Draft ER390 Project Proposal

Vizualizing HAT's GOE Monitoring Data

Wendy Anthony

2025/09/04

Contents

The Idea behind submitting this project proposal

- I am a student in UVic's Restoration of Natural Systems program, as well as having the equivalent of a double major in Environmental Studies and Geography, and being a life-long learner with an interest in making a significant contribution to the local restoration community, and I'm looking to assist a willing community partner through the creation of my Final RNS Restoration Project.
- Nancy Shackelford (RNS Director) suggested that I might consider helping HAT GOE Monitoring Project of Southern Vancouver Island by creating a tool for vizualizing GOE Monitoring Data.
- I propose starting this project in the Fall term, and while HAT is preparing to share data later in October, I could continue on the coding work I started in August (over 35 hours logged so far), and create automatic code function snippets that can be reused by just changing the choice of variable, which could then be incorporated into an R Package. When HAT is ready for me in October, I should be able to find ways to use R code to help clean their data. See Proposed Gantt Timeline Chart

Background

- Data used for this proposal is from two tables in the report that HAT used as a model for their GOE Monitoring Project (Malloff & Shackelford, 2024). In order to compare monitoring results from multiple restoration projects, consistent measurements of vegetation composition (e.g. native and exotic species cover) may help identify responses to the stressors affecting already compromised and rare Garry Oak Ecosystems. Representing data in a visual manner can help make complex data more understandable than a table of numbers, and may reveal patterns or trends resulting from different restoration management actions (Malloff & Shackelford, 2024).
- Data visualization can represent abstract information visually, allowing for data exploration and analysis, modelling restoration management alternatives, data-based story-telling, and communicating potential insights, to help understand and make restoration management decisions. Using RMarkdown to generate reports helps create research methods that are reproducible (McKay, 2020).

Project Partner

- I have already spoken to Vanessa Brownlee, Habitat Restoration Coordinator at Habitat Acquisition Trust (HAT) on 2025-08-27, who was very interested in my proposed project idea, and Vanessa seemed willing to consider letting me start in Fall 2025.
- I have also already spoken briefly to Board Chair Andy MacKinnon at the HAT Social 2025-08-27, and he seemed keen to know I was interested in doing my RNS Final Project with HAT and was looking forward to know more about what I had in mind.
- I am prepared to be guided by HAT's preferences as I offer my help to create a tool to automate some data indices calculations and visualization that would hopefully prove to be useful, and adaptable, to HAT's long-term monitoring programs.

Project Goals and Overall Purpose

- I propose to create a customized, automatic coding and reporting tool, using open-source tools R package libraries, which can be used by non-coders at HAT for data formatting and interactively visualizing data collected at multiple sites for their Garry Oak Ecosystem (GOE) Monitoring Program, including tutorial instructions, to determine what stories the collected data may illustrate, as an easier way for communication data information than looking at tables of numbers.
- This project proposal has been completed using code in RMarkdown, as a showcase of the types of visualizations tools and reports that I propose to offer HAT, adapting the code to be based on HAT monitoring data.

Project Objectives (SMART actionable targets)

- See Work flow Flowchart, Figure 1, and Gantt Chart Proposed Timeline, Figure 2)
- **Analyze the collected data** to understand the underlying structure, while ensuring the use of consistent naming protocols for variables, and formatting data to be usable for use with data visualization code.
- **Write data visualization code**, using RStudio, RMarkdown, and RShiny, to create appropriate visualizations to enable a better understanding of the collected data.
- **Customize data tools** to reflect a consistent style that uses HAT's colours, logos, and plotting themes.
- **Create RMarkdown documents**, with reusable code, useable without needing to know how to code to automatically download and format data, create data visualizations, and produce reports for the whole project and individual sites.
- **Create scaleable data visualization tools** that can have more sites added, with different measuring and sensitivity criteria (e.g. First Nations GOE sites), as well as multi-year data.
- **Track differences between individual sites** using variables such as native species, exotic species, and other metrics or indices that HAT measures
- **Create custom R functions** for reading, analyzing, and plotting data, which could be incorporated into a customized R package for HAT.
- **Create tools for next season use** to ensure consistent use of data variable names, and to incorporate any lessons learned in the process of working with it data in year one.
- **Create an interactive, web-browser-based Shiny App** for user-chosen, custom-make data visualization plots and maps for selected location(s) and variables, with the ability to save and download the resulting data, html widgets, plot images, interactive web maps, and perform statistical analysis.
- **Recognize the sensitive and confidential nature of First Nations data** by, after receiving consent, creating separate data visualization code, that may also require using different variables depending on the individual land project goals.

