NMFS R User Group Lightning Talks: Open Science Work in R

January 31, 2023

3:00PM Eastern

What is the NMFS R User Group?

- Community of R users within NOAA Fisheries
- Monthly meetings, Google Space, calendar of R related events
- Next month: Connecting to REST APIs with R. Feb 28 at 3 pm EST

Ways to Join:

Link to form on our GitHub page (https://nmfs-opensci.github.io/NMFS-R-UG/)

OR reach out to emily.markowitz@noaa.gov, eli.holmes@noaa.gov, kathryn.doering@noaa.gov, or josh.london@noaa.gov

A personal take on science and society

World view

Why 2023 is the US Year of Open Science

Here's how NASA is incentivizing open science, and how you can too.

In Nature by Chelle Gentemann

Read the article: https://doi.org/10.1038/d41586-023-00019-y

For more information about NOAA Fisheries Open Science efforts:

NMFS OPEN SCIENCE

https://nmfs-opensci.github.io/

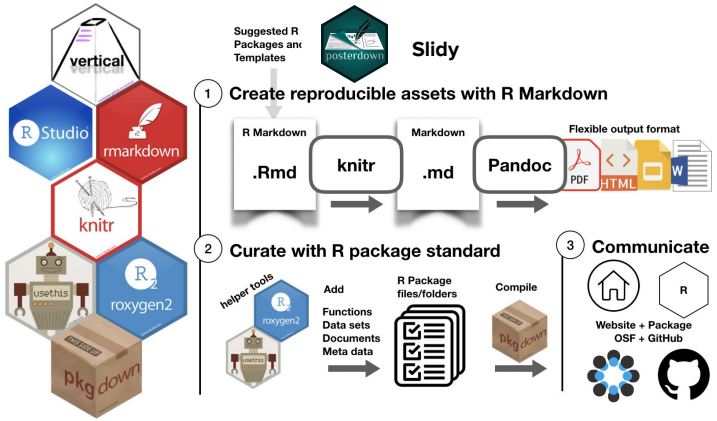
Upcoming lightning talks:

- Hem Nalini Morzaria-Luna (NWFSC) “Using R package vertical for reproducibility of scientific papers”
- Andy Beet (NEFSC) “stocksmart: an R data package”
- Andrea Havron (OST) “Fisheries Integrated Modeling System (FIMS)”
- Alana Santana and Rory Spurr (University of Washington and WCR) “Research permit visualization app”
- Greg Williams (NWFSC) “Automating California Current Ecosystem Status Reports for the PFMCouncil”
- Eli Holmes (NWFSC) “Creating NOAA reports with R and Quarto using the quarto_titlepages extension”
- Em Markowitz (AFSC) “Converting data reports to dynamic R Markdown feat. the Bering Sea bottom trawl survey data report”
- Sean Rohan (AFSC) “coldpool: Cold pool area and temperature from the eastern Bering Sea”
- Elizabeth Gugliotti (OST) “Creating a Posit Connect API and using it to run a model/get model results”
- Meg Oshima (PIFSC) “Automating SS model development workflow”
- Felipe Quezada (SWFSC) “CPS cluster and bayesian modelling for landings”
- Desiree Tommasi (SWFSC) “Bluefin tuna MSE”
- Brian Smith (NEFSC) “Knee-deep in fish guts: sharing metadata and creating interactive products with Shiny and RMarkdown”
- Catherine Foley (NEFSC) “Operational Tools for the NEFSC Northeast Bottom Trawl Survey”

Using R package *vertical* for reproducibility of scientific papers

Hem Nalini Morzaria-Luna, Github: hmorzaria.
HemNalini.MorzariaLuna@noaa.gov

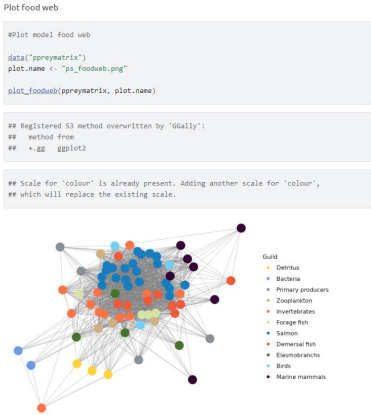
- R-based structured workflow for creating and sharing research
- FAIR (Findable, Accessible, Interoperable, Reusable) guidelines
- Document data, analysis and results in one place

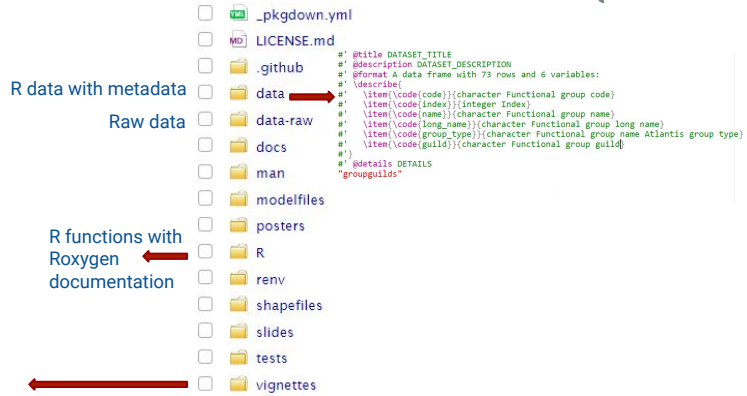


Vuorre, M. and Crump, M.J., 2021. Sharing and organizing research products as R packages. *Behavior research methods*, *53*, pp.792-802.

https://crumplab.com/vertical/articles/vertical.html

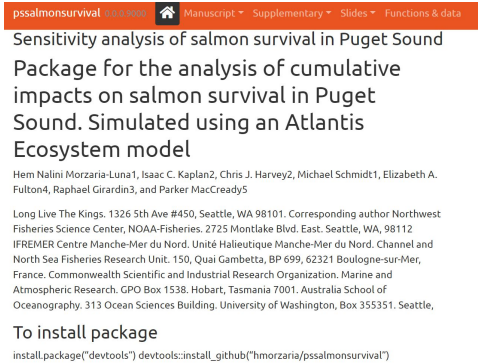
Vignette (R notebook in Rmarkdown) keeps track of data wrangling, analysis, and figures





