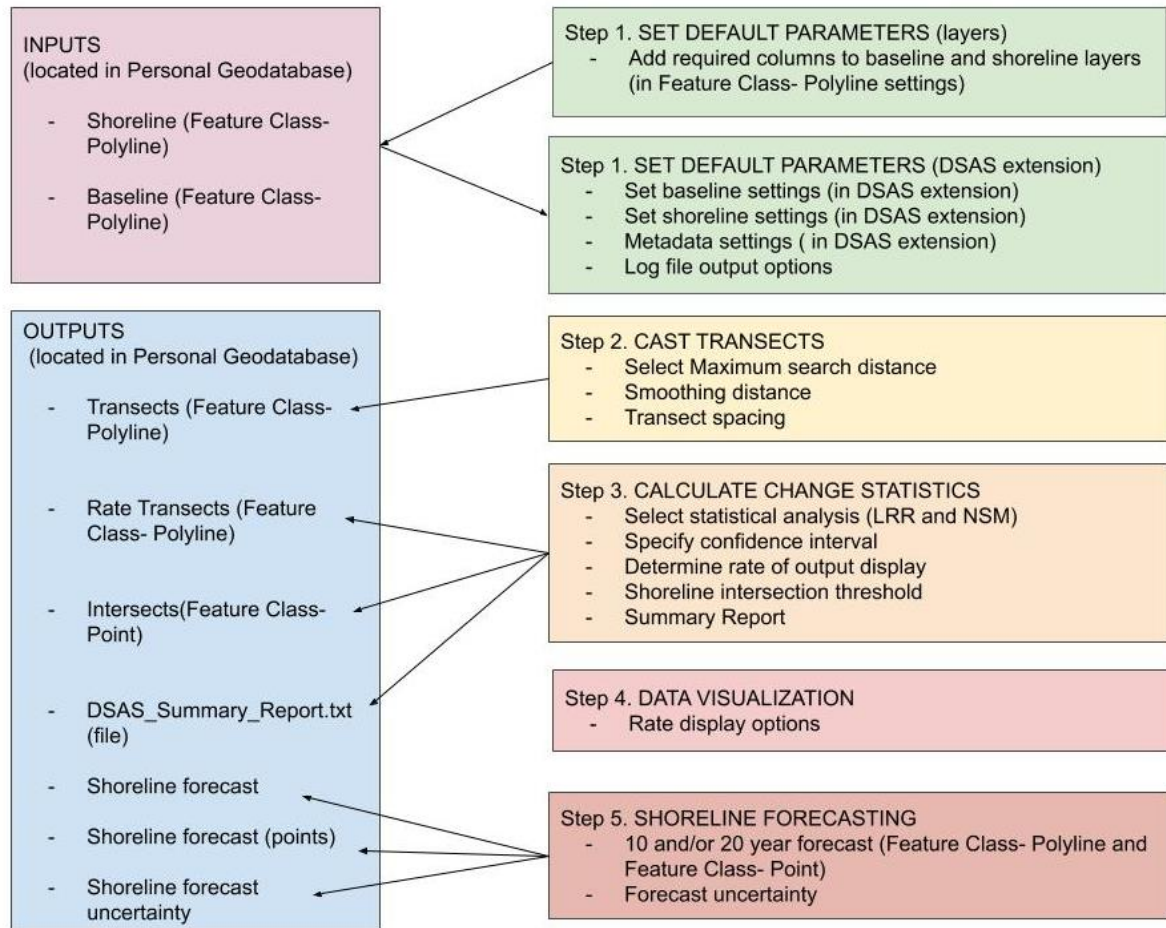
Figure 3-4. Modified DSAS components and operational workflow ("Digital Shoreline Analysis System (DSAS) Version 5.0 User Guide.", 2021).
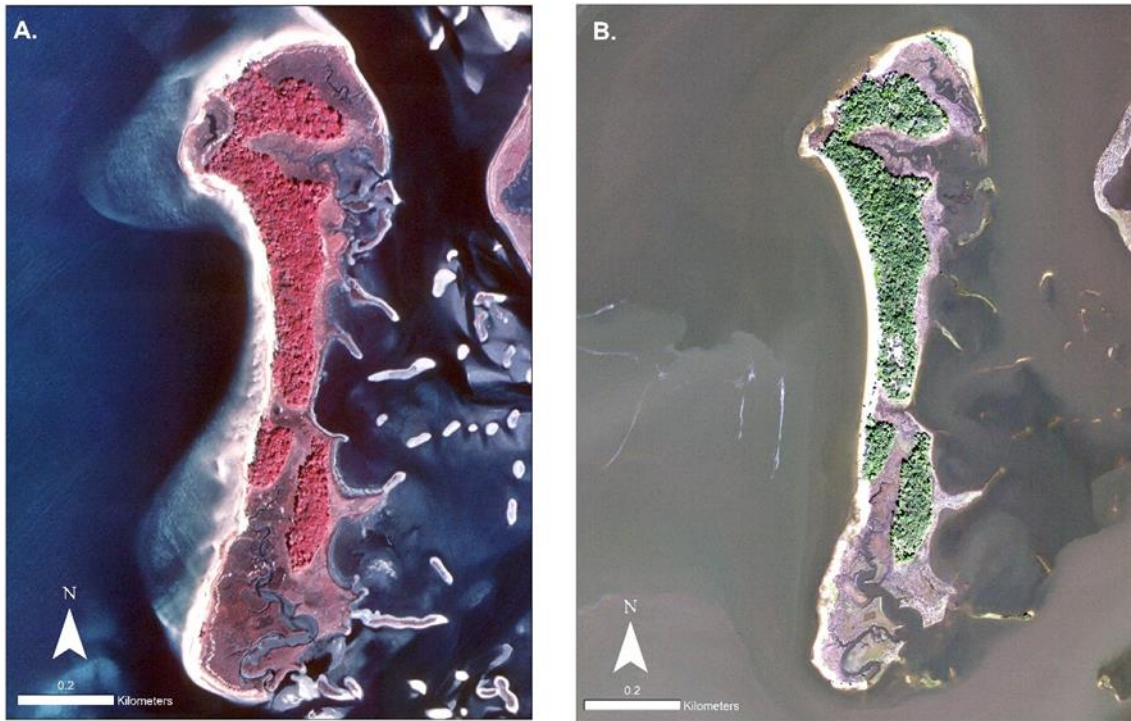
Figure 3-5. NAIP aerial photographs, map scale 1:3,000, and projection NAD 1983 UTM, Zone 17N; A) 1994 aerial imagery, max wind speed of 19.31 KPH and a median river discharge of 274.9 m^3/s on the day the imagery was captured (Table 3-1); B) 2019 aerial imagery aerial imagery, max wind speed of 11.27 KPH and a median river discharge of 146.9 m^3/s on the day the imagery was captured (Table 3-1). The time of day when the images were taken are not accessible; tidal height information is not available.