Project Methods for HAT Data visualization

- See Work flow Flowchart, Figure 1, and Gantt Chart Proposed Timeline, Figure 2)
- Continue to research and learn how to code in Shiny reactively in order to make interactive data choices (e.g. a drop-down list of Sites) respond to a prior data choice (e.g. Subregions), in order to filter the results of specific interest. I'd like to extend these interactive choice reactions to the selection of monitoring variables being measured in order to be able to make custom-made visualizations and analysis with only a few clicks
- Sign any data confidentiality agreements as may be required by HAT
- Consult with HAT restoration monitor team to explore and understand the nature of the collected data, data format (forms, spreadsheets), the variables that are being measured, any indices to be calculated using the collected data, what questions they would like to explore with the data
- Create a Project folder, with subfolders for R scripts, raw data, processed data, images, reports, download a copy of the current data (including any meta data e.g. measurement units, description of measurement variable, date collected, etc)
- Analyze and clean data, check for errors, inconsistencies, data quality, calculate indices values from raw data measurements
- create file and variable naming protocols
- Create data visualization plot charts that appropriate to reflect the nature of the variables being measured
- Documents my methods, successes, and failures in the process of the project
- Continual research for finding appropriate R coding methods, and to contribute to writing my final paper

Proposed Deliverables

- See Work flow Flowchart, Figure 1, and Gantt Chart Proposed Timeline, Figure 2)
- All of the following deliverables will be created using RStudio

Data Visualizations

- See example visualization results in next section
- data tables, chart plots, interactive maps
- for individual sites
- for comparing sites (all together, or grouped by subregion, or separately by First Nations if they choose to participate)
- statistical analysis, and data summary stats (depending on HAT goals)

RMarkdown Data Processing Document

- for data downloading, cleaning, and formatting data visualizations
- explain process for preparing data for use
- can be used to create HTML, pdf, or Word reports, with table of contents, plot output images, and even code if desired
- HTML Document includes a linked Table of Contents for easy navigation between sections

Customizations for Plot-Charts, Reports, and Shiny App

- colour palette for background app
- change default colour palette in data vizualization plots from pink and blue to some other colour appropriate to HAT and the data
- add HAT logo
- add colour to Shiny tabs
- create custom HAT plot chart theme, including colours, backgrounds, etc
- Individual site reports

Shiny App Interactive Data Viz Tool

- https://goe-interact.shinyapps.io/Test-Shiny-GOE-Monitor/
- using data processed with RMarkdown document
- for choosing data variables to visualize in tables, plot charts, and maps

Tutorials for using Data visualization Tools

- easy to understand
- customized

Create Report Templates

- for customized individual site reports, as well as total project reports, and others as requested by HAT

Best practises for file naming, variable, naming, etc.

- ensure consistency in naming files, variables, file folder structure

Write R program for Hat's requirements

- Given time, and HAT's interest, I would be prepared to create a HAT-specific R program library, with all the functions required to create the data visualizations and reports for HAT's monitoring program

Presentation

- The completed project will be presented to my community partner HAT
- The presentation would include a slide show, live demonstrations of Data Visualization Tools, and a draft copy of my final written report, with the final approved copy sent to HAT when finished and ready to be published.