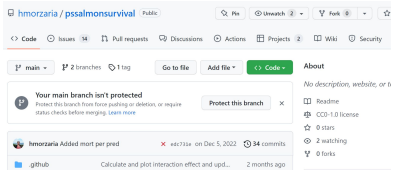
Website to share vignette, supplement, slides and manuscript

https://hmorzaria.github.io/pssalmonsurvival/



GitHub action

Data and code archiving - DOI



zenodo

stocksmart: An R data package

https://github.com/NOAA-EDAB/stocksmart/

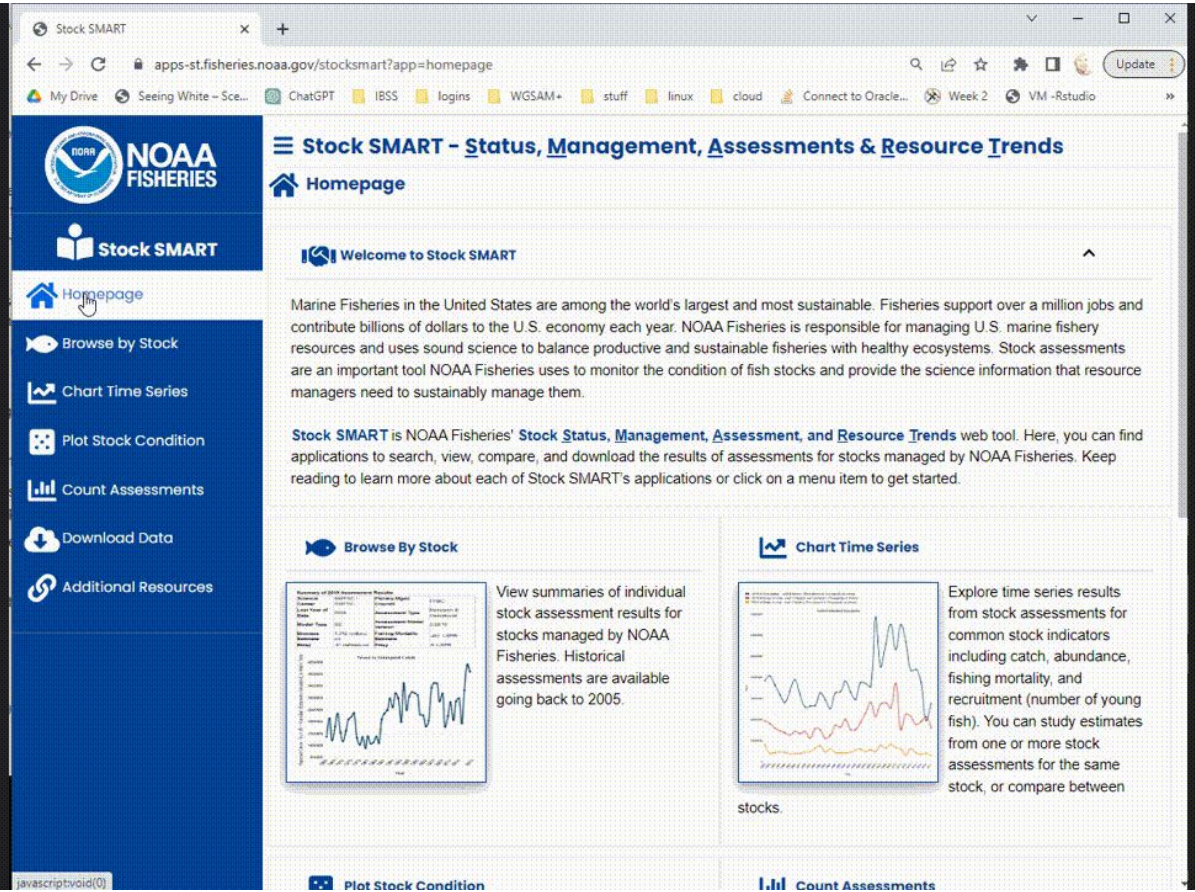
Andy Beet (EDAB, NEFSC)

With help from the stock SMART team:

Kristan Blackhart

Wei Qiu

Jeffrey Vieser



FIMS R Package

Linking in TMB

Andrea Havron, ECS Federal, OST

main | Go to file | Add file | Code

ChristineStawitz-NOAA Merge pull req... ✓ 3 weeks ago 397

.devcontainer	Population loop module (#237)	3 weeks ago
.github	small gha changes: branches	last month
R	Population loop module (#237)	3 weeks ago
data-raw	style and docs: run devtools::docu...	6 months ago
data	107 feature implement empirical w...	7 months ago
inst	style: run clang format	3 weeks ago
man	style and docs: run devtools::docu...	3 weeks ago
src	Population loop module (#237)	3 weeks ago
tests	style: run clang format	3 weeks ago
.Rbuildignore	fix, #209: changes recommended b...	4 months ago
.gitignore	Population loop module (#237)	3 weeks ago
CMakeLists.txt	add more c++ and R tests (#248)	3 months ago
CONTRIBUTIN...	make shorter and link to collaborat...	2 months ago
DESCRIPTION	style and docs: run devtools::docu...	last month
FIMS.Rproj	make into an R package	last year
LICENSE	getting rid of .md to make CRAN h...	10 months ago
NAMESPACE	fix, #209: changes recommended b...	4 months ago
README.md	90 cleanup readme add badges an...	8 months ago
codecov.yml	90 cleanup readme add badges an...	8 months ago

README.md

call-r-cmd-check passing | codecov 86% | lifecycle experimental

[FIMS landing page](), [FIMS development repo]()
[TMB wiki]() on distributing code

src/

FIMS.cpp ← `#define TMB_LIB_INIT R_init_FIMS`

Makevars

DESCRIPTION

```
Imports: TMB
LinkingTo:
  TMB
  RcppEigen
```

NAMESPACE

```
useDynLib(FIMS, .registration = TRUE)
```

R/FIMS-package.R

```
## usethis namespace: start
#' @useDynLib FIMS, .registration = TRUE
#' @importFrom Rcpp sourceCpp
#' @importFrom utils head
#' @import stats
#' @import methods
#' @importFrom ggplot2 .data
#' @importFrom usethis use_template ui_stop
## usethis namespace: end
NULL
```

`roxygen2::roxygenize()`

inst/include

common	3 weeks ago
distributions	3 weeks ago
interface	3 weeks ago
population_dynamics	3 weeks ago

TMB model compiles during package install

```
> remotes::install_github("NOAA-FIMS/FIMS")
```