A.

LRR (m/yr) for years 1994 to 2007

LRR
- ── -5.0 < LRR <= -3.0
- ── -3.0 < LRR <= -2.0
- ── -2.0 < LRR <= -1.0
- ── -1.0 < LRR <= -0.5
- ── -0.5 < LRR <= 0.5
- ── 0.5 < LRR <= 1.0
- ── 1.0 < LRR <= 2.0
- ── 2.0 < LRR <= 3.0
- ── 3.0 < LRR <= 4.0

N
W ── E
S

0 0.1 0.2   0.4 Kilometers

B.

NSM (meters) for years 1994 to 2007

NSM
- ── -70.5 < NSM <= -35.9
- ── -35.9 < NSM <= -23.9
- ── -23.9 < NSM <= -12.0
- ── -12.0 < NSM <= -10.0
- ── -10.0 < NSM <= 2.2
- ── 2.2 < NSM <= 4.3
- ── 4.3 < NSM <= 6.4

N
W ── E
S

0 0.1 0.2   0.4 Kilometers

Figure 3-6. Results of DSAS model results for the years 1994 to 2007. The transects (n= 82) display where shoreline erosion and accretion have been modeled. The shorelines (green) include digitized shorelines for the years 1994 to 2007 in this figure. A) Linear Regression Rates model (m/yr) displaying transects with legend colors corresponding to Figure 3-8, A. B) Net Shoreline model (meters) displaying transects with legend colors corresponding to Figure 3-8, B.

Figure 3-7. Figure of the DSAS statistics for the years 1994 to 2007. The total amount of transects calculated by DSAS is 82. Results are displayed in the percentage of transects in each bin where erosion is negative x-values and accretion is positive x-values. A) LRR model results with bar colors corresponding to Figure 3-7, A, B) NSM model results with bar colors corresponding to Figure 3-7, B.

A.

LRR (m/yr) for years 2010 to 2019

LRR
- ── -4.0 < LRR <= -3.0
- ── -3.0 < LRR <= -2.0
- ── -2.0 < LRR <= -1.0
- ── -1.0 < LRR <= -0.5
- ── -0.5 < LRR <= 0.5
- ── 0.5 < LRR <= 1.0
- ── 1.0 < LRR <= 2.0
- ── 2.0 < LRR <= 3.0
- ── 3.0 < LRR <= 4.0

0   0.1   0.2   0.4 Kilometers

B.