Written Report

- The final RNS report will document the work I did to complete this project
- Report will follow the expectations of RNS program, including format (submit as Word document) and length (5000 words), due 2026-08-15, to eventually be published in the online journal Ecorestoration: RNS Technical Series

Bibliography

- Malloff, J., & Shackelford, N. (2024). Feeling the Pulse: Monitoring methods and initial outcomes in oak meadow ecosystems. Restoration Futures Lab at the University of Victoria.
- McKay, S. K. (2020). Data visualization for ecological analysis and restoration. US Army Corp of Engineers:Ecosystem Management and Restoration Research Program (EMRRP) Webinar Series. https://cw-environment.erdc.dren.mil/webinars/20Feb5-EcoMod-DataViz.pdf

Workflow Graph

- See Objectives, Methods, and Deliverables
- The workflow shows how the raw data, from HAT monitoring measurements, would be transformed and included in code to produce a variety of deliverable products.
- This graph is made with code from DiagrammeR, using a colour-blind friendly colour palette.

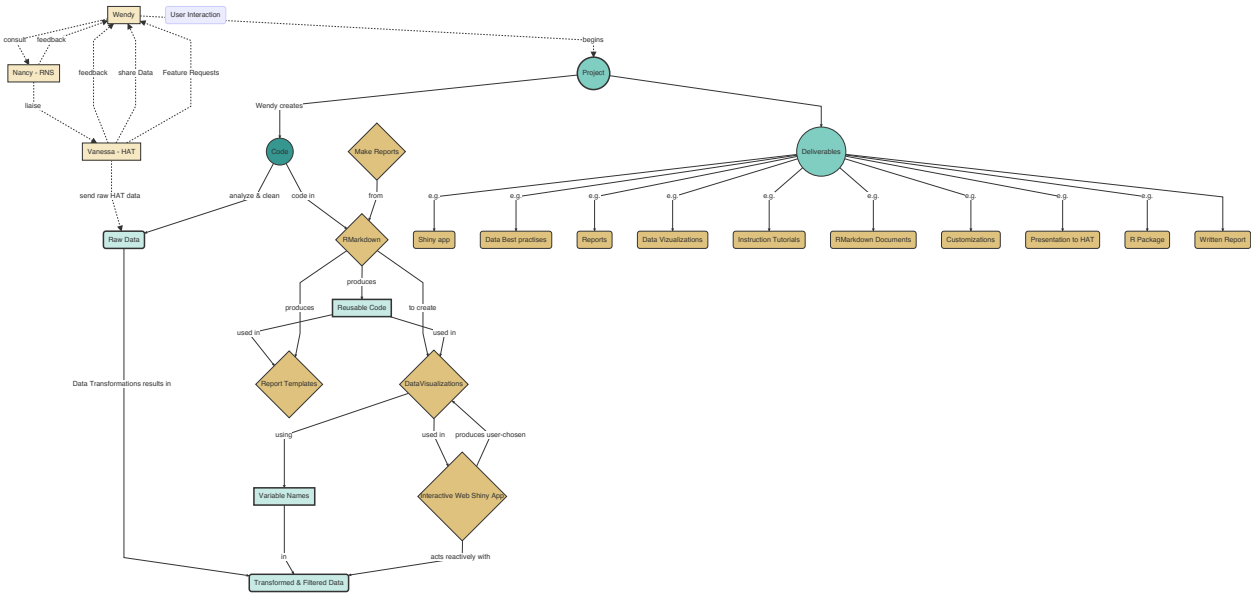

Figure 1: Project method workflow diagram for ER390 data visualization for HAT

Proposed Timeline (Look at SMART Objectives)

- red coloured items are considered critical
- grey items are completed
- Note: This chart was coded with DiagrammeR using an imported csv file of data
- See Objectives, Methods, and Deliverables