« Documents › R › R-4.1.3 › library › FIMS › libs › x64

Name	Date modified
FIMS.dll	1/26/2023 10:19 AM

Develop and run FIMS virtually

.devcontainer

Go to file | Add file | Code

Local | Codespaces

Codespaces
Your workspaces in the cloud

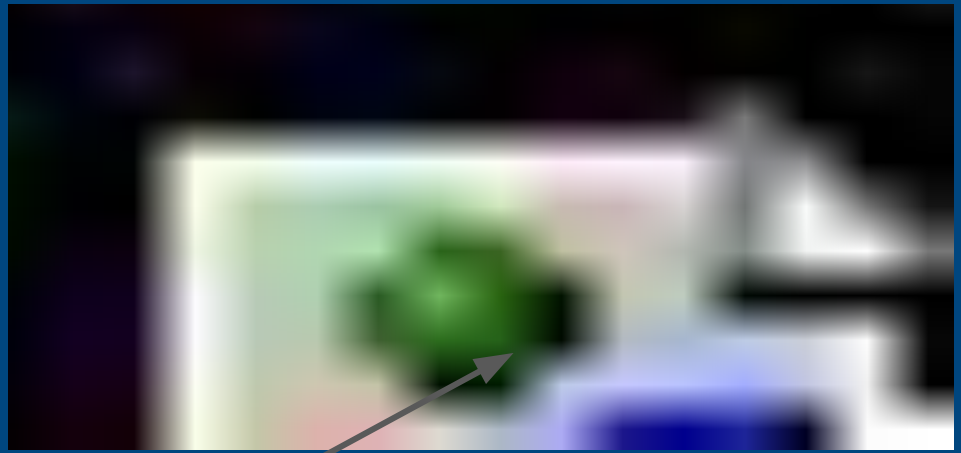
On current branch

laughing tribble ● Active
main No changes

Visualizing ESA-Listed Fish Research In the West Coast Region

Alana Santana (Github: asantan8) and Rory Spurr (Github: rory-spurr)

UW advisor: Dr. Anne Beaudreau | NMFS lead: Diana Dishman



Leveraging the power of Shiny and Leaflet together to create interactive maps

Using plotly and R Shiny to develop dynamic and interactive time series plots

Objectives:

- Support the decision-making process for scientific research permits in NOAA West Coast Region.
- Provide more transparency to researchers as well as state and tribal government employees about the permitting process.
- Educate the public about the role of research to inform the management of ESA-listed species.

CCIEA - Ecosystem Status Report Automation

EBFM OBJECTIVE -

Compile and integrate data from >90 scientists in 6-8 weeks
Synthesize ecosystem status and trends report for PFMC

2° Objective -

Openscapes - Community of Practice
Open Data Science

Greg Williams (team: N. Tolimieri, L. DeWitt, C. Harvey, A. Leising)

Data

CCIEA Data Uploader

INTEGRATED ECOSYSTEM ASSESSMENT

Example.Data

	A	C
1	year	SEup
2	2004	34.26655756
3	2005	33.98078964
4	2006	35.77528111
5	2006	35.39128618	37.45690339
6	2007	35.27921512	36.84814213
7	2008	35.16213729	37.17317949
8	2009	36.97578032	38.78720003
9	2010	36.76255016	38.5401725
10	2011	35.7290419	37.83469224

ERDDAP
Easier access to scientific data

ERDDAP

ERDDAP is a data server that gives you a simple, consistent way to download subsets of scientific datasets in common file formats and make graphs and maps. This particular ERDDAP installation has oceanographic data (for example, data from satellites and buoys).

Easier Access to Scientific Data

Tables & Text Submissions

CCIEA-esr / Projects / ESR Data Deadline Management 2022-23

ESR Data Deadline Management 2022-23

	Title
1	Overview Summary Graphic
2	Climate Change Appx: Leising
3	Climate / Fire/ Drought section
4	CA salmon tables - Correigh
5	Data uploader trials: Tom - Seabirds

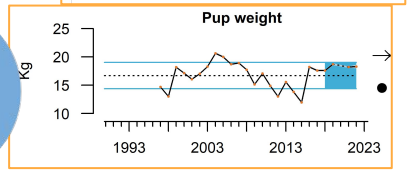
CCIEA-esr

Overview | Repositories 3

Find a repository...

Ecosystem-status-report-FY23 Private
Repo for preparing and archiving ESR FY2023 2022-2023
TeX 0 1 0 0 Updated 5 hours ago

ESR-Figures-Code Private
Repo for producing figures for the ESR
R 0 1 0 0 Updated on Nov 23, 2022



Rmd Code

Figure Code

Final Report, Web Products

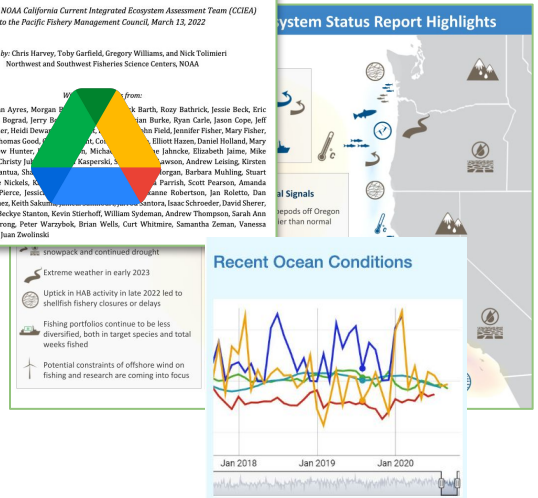
Agenda Item H.2.a
CCIEA Team Report 1 (Electronic Only)
March 2022

2021-2022 CALIFORNIA CURRENT ECOSYSTEM STATUS REPORT

A report of the NOAA California Current Integrated Ecosystem Assessment Team (CCIEA) to the Pacific Fishery Management Council, March 13, 2022

Edited by: Chris Harvey, Toby Garfield, Gregory Williams, and Nick Tolimieri
Northwest and Southwest Fisheries Science Centers, NOAA

Written by and contributions from:
Kelly Andrews, Dan Ayres, Morgan B... Barth, Roxy Bathrick, Jessie Beck, Eric Bjorkstedt, Steven Bograd, Jerry B... Burke, Ryan Carle, Jason Cope, Jeff Cowen, David Demer, Heidi Dewa... Field, Jennifer Fisher, Mary Fisher, Zachary Forster, Thomas Good... Elliott Hazen, Daniel Holland, Mary Hunsicker, Matthew Hunter, ... Jahncke, Elizabeth Jaime, Mike Johns, Tim Jones, Christy Ju... Kasperski, ... Lawson, Andrew Leising, Kirsten Lindquist, Nate Mantua, Sh... Morgan, Barbara Muhling, Stuart Munsch, Catherine Nickels, ... Parrish, Scott Pearson, Amanda Phillips, Stephen Pierce, Jessi... Anne Robertson, Jan Roletto, Dan Rudnick, Lauren Saez, Keith Sakuma, ... Santora, Isaac Schroeder, David Sherer, Kayleigh Somers, Beckye Stanton, Kevin Sterhoff, William Sydeman, Andrew Thompson, Sarah Ann Thompson, Duy Trong, Peter Warzybok, Brian Wells, Curt Whitmire, Samantha Zeman, Vanessa Zubkousky-White, Juan Zwolinski