NSM (meters) for yeras 2010 to 2019

NSM
- ── -41.8 < NSM <= -20.1
- ── -20.1 < NSM <= -13.4
- ── -13.4 < NSM <= -6.7
- ── -6.7 < NSM <= 2.9
- ── 2.9 < NSM <= 5.8
- ── 5.8 < NSM <= 8.7
- ── 8.7 < NSM <= 9.9
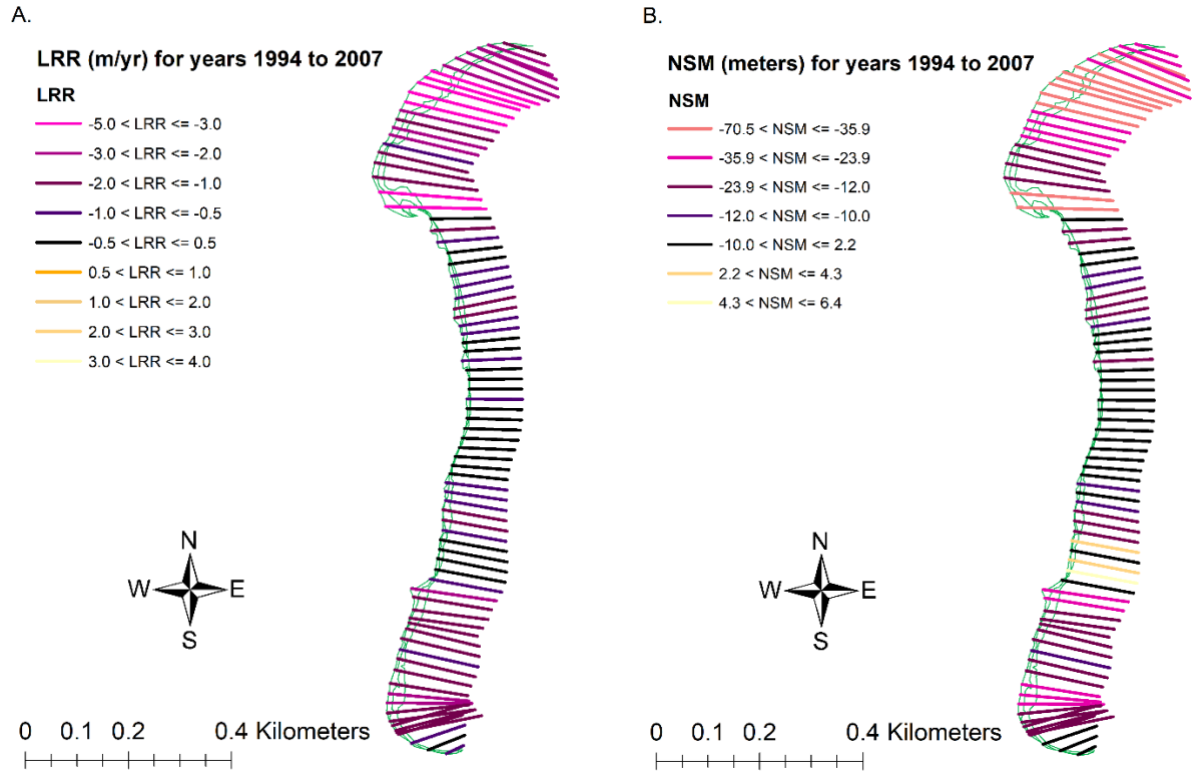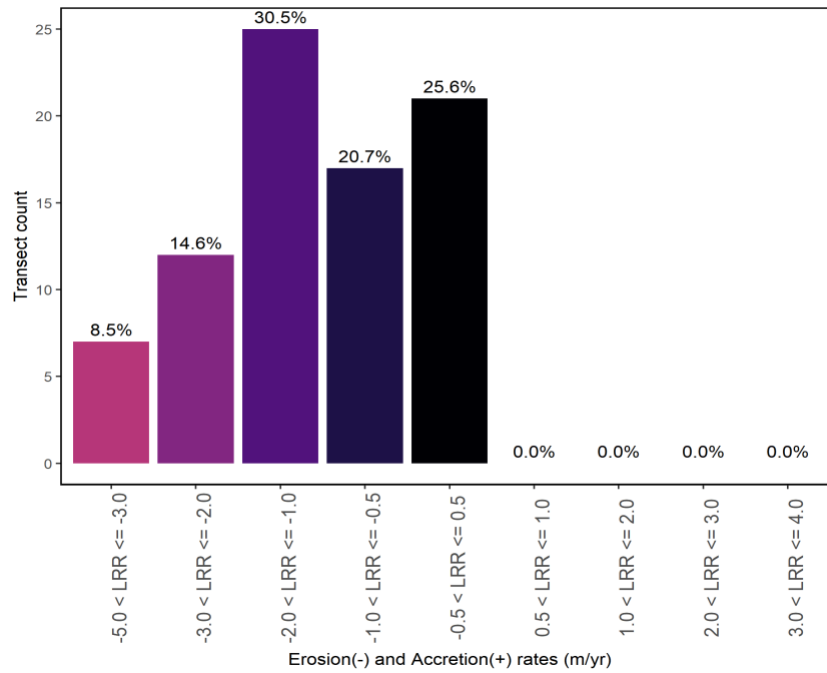
0   0.1   0.2   0.4 Kilometers