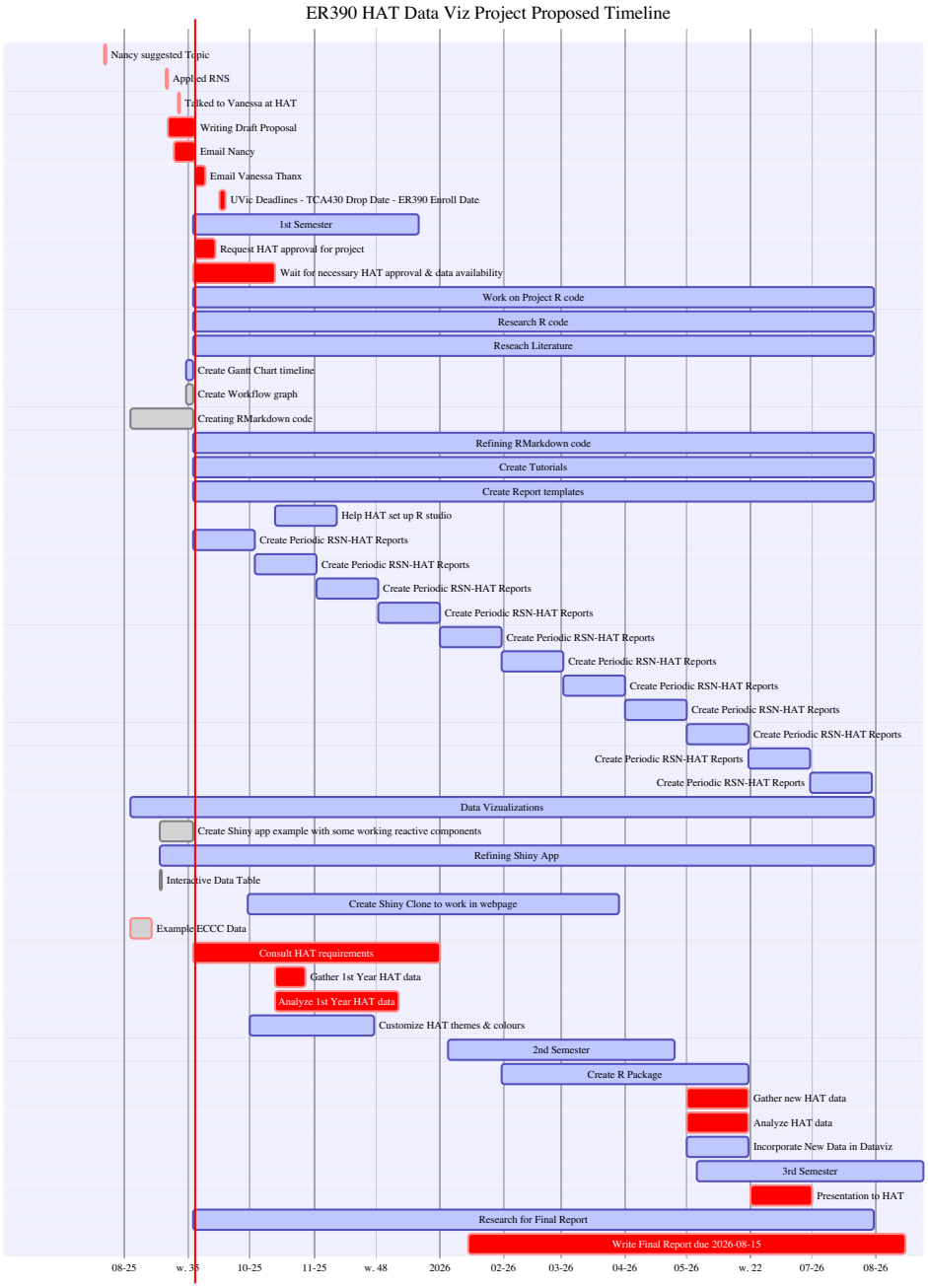

Figure 2: Gantt Proposed ER390 Timeline

Example Data Visualization Results

Data download and formatting

- These example visualizations use GOE monitoring data from report by Malloff & Shackelford (2024)
- Data (in excel or csv format) is uploaded
- Data is formatted, saved, and managed to be useful for creating visualizations
- The code has been hidden for the HTML report, highlighting only the resulting plots, though is visible in PDF document. The R code has been documented using hidden comments.