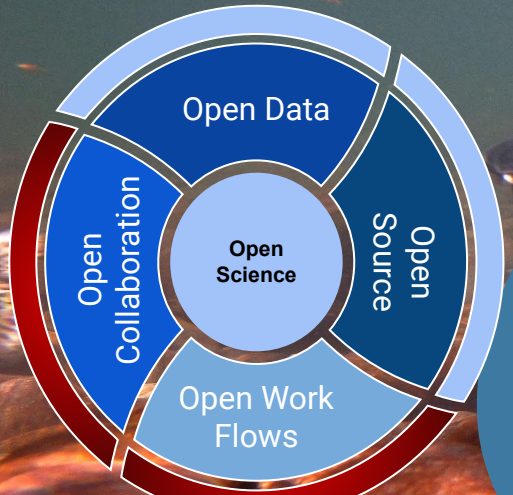


ERDDAP: https://oceanview.pfeg.noaa.gov/erddap/index.html
CCIEA Website: https://www.integratedecosystemassessment.noaa.gov/regions/california-current

A Quarto extension for reproducible government reports: *quarto_titlepages*

Eli Holmes, NWFSC, eli.holmes@noaa.gov

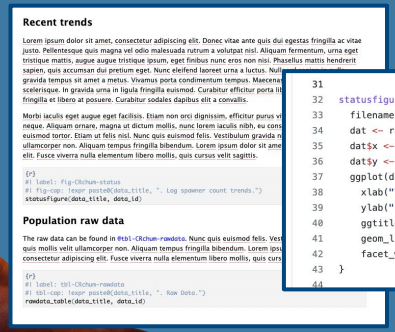
Github: eeholmes



quarto

CONTENT

Text, data, code for tables figures



Recent trends

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec vitae ante quis dui egestas fringilla ac vitae justo. Pellentesque quis magna vel odio malesuada rutrum a volutpat nisl. Aliquam fermentum, urna eget tristique mattis, augue augue tristique ipsum, eget finibus nunc eros non nisi. Phasellus mattis hendrerit sapien, quis accumsan dui pretium eget. Nunc eleifend laoreet urna a luctus. Nul
gravida tempus sit amet a metus. Vivamus porta condimentum tempus. Maecena
scelerisque. In gravida urna in ligula fringilla euismod. Curabitur efficitur porta lit
fringilla et libero at posuere. Curabitur sodales dapibus elit a convallis.

Morbi iaculis eget augue eget facilisis. Etiam non orci dignissim, efficitur purus vi
neque. Aliquam ornare, magna ut dictum mollis, nunc lorem iaculis nibh, eu cons
euismod tortor. Etiam ut felis nisl. Nunc quis euismod felis. Vestibulum gravida n
ullamcorper non. Aliquam tempus fringilla bibendum. Lorem ipsum dolor sit ame
elit. Fusce viverra nulla elementum libero mollis, quis cursus velit sagittis.

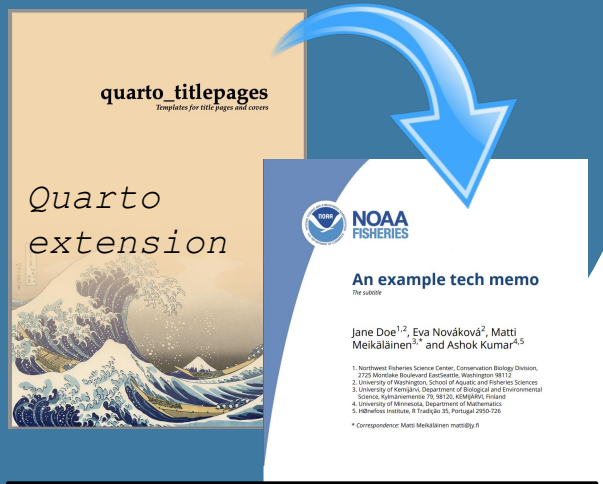
```
{r}
#| label: fig-CRchum-status
#| fig-cap: !expr paste0(data_title, ". Log spawner count trends.")
statusfigure(data_title, data_id)
```

Population raw data

The raw data can be found in #tbl-CRchum-rawdata. Nunc quis euismod felis. Ves
quis mollis velit ullamcorper non. Aliquam tempus fringilla bibendum. Lorem ipsu
consectetur adipiscing elit. Fusce viverra nulla elementum libero mollis, quis curs

```
{r}
#| label: tbl-CRchum-rawdata
#| tbl-cap: !expr paste0(data_title, ". Raw Data.")
rawdata_table(data_title, data_id)
```

```
statusfigure <- function(title, id, x="YEAR", y
  filename <- here::here("data", paste0(id, ".c
  dat <- read.csv(filename, stringsAsFactors =
  dat$x <- dat[[x]]
  dat$y <- log(dat[[y]])
  ggplot(dat, aes(x=x, y=y)) +
    xlab("Year") +
    ylab("log(Spawners)") +
    ggtitle(title) +
    geom_line() +
    facet_wrap(~COMMON_POPULATION_NAME)
}
```