Figure 3-8. Results of DSAS model results for the years 2010 to 2019. The transects (n= 82) display where shoreline erosion and accretion have been modeled. The shorelines (green) include digitized shorelines for the years 2010 to 2019 in this figure. A) Linear Regression Rates model (m/yr) displaying transects with legend colors corresponding to Figure 3-10, A. B) Net Shoreline model (meters) displaying transects with legend colors corresponding to Figure 3-10, B.
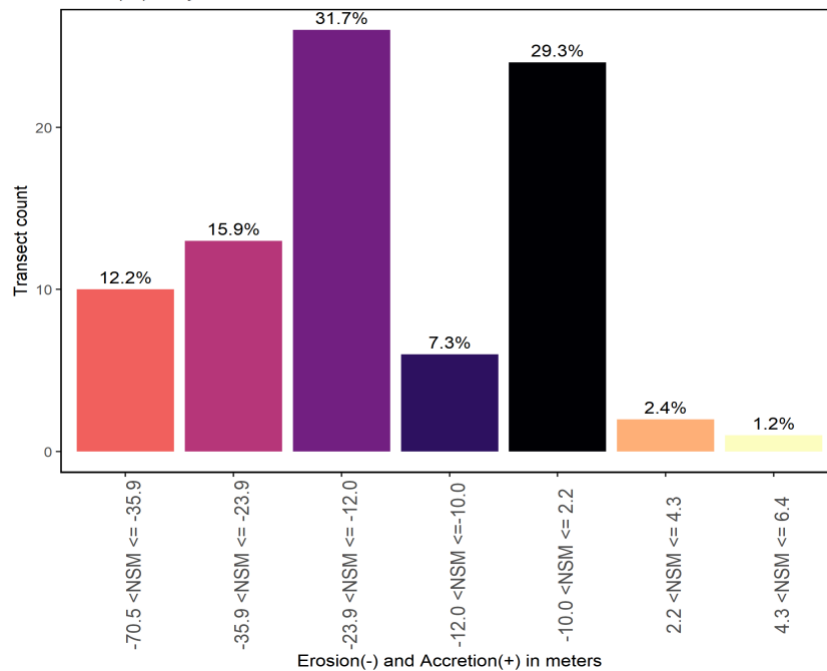
Figure 3-9. Figure of the DSAS statistics for the years 2010 to 2019. The total amount of transects calculated by DSAS is 82. Results are