Load Site Data

Click open to see R code for `Loading Data`

```
# Load data and create 2 data frames
# Load Site Data
geo_eccc <- read.csv("data/SiteData-TEST-GOE-Monitor.csv", header = TRUE, sep = ",")
# Load Site Details
Sites_eccc <- read.csv("data/SiteDetails-TEST-GOE-Monitor.csv", header = TRUE, sep = ",")
```

Data Cleaning and Transformation

Click open to see R code for `Data Cleaning and Transformation`

```
# create new column with proportions changed to percentage as a whole number
Percentage_ns <- geo_eccc$Proportion_of_native_species * 100

# add-new column for Percentages
geo_eccc <- cbind(geo_eccc, Percentage_ns)

# arrange order of site names
Sites_eccc_1 <- Sites_eccc %>% arrange(Sites_eccc$Site)
geo_eccc_1 <- geo_eccc %>% arrange(geo_eccc$Site)

# bind 2 dataframes together
# this requires Sites rows from both data frames to be in alphabetical order
## to assign the data rows correctly
geo_eccc_sites <- cbind(Sites_eccc_1, geo_eccc_1)

# Remove duplicate columns resulting from binding 2 data frames
geo_eccc_sites_1 <- geo_eccc_sites[c(-9,-10)]

# Rename columns to remove dot-format
names(geo_eccc_sites_1)[names(geo_eccc_sites_1) == "Land.Manager"] <- "Land_Manager"
names(geo_eccc_sites_1)[names(geo_eccc_sites_1) == "Main.Restoration.Type"] <- "Main_Restoration_Type"
names(geo_eccc_sites_1)[names(geo_eccc_sites_1) == "Restoration.Intensity"] <- "Restoration_Intensity"
names(geo_eccc_sites_1)[names(geo_eccc_sites_1) == "Area.ha"] <- "Area_ha"

# write csv file with combined data
## to create new data file which is used for creating the visualizations
write.csv(geo_eccc_sites_1, "data/geo_eccc_all_site_data.csv", row.names = FALSE)
```

Data Visualization Example Results

Data Table

- using data from Malloff & Shackelford (2024)

Full Data Table

- The full table is too wide for use in pdf because current variable names are too long.

Subregion Site Land_Manager Main_Restoration_Type Restoration_Intensity Area_ha Lat Lng Proportion_of_native_species Cultural_species_richness Exotic_species Trampling Herbivory Composite_Index Year Percentage_ns

Table Head of Site Information

Subregion	Site	Land.Manager	Main.Restoration.Type	Restoration.Intensity	Area.ha	Lat	Lng
Gulf Islands	Anniversary Island	GINPR	herbivore reduction	high	4.39	48.82292	-123.1823
Gulf Islands	AVNR	Saltspring conservancy	herbivore reduction	minimal	20.54	48.80351	-123.4425
Saanich Peninsula	Bear Hill Park	CRD	invasive removal	low	3.80	48.54639	-123.4078
Gulf Islands	Brackman Island	GINPR	invasive removal	high	4.41	48.71897	-123.3864
Saanich Peninsula	Camas Hill	HAT	invasive removal	high	10.10	48.40173	-123.5975
Gulf Islands	Crows Nest	Trinity Western University	herbivore reduction	high	15.34	48.78237	-123.4612