TYPESETTING

PDFs with the title pages, cover pages, copyright, etc

NMFS OPEN SCIENCE

Find this and more templates at *nmfs-opensci.github.io*

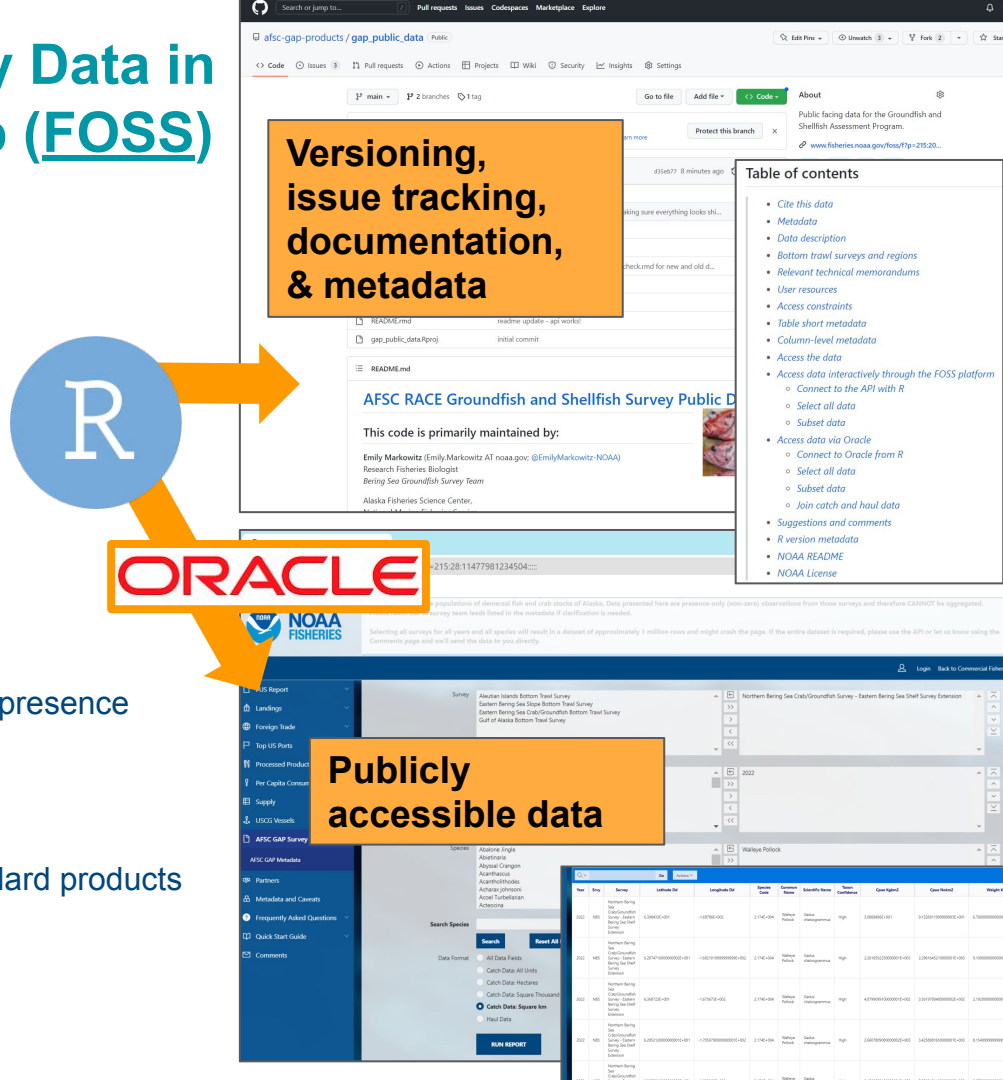
NOAA FISHERIES

AFSC Groundfish Survey Data in Fisheries One Stop Shop (FOSS)

Em Markowitz

Objective: Share standard station-level catch, environmental, and catch-per-unit data from our surveys with the public (more info).

- **Modernized data accessibility**
  User-friendly interactive table & API connections
- **Transparency-forward documentation**
  Descriptive user metadata (GitHub readme)
- **Streamlined data distribution**
  Scientist make tables; FOSS manages public-facing presence
- **Reproducible workflows**
  Functionalized R scripts shared on GitHub
- **Minimizing data requests**
  Eliminates time scientists spend responding on standard products
- **Welcoming collaboration**
  Inviting community feedback and GitHub issues



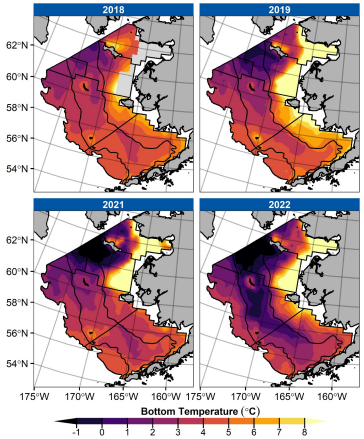
coldpool: Temperature data products for the eastern Bering Sea

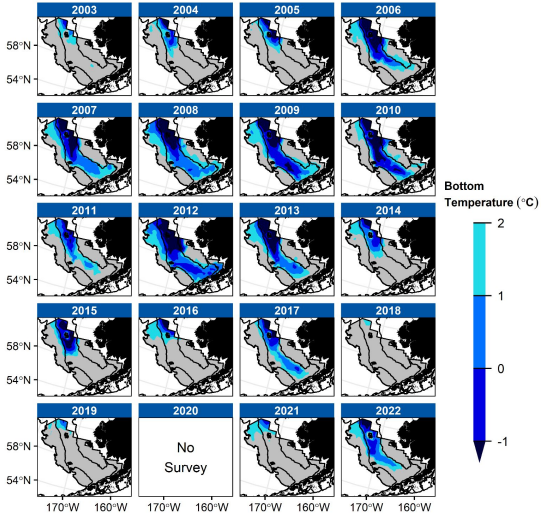
What does it do?

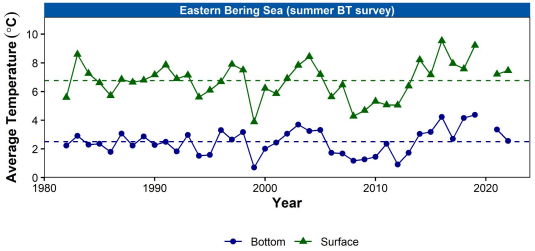
- Calculates annual bottom trawl survey temperature data products.
- Provides data products as built-in data sets (w/ documentation).
- Supplies free-to-use plots.