displayed in the percentage of transects in each bin where erosion are negative x-values and accretion are positive x-values. A) LRR model results with bar colors corresponding to Figure 3-9, A, B) NSM model results with bar colors corresponding to Figure 3-9, B.
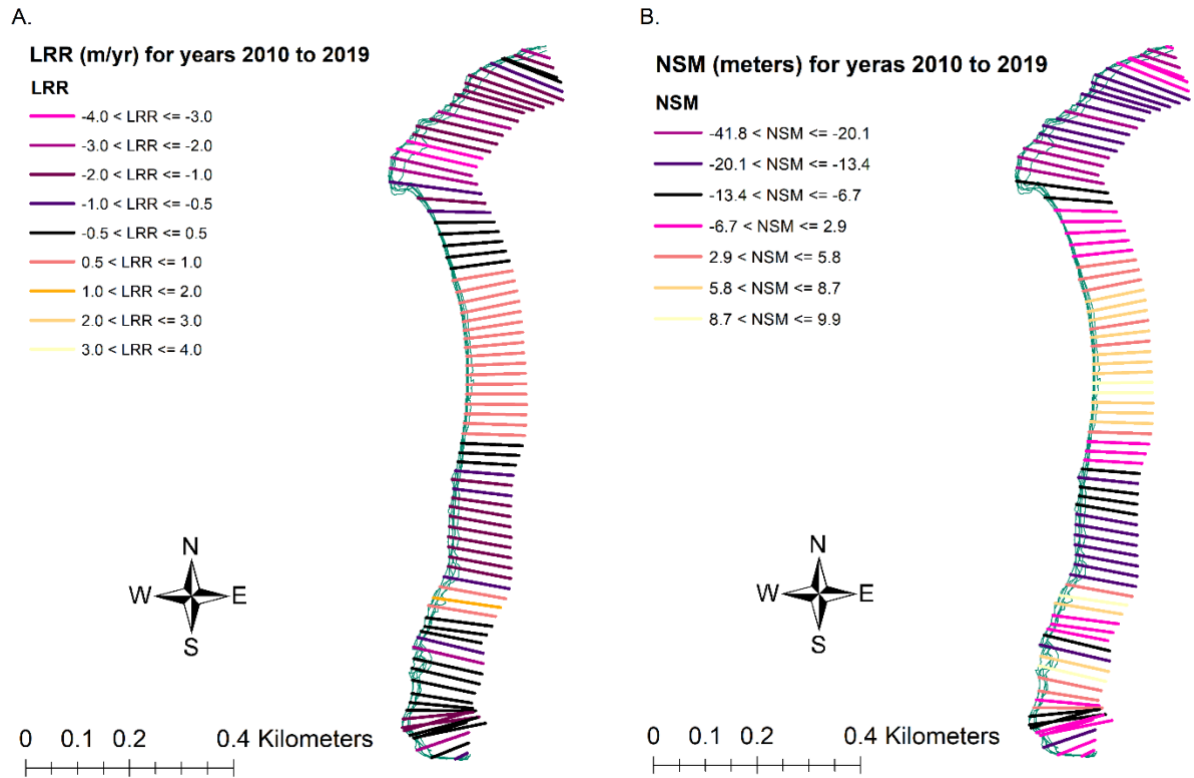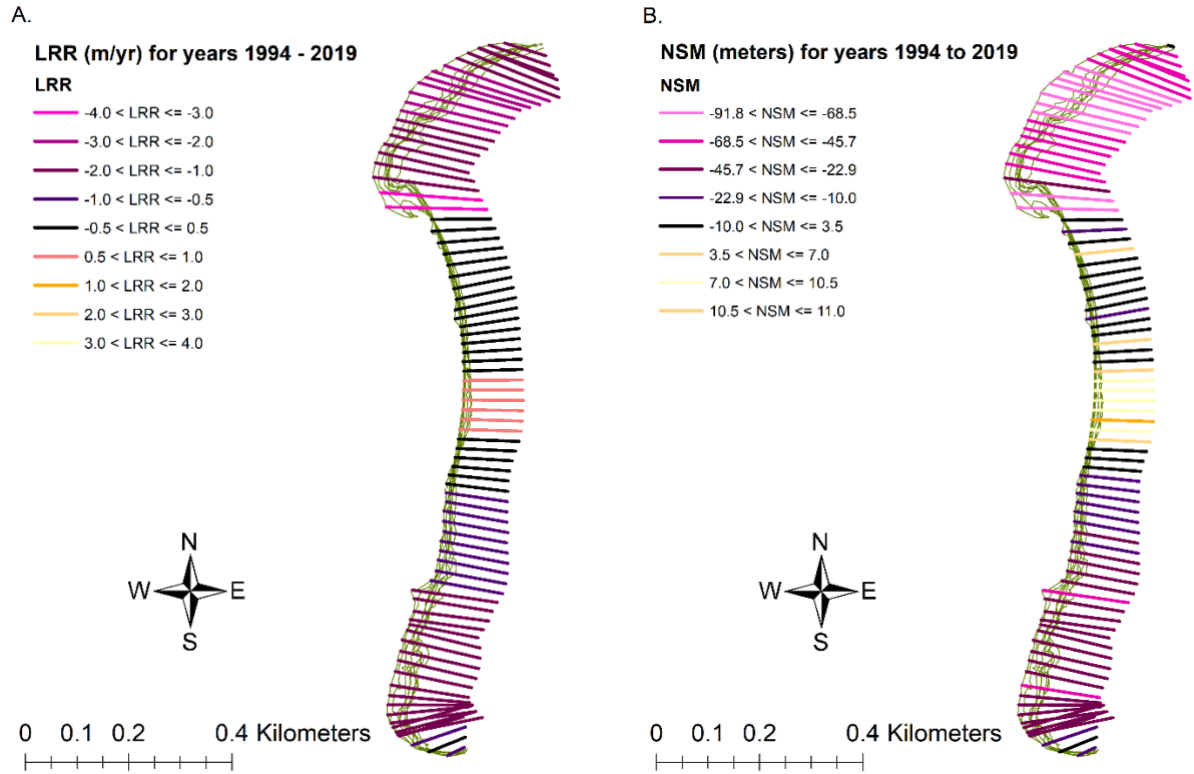
Figure 3-10. Results of DSAS model results for years 1994 to 2019. The transects (n= 82) display where shoreline erosion and accretion have been modeled. The shorelines (green) include digitized shorelines for the years 1994 to 2019 in this figure. A) Linear Regression Rates model (m/yr) displaying