Table Head of Site Data

Subregion	Site	Proportion_of_native_species	Cultural_species_richness	Exotic_species	Trampling	Herbivory	Composite_
Gulf Islands	Anniversary Island	0.94	2.71	0.17	5.00	1.14	
Gulf Islands	AVNR	0.55	1.30	57.19	4.70	36.50	
Saanich Peninsula	Bear Hill Park	0.62	1.86	24.46	7.57	9.50	
Gulf Islands	Brackman Island	0.77	2.80	11.17	5.40	0.00	
Saanich Peninsula	Camas Hill	0.61	2.43	27.19	7.71	4.71	
Gulf Islands	Crows Nest	0.66	1.07	31.51	1.64	26.71	

Figures 3,4,5: Example Data Tables

Map of Site Locations

Leaflet Map with Circle Markers

- The Leaflet Map is created using R code in RMarkdown.
- Data is summarized in map marker popups with summary of site data

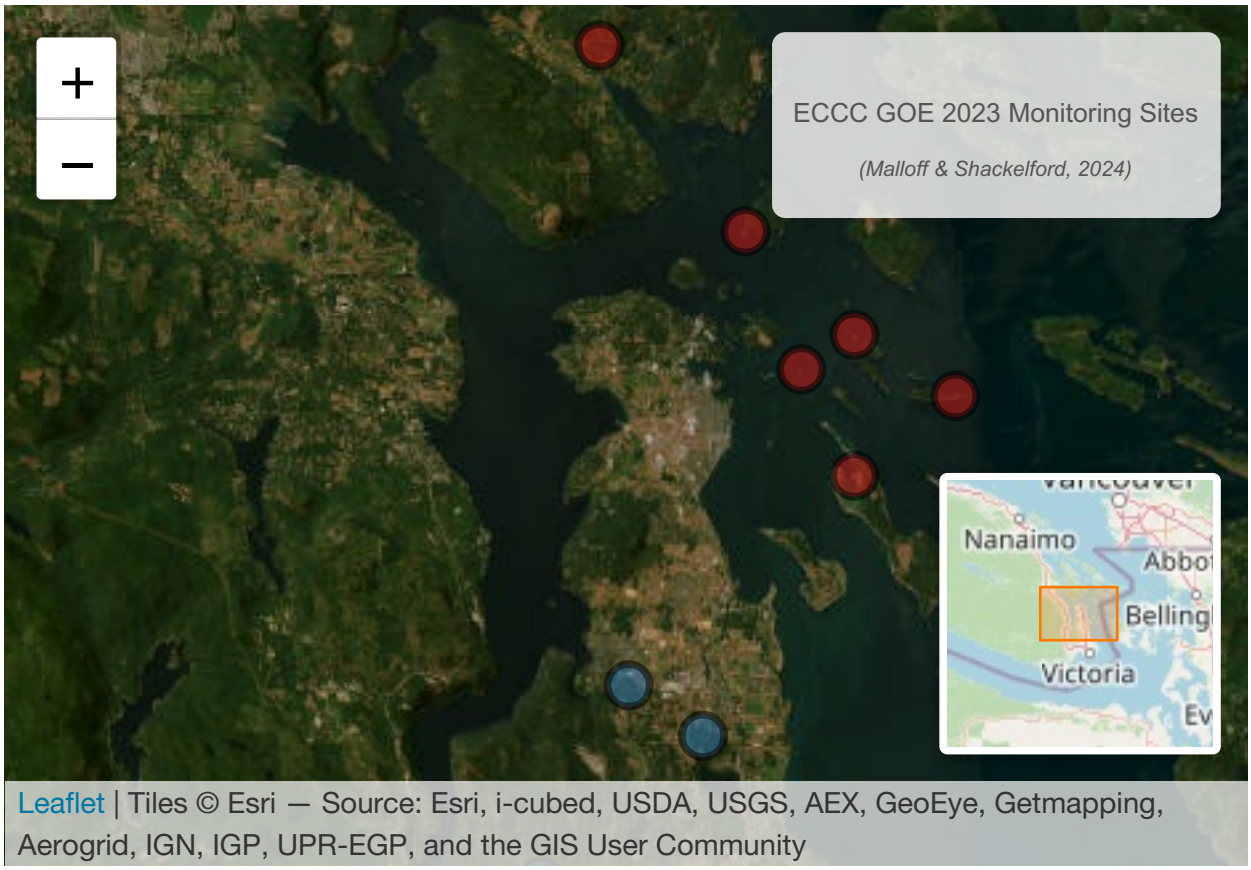

Figure 6: Interactive Leaflet Map with Example Sites Data

Subregion Plots

- These plot compare monitoring variable measurement results between two Subregions.

Violin/Bar Plot: Compare Subregions Exotic and Native Species

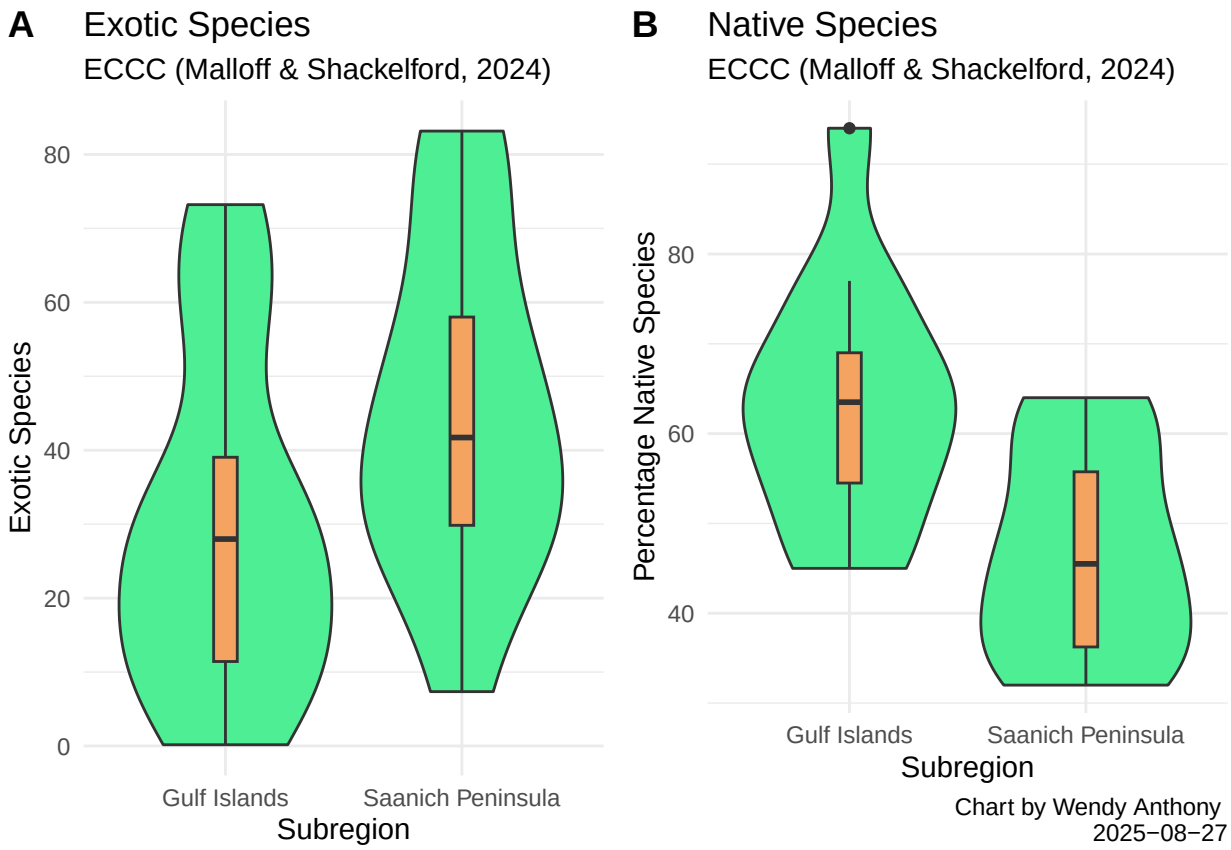

Figure 7: Violin / bar plot comparing variables of 2 Subregions

Site Plots

Point Plot: Compare Individual Sites by Restoration Intensity

- Sites are ordered by the level of restoration intensity, ranging from high to minimal
- This plotly-style chart becomes interactive by clicking data points for view a tool-tip of data results.