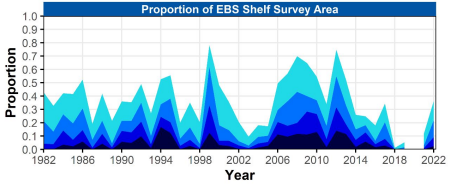
Uses:

- Covariates for stock assessment
- Risk Tables in stock assessments
- Ecosystem Status Reports
- Primary research
- Presentations and reports









https://github.com/afsc-gap-products/coldpool

Sean Rohan (@Sean-Rohan-NOAA)

Using the Posit Connect API to run Stock Synthesis

Elizabeth Gugliotti and Kathryn Doering

Setup locally

[See github repository]

Plumber creates the API

Deployed on Posit Connect

Create API key on Posit Connect

Interact with API through R

Source on Save | Run

```r
library(httr)
library(jsonlite)

setwd("C:/Users/elizabeth.gugliotti/Desktop/R_fun")
source("C:/Users/elizabeth.gugliotti/Desktop/R_fun/connect_api_key.R")

connect_path <- "https://dev-connect.fisheries.noaa.gov/content/4145fffa-489f-4f9b-9b38-66d87c9d4b01/"

run_ss <- GET("https://dev-connect.fisheries.noaa.gov/content/4145fffa-489f-4f9b-9b38-66d87c9d4b01/ss",
  add_headers(Authorization = paste("Key", ConnectAPIKey)))

view_result <- GET("https://dev-connect.fisheries.noaa.gov/content/4145fffa-489f-4f9b-9b38-66d87c9d4b01/results",
  add_headers(Authorization = paste("Key", ConnectAPIKey)))

data = fromJSON(rawToChar(view_result$content))
```

15:48 (Top Level)

Console | Terminal | Background Jobs

R 4.2.2 · C:/Users/elizabeth.gugliotti/Desktop/R_fun/

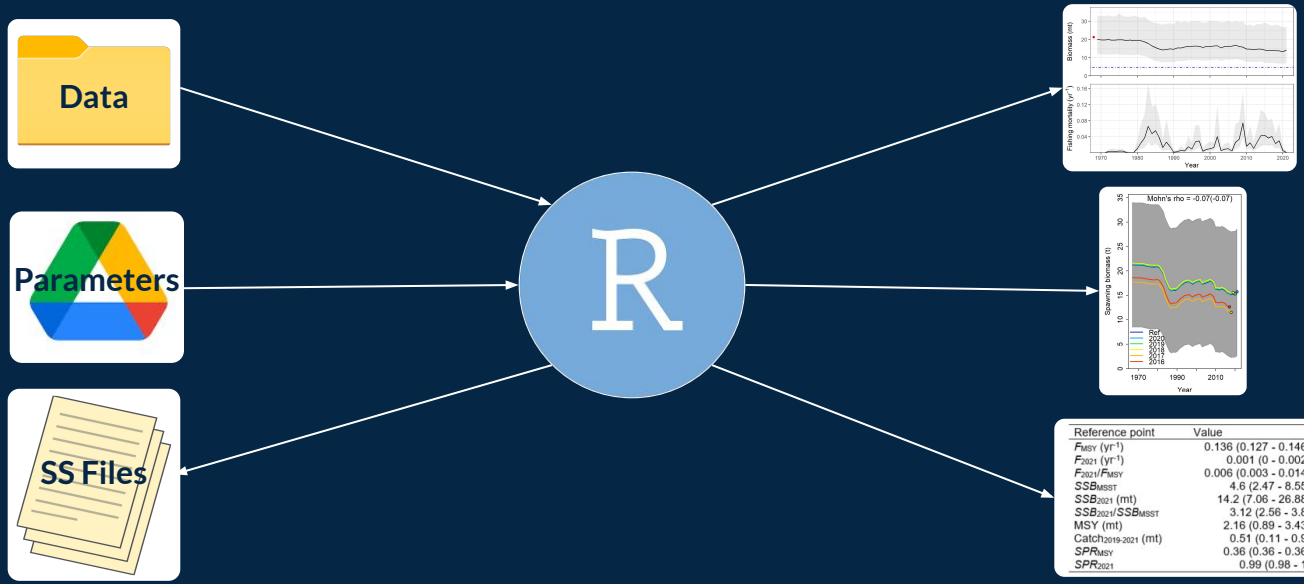
```
> data
 [1] "#V3.30.20.00;_safe;_compile_date:_Sep 30 2022;_Stock_Synthesis_by_Richard_Methot_(NOAA)_using_ADMB_13.0"

 [2] "#C generic forecast file"

 [3] "# for all year entries except rebuilder; enter either: actual year, -999 for styr, 0 for endyr, neg number for rel
```

Automating SS Model Development Workflow for Many Models

Meg Oshima and Marc Nadon

problem:

solution:

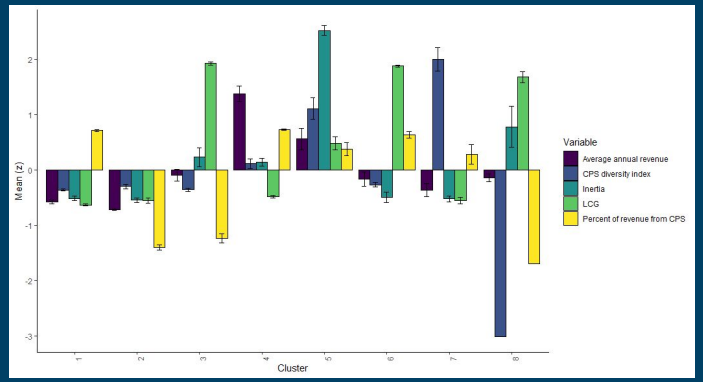


Multilevel Bayesian modeling for CPS landings using cluster analysis

Felipe Quezada

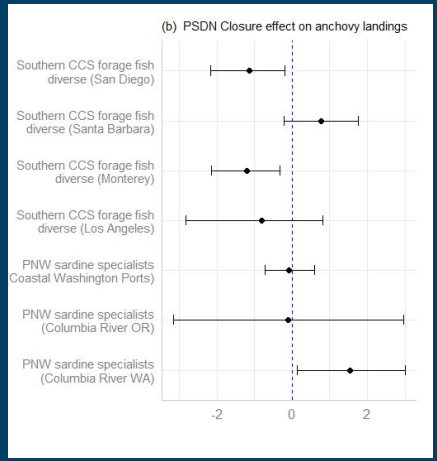
- Allocate vessel to a cluster using R package *cluster* (function PAM), based on different vessel characteristics (i.e., inputs)
- Optimal number of cluster have to be defined by the user (e.g., based on the average silhouette method)

```
#-----------------------------------------------------------
### PAM Clustering
Clusters<-pam(Distance_matrix, n.clust, keep.diss = TRUE)
```



- Multilevel (at cluster/ports level) bayesian model for CPS landings using the R package *brms* based in Stan.

```
price_model <- bf(MSQD_Price_z ~ 1 + Price.Fishmeal.AFI_z + (1 | port_ID))
landing_model <- bf(log(MSQD_Landings) ~
                    1 + MSQD_SPAWN_SDM_90 + MSQD_Price_z + ...
                    (1 + MSQD_SPAWN_SDM_90 + MSQD_Price_z + ... | port_cluster_ID))

fit_qMSQD <-
   brm(data = dataset_msqd_landing,
       family = gaussian,
       price_model + landing_model + set_rescor(TRUE),
       prior = prior_lognormal,
       iter = 2000, warmup = 1000, chains = 4, cores = 4,
       control = list(max_treedepth = 15, adapt_delta = 0.99),
       file = "Estimations/fit_qMSQD_wages_prior")
```