transects with legend colors corresponding to Figure 3-12, A. B) Net Shoreline model (meters) displaying transects with legend colors corresponding to Figure 3-12, B.
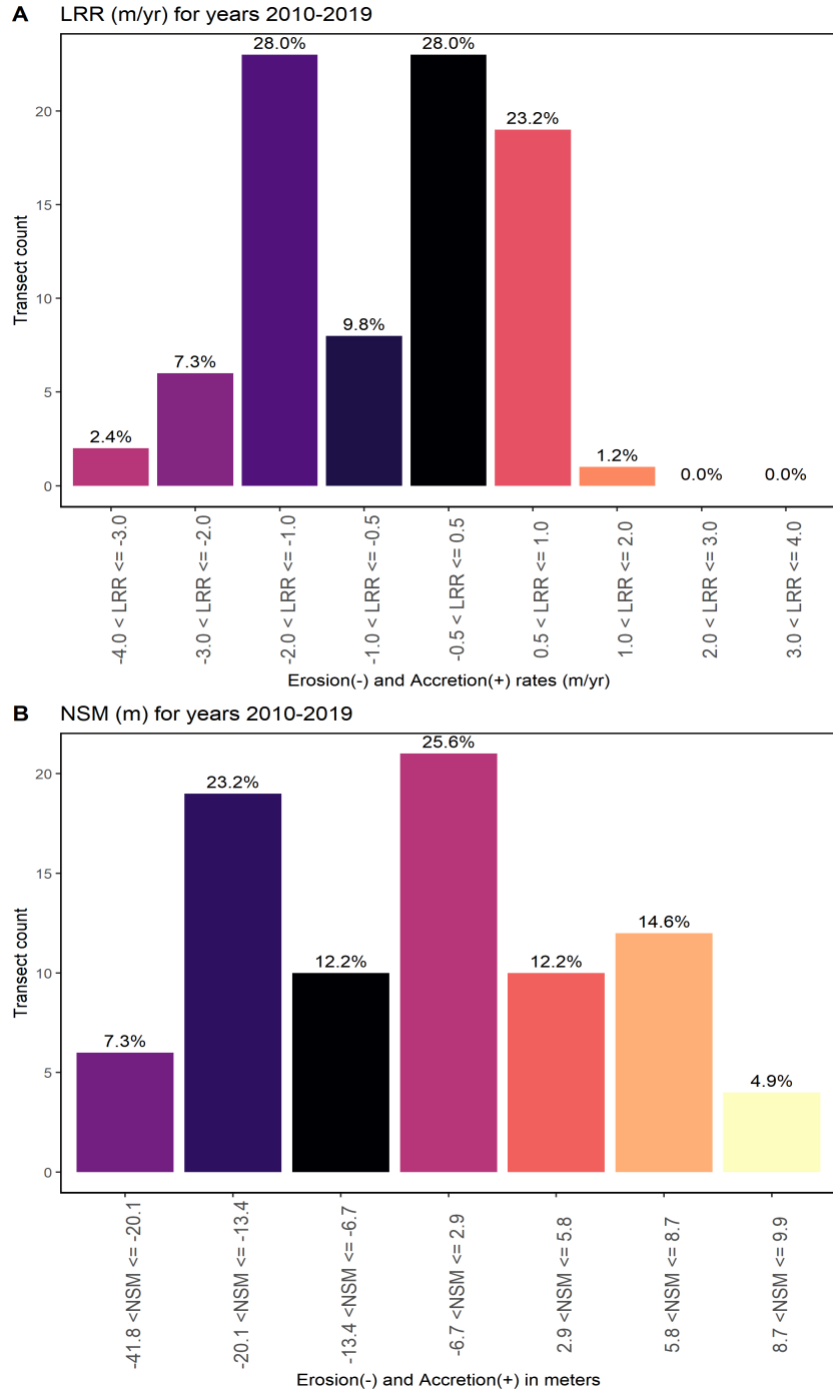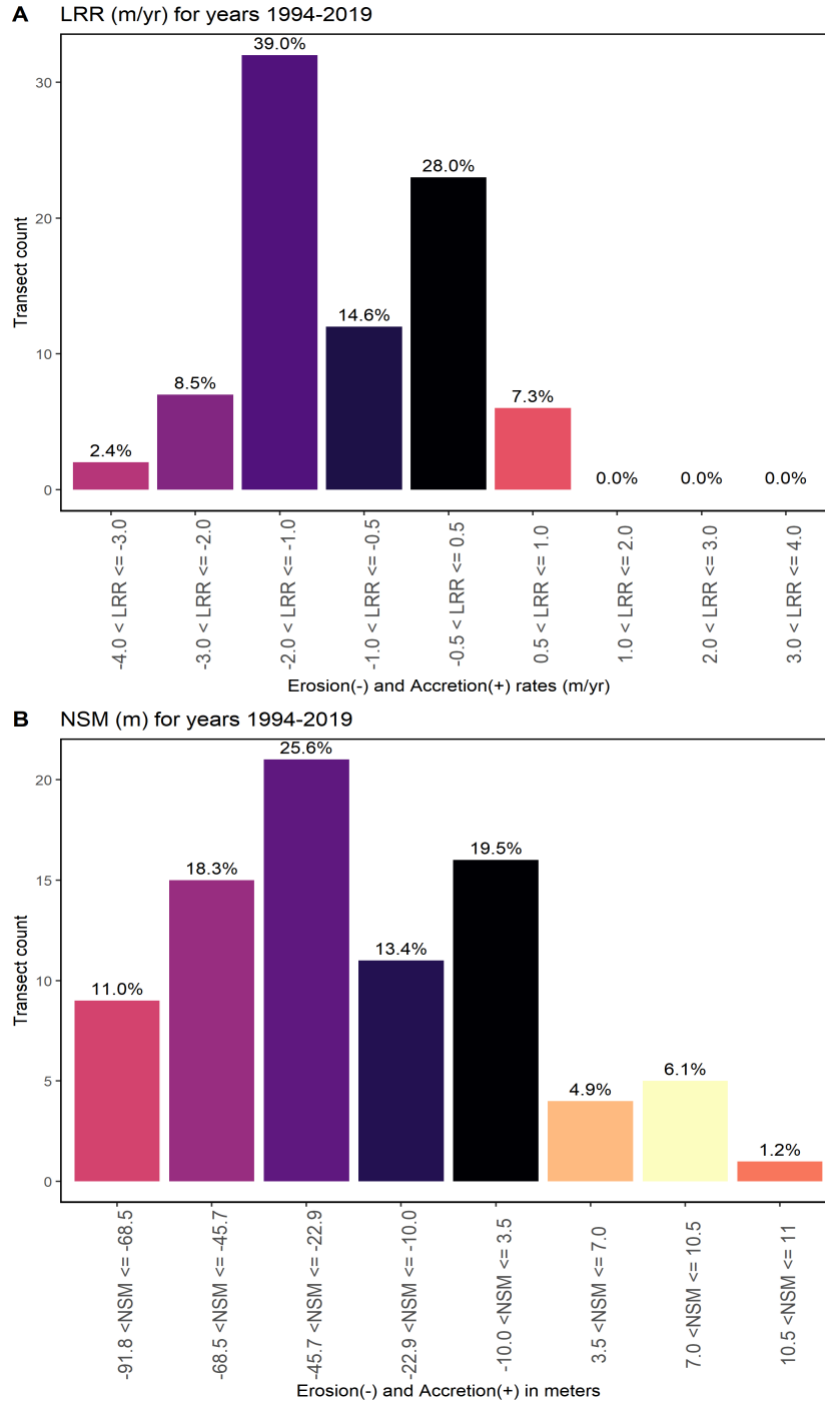
Figure 3-11. Figure of the DSAS statistics for the years 1994 to 2019. The total amount of transects calculated by DSAS is 82. Results are displayed in the percentage of transects in each bin where erosion is negative x-values and accretion is positive x-values. A) LRR model results with bar colors corresponding to Figure 3-11, A, B) NSM model results with bar colors corresponding to Figure 3-11, B.