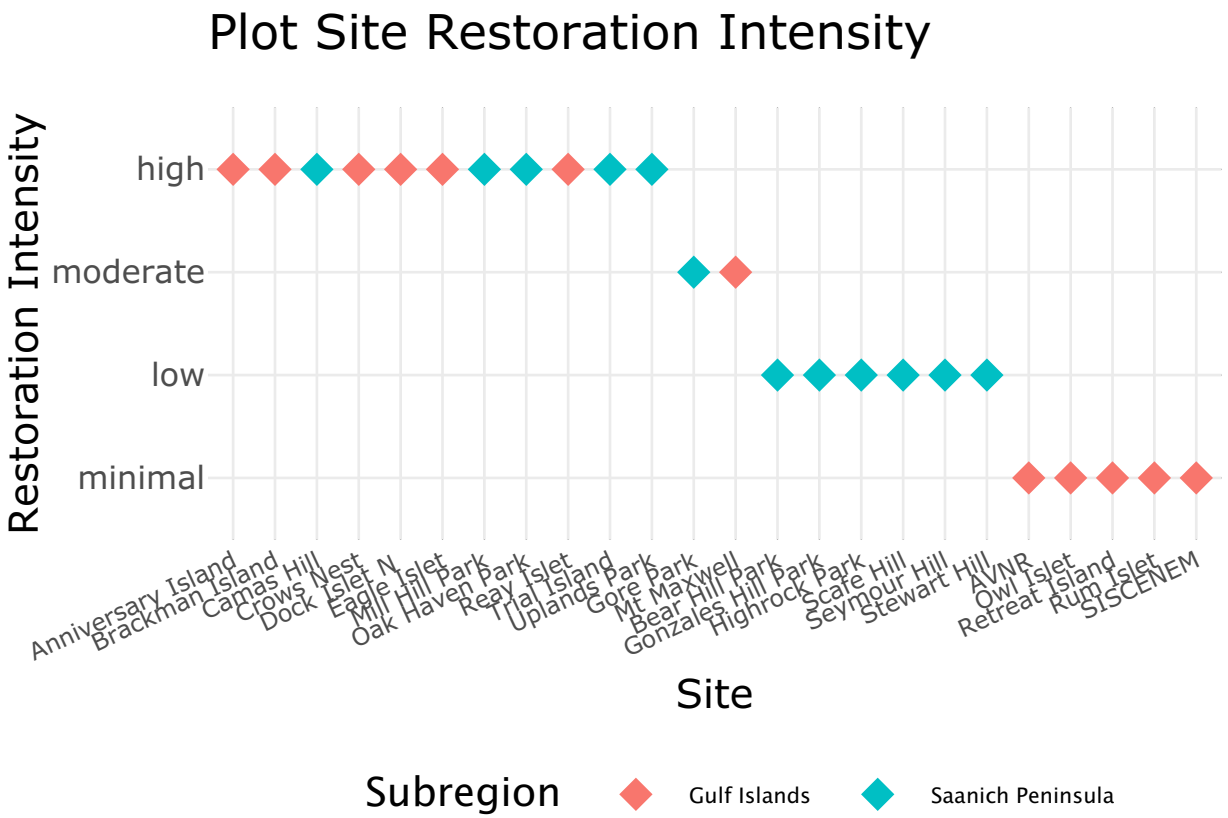

Figure 8: Point Plot of Restoration Intensity of Example Sites