Pacific Bluefin Tuna Management Strategy Evaluation – Desiree Tommasi

International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean

NOAA — National Oceanic and Atmospheric Administration, U.S. Department of Commerce

University of California Santa Cruz

Simulated "True" Population and Fisheries Dynamics

Stock Synthesis

Data Generation

R

Implementation

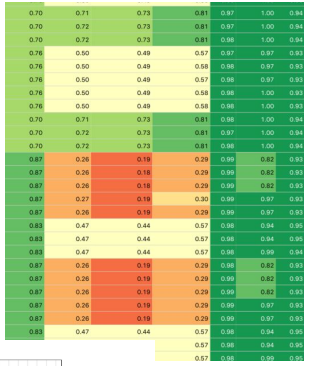
Management Procedure

Stock Synthesis

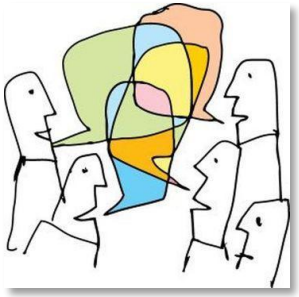
Assessment model

Harvest Control Rule

Allocation

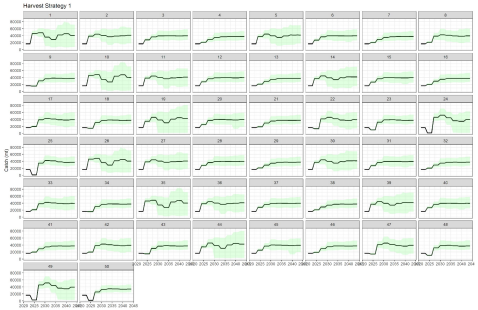


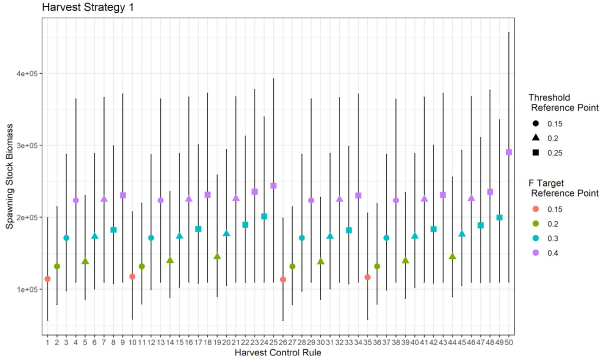
0.70	0.71	0.73	0.81	0.97	1.00	0.94
0.70	0.72	0.73	0.81	0.97	1.00	0.94
0.70	0.72	0.73	0.81	0.98	1.00	0.94
0.76	0.50	0.49	0.57	0.97	0.97	0.93
0.76	0.50	0.49	0.58	0.98	0.97	0.93
0.76	0.50	0.49	0.57	0.98	0.97	0.93
0.76	0.50	0.49	0.58	0.98	1.00	0.93
0.76	0.50	0.49	0.58	0.98	1.00	0.93
0.70	0.71	0.73	0.81	0.98	1.00	0.94
0.70	0.72	0.73	0.81	0.97	1.00	0.94
0.70	0.72	0.73	0.81	0.98	1.00	0.94
0.87	0.26	0.19	0.29	0.99	0.82	0.93
0.87	0.26	0.18	0.29	0.99	0.82	0.93
0.87	0.26	0.18	0.29	0.99	0.82	0.93
0.87	0.27	0.19	0.30	0.99	0.97	0.93
0.87	0.26	0.19	0.29	0.99	0.97	0.93
0.83	0.47	0.44	0.57	0.98	0.94	0.95
0.83	0.47	0.44	0.57	0.98	0.94	0.95
0.83	0.47	0.44	0.57	0.98	0.99	0.94
0.87	0.26	0.19	0.29	0.98	0.82	0.93
0.87	0.26	0.19	0.29	0.99	0.82	0.93
0.87	0.26	0.19	0.29	0.99	0.82	0.93
0.87	0.26	0.19	0.29	0.99	0.97	0.93
0.87	0.26	0.19	0.29	0.99	0.97	0.93
0.83	0.47	0.44	0.57	0.98	0.94	0.95
			0.57	0.98	0.94	0.95
			0.57	0.98	0.99	0.95



Management Objectives

Performance Metrics





Repo: https://github.com/detommas/PBF_MSE

B. Smith, NEFSC

Fish guts, metadata, and interactive products with Shiny and R Markdown

NOAA FISHERIES

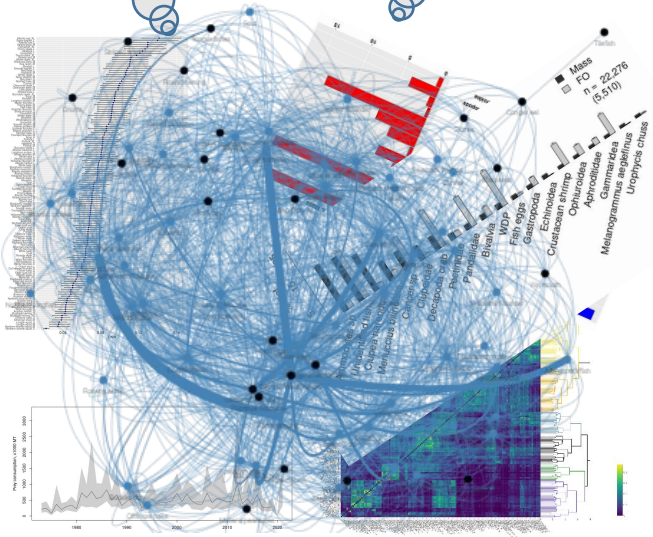
Data, Data, DATA!!