A.
2019 Aerial
Photograph
with 10-year
forecast overlay

N
W E
S

0 0.05 0.1 0.2 Kilometers
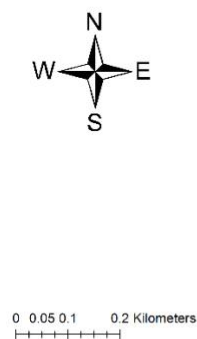
B.
2019 Aerial
Photograph
with 20-year
forecast overlay

N
W E
S

0 0.05 0.1 0.2 Kilometers
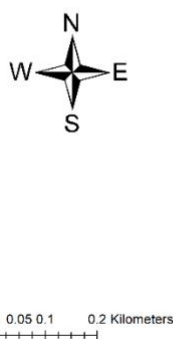
Figure 3-12. DSAS shoreline prediction forecast. Modeled Shorelines are located on the west side of each panel. A) Shoreline forecast for a 10-year prediction (thick black line) and its uncertainty (yellow shaded region) overlayed aerial imagery (2019) to display the predicted shoreline loss in comparison to the latest imagery selected. B) Shoreline forecast for a 20-year prediction (thick black line) and its uncertainty (yellow shaded region) overlayed aerial imagery (2019) to display the predicted shoreline loss in comparison to the latest imagery selected.

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. 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	Entity ID: DI00000000018672 (found in DOQ)	
		Precipitation (cm)- 0.00	Map Projection	UTM
		Max Wind Speed	Projection Zone	17N
		(KPH)- 19.31	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	Entity ID: DI00000001164809 (found in DOQ)	
		Precipitation (cm)- 0.00	Map Projection	UTM
		Max Wind Speed	Projection Zone	17N
		(KPH)- 25.75	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	Entity ID: N_2908356_NW_17_1_20071102 (found in NAIP)	
		Precipitation (cm)- 0.00	Map Projection	UTM
		Max Wind Speed	Projection Zone	17N
		(KPH)- 22.53	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	Entity ID: M_2908356_NW_17_1_20100919 (found in NAIP)	
		Precipitation (cm)- 0.00	Map Projection	UTM
		Max Wind Speed	Projection Zone	17N
		(KPH)-24.14	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. Color band and computer channel information to attain true color (National Agriculture Imagery Program, 2017)

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

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 a 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

CHAPTER 4
CONCLUSION

Data management workflows are imperative to every ecological project. Ecological projects may be prone to mismanaged data and file organization because many projects change their focus as hypotheses evolve, instrumentation changes, and staff turnovers occur. Data management workflows increase efficiency, reproducibility, and transparency within ecological projects by requiring project administrators to think through and document all steps starting from the start of data collection to the end of the workflow, including analysis and reporting while considering what data types will be collected and what frequency help guide data management workflows creation.

Ecological project workflows should include a way to save iterations of changing project files and data. Using version control software like GitHub to store data and project files increases reproducibility and transparency in ecological projects by having an online location where collaborators can view and contribute to project files. Version control workflows are a standard in many disciplines and these workflows benefit ecological projects in the same way other than disciplines by tracking and managing project file changes.

The analysis in Chapter 3 follows the approach of creating and following a workflow to create a reproducible study for the study area and additional shorelines. Geographic geodatabases require their workflow to store the geographic data because all shapefiles used to create the map need to be located within the geodatabase. The workflow of Chapter 3 was designed by USGS, but additional steps were created for naming and storing the geodatabase. The study is organized with a geodatabase which contains standardized naming conventions as described in Chapter 2. Standardizing a

naming convention for the maps and shapefiles, as described in Chapter 2, will help team members locate files and be able to identify files without having to click on each one while keeping maps and shapefiles names organized and standardized. Saving finalized maps in a version control environment will help investigators share their maps with other researchers by providing a centralized place with finalized maps available online (e.g., through GitHub).

Developing a proactive approach to creating workflows for the data collection, data types, and various file types for a project is likely much easier than assembling a workflow retrospectively. As mentioned in Chapter 1, it is essential to consider the workflow dynamic of all employees in the project, what files/data they will be handling, where the files/data will be located, and how to track and manage those changes. The methods and suggestions in Chapters 1 and 2 may inspire investigators to plan out all aspects of data collection from start to finish. Planning these steps will also help with project requirements from funding agencies who expect the research investigators to create a blueprint of the research proposed and research processes. The ability to continue doing the same analysis over a period of time is precisely the type of reproducible stance that should be encouraged in ecological projects.

The Big Bend region of Florida has experienced significant changes in recent decades and observed changes such as loss of named islands, including Derrick Key (Vitale et al., 2020), erosion of Deer Island as shown in this study, large scale losses of oyster reefs (Seavey et al.,2011; Moore et al., 2020), and increasing rate of SLR. Restoration efforts for shoreline habitats are ongoing, as demonstrated by the Lone Cabbage Reef project, with more restoration efforts proposed. My thesis demonstrates

how the Big Bend region is changing (Chapter 3) and how a data workflow and

management system can be used to inform a restoration program to facilitate the

restoration of coastal features (Chapters 1 and 2). In this way, my hope is that future

restoration decisions can be informed both by restoration needs such as documenting

change in Deer Island and then effective on-the-ground implementation and evaluation

of the restoration through sound data science practices.