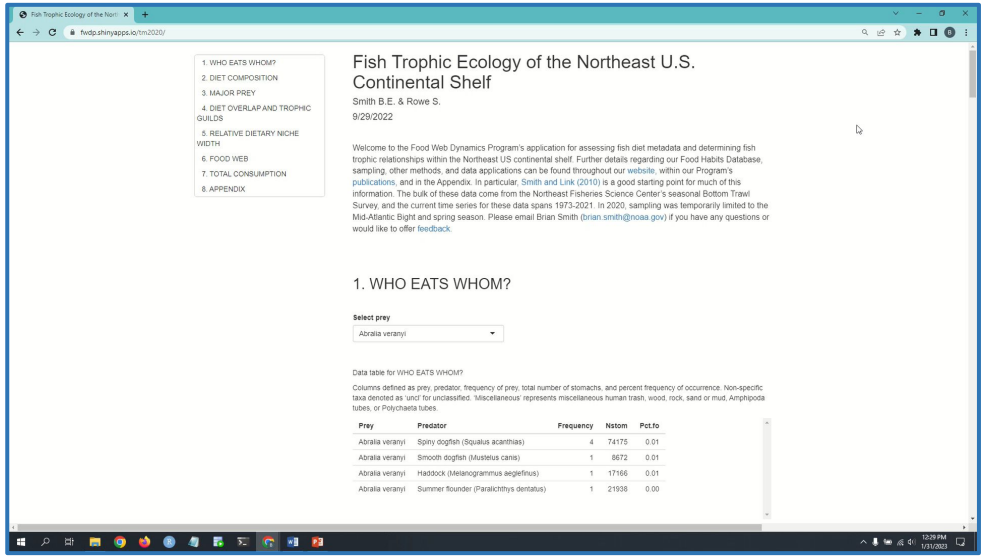
No guts, no glory!

- 187 predators
- 1,373 prey taxa
- 50 years
- ~1 M records.

..but where to start?

1 cup of R magic





- Provides an introduction, metadata, and statistics
- Reduces effort and improves data sharing
- Living document (always current)
- Maximizes data visualization
- Globally accessible

https://fwdp.shinyapps.io/tm202

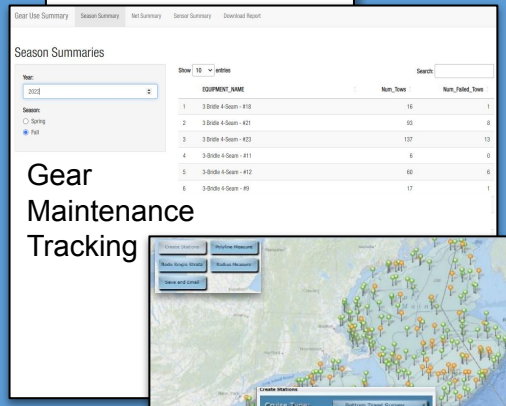
Operational Tools for the NEFSC Northeast Bottom Trawl Survey

Catherine Foley, NEFSC

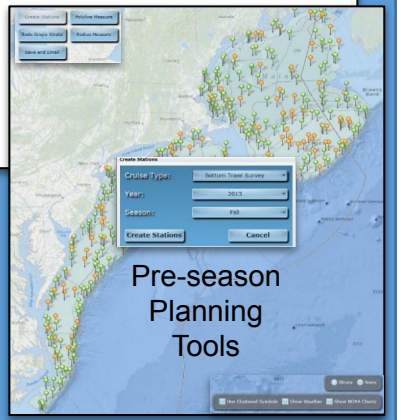
Creating a transparent, flexible framework for survey station creation

Ongoing Projects

Survey Report Automation



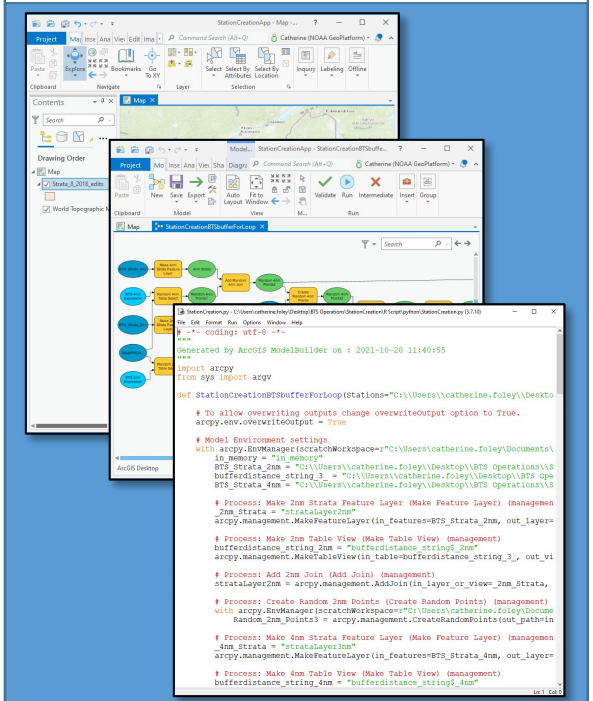
Gear Maintenance Tracking



Pre-season Planning Tools

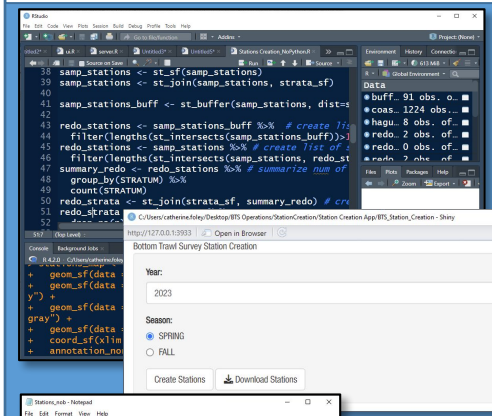
Github: cfoley2

Framework (the old)



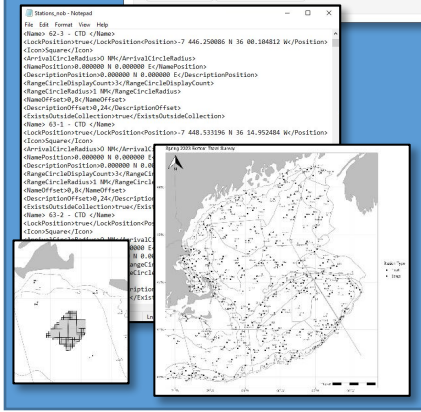
Microsoft Silverlight front end
ArcGIS ModelBuilder → python

Product (the new)



R wrapper for python script → standalone R script

Shiny UI



Output Files:

- FSCS-ready stations (csv)
- Navigation files (rosepoint, nav, nob, olex)
- Chief scientist charts
- Reports (State Dept., SBNMS, wind development areas)

Thanks to all our presenters!
Any questions?