LIST OF REFERENCES

Adams, P. N., Keough, K. M., & Olabarrieta, M. (2016). Beach Morphodynamics influenced by an ebb-tidal delta on the north Florida Atlantic coast. Earth Surface Processes and Landforms, 41(7), 936-950.

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

Berg, R. (2017). National Hurricane Center Tropical Cyclone Report. Hurricane Hermine. Miami: National Hurricane Center.

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

Brown, S., Nicholls, R. J., Woodroffe, C. D., Hanson, S., Hinkel, J., Kebede, A. S., & Vafeidis, A. T. (2013). Sea-level rise impacts and responses: a global perspective. In Coastal hazards (pp. 117-149). Springer, Dordrecht.

Bureau of Economic and Business Research, Population Studies Program, 2021, 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

Cahoon, D. R., & Guntenspergen, G. R. (2010). Climate change, sea-level rise, and coastal wetlands. National Wetlands Newsletter, 32(1), 8-12.

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.

Curtis, K. J., & Schneider, A. (2011). Understanding the demographic implications of climate change: estimates of localized population predictions under future scenarios of sea-level rise. Population and Environment, 33(1), 28-54.

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

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

Elko, Nicole, et al. "A century of US beach nourishment." Ocean & Coastal Management 199 (2021): 105406.

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, Steven L., and Albert C. Hine 1995 Coastal Storm Deposition: Salt-marsh Response to a Severe Extratropical Storm, March 1993, West-Central Florida. Geology 23(8):679-682.

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

Habel, S., Fletcher, C. H., Anderson, T. R., & Thompson, P. R. (2020). Sea-Level Rise induced Multi-Mechanism flooding and contribution to Urban infrastructure failure. Scientific reports, 10(1), 1-12.

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

IPCC 4th Assessment Report (2007), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M.Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Cambridge Univ. Press, Cambridge, U. K.

IPCC, 5th Assessment report (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by T.F.Stocker, D.Qin,G.-K. Plattner, M.Tignor,S.K.Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Cambridge Univ. Press, Cambridge, U. K., New York, USA in press, available at http://www.ipcc.ch/report/ar5/wg1/#.UsHKO_tW-Ag.

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.

Knutson, T., Camargo, S. J., Chan, J. C., Emanuel, K., Ho, C. H., Kossin, J., ... & Wu, L. (2020). Tropical cyclones and climate change assessment: Part II: Projected response to anthropogenic warming. Bulletin of the American Meteorological Society, 101(3), E303-E322.

Kopp, R. E., Gilmore, E. A., Little, C. M., Lorenzo-Trueba, J., Ramenzoni, V. C., & Sweet, W. V. (2019). Usable science for managing the risks of sea-level rise. Earth's future, 7(12), 1235-1269.

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

Lyons, Dylan. "Tropical Storm ETA Had Little Impact on Cedar Key." Https://Www.wcjb.com, www.wcjb.com/2020/11/12/tropical-storm-eta-had-little-impact-on-cedar-key/.

Marchand, M. (2010). Concepts and science for coastal erosion management.

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

Mimura, N. (2013). Sea-level rise caused by climate change and its implications for society. Proceedings of the Japan Academy, Series B, 89(7), 281-301.

Monés, Micah P., Neill J. Wallis, and Kenneth E. Sassaman 2012 Archaeological Investigations at Deer Island, Levy County, Florida. Technical Report 15. Laboratory of Southeastern Archaeology, Department of Anthropology, University of Florida, Gainesville.

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.

National Agriculture Imagery Program (NAIP) aerial imagery band wavelength ranges in units (µm), 2017. https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/APFO/support-documents/pdfs/fourband_infosheet_2017.pdf, (Accessed on April 21, 2021).

National Geophysical Data Center (NGDC) produced by NOAA Satellite and Information Service, 2017, http://rimmer.ngdc.noaa.gov/mgg/coast/getcoast.html, (Accessed April 21, 2021)

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

Seavey, J. R., Pine III, W. E., Frederick, P., Sturmer, L., & Berrigan, M. (2011). Decadal changes in oyster reefs in the Big Bend of Florida's Gulf Coast. Ecosphere, 2(10), 1-14.

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

Ury, E. A., Yang, X., Wright, J. P., & Bernhardt, E. S. (2021). Rapid deforestation of a coastal landscape driven by sea-level rise and extreme events. Ecological applications, e02339.

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.

Wadhams, P., N. Hughes, and J. Rodrigues (2011), Arctic sea ice thickness characteristics in winter 2004 and 2007 from submarine sonar transects, J. Geophys. Res.,116, C00E02, doi:10.1029/2011JC006982.

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.

Williams, K., Ewel, K. C., Stumpf, R. P., Putz, F. E., & Workman, T. W. (1999). Sea-level rise and coastal forest retreat on the west coast of Florida, USA. Ecology, 80(6), 2045-2063.

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

Melissa Moreno was born and raised in Miami, Florida. She graduated from the University of Florida with a Bachelor of Science in Wildlife Ecology and Conservation in 2017. She is in her third year of Master of Science program with SNRE at the University of Florida. She graduated from the University of Florida with a Master of Science degree in Interdisciplinary Ecology, focusing on forestry and data management in 2021. Her most significant and relevant work experience is working for the Lone Cabbage Reef restoration project as a data steward. After graduation, Melissa plans to take time off and enjoy artistic ventures.