

Source: Ichthyoplankton Information System, Alaska Fisheries Science Center

ENVIRONMENTAL RELATIONSHIPS OF FISHES

Laura Vary

Analytical Workflows

December 2, 2021

BRIEF REMINDER OF RESEARCH OBJECTIVES

How flexible is
spawning geography
and phenology?

Do water mass
characteristics better
explain larval
biogeography?

INITIAL AND MID-TERM GOALS

- Create a repository with all analytical workflow processes
- Create a separate folder with data cleaning practices
- Create a style guide with a workflow map and figure conventions
- Add comments and introductions to all code files
- Create a streamlined figure creation workflow

Decide Subsetting Process of Data and Prepare for Models

Run GAMs for Eggs and Larvae

Compare AIC Values

Compare Reduced MSE Values

Generate Predictions and Clean Figures

Ground-truth with Biological Knowledge

Seems good!










Generate Final Figures and Model Output Tables

Trim to Better Match Biological Knowledge

Modify Wiggleness of Splines

Something is weird

ACCOMPLISHED GOALS: CLEAN REPOSITORY

	varyl97 Improved figure outputs	f33ad0f 3 days ago	 50 commits
	Environmental Data	Improving code, rerunning GAMs	last month
	Figures	Added in a figure	last month
	GAM Models	Completed analyses for all species. Updating code files and working o...	20 days ago
	Ichthyo Data	Completed analyses for all species. Updating code files and working o...	20 days ago
	Organized Thesis Analyses (Code)	Improved figure outputs	3 days ago
	.gitignore	Initial commit	2 months ago
	README.md	Update README.md	last month

```
##Bering Sea Yellowfin Sole:
#the following code loads and cleans the yellowfin sole egg and larval data from ecoFOCI cruises.
#at the end of this script are finalized datasets to load in the future.
###CTD Loading: #####
allctd<-read.csv(file="All_CTD_Data_8302021.csv")
names(allctd)
allctd<-allctd[c('Latitude','Longitude','Date','Time','Pressure','Depth',
                'Temperature','Conductivity','Salinity','Sigma.T',
                'Flag','Cruise','Station','Haul','Grid','FOCI_HAUL_ID',
                'FOCI_file'')]

allctd$Year<-NA
allctd$Year<-str_sub(allctd$Date,start=-4) #to get easy year reference

allctd<-allctd[allctd$Temperature<14,] #these 2 lines remove anomalous temp/sal values
allctd<-allctd[allctd$Salinity>29&allctd$Salinity<36,] #this loads a compilation of all CTD data from ecoFOCI trawls
#ultimately, these data will be matched with larval data for larval biogeography GAMs.

###loading data, subsetting and cleaning properly
yfeggraw<-read.csv(file='YFSole_Egg_Catch.csv',header=TRUE,check.names=TRUE)
yflarvraw<-read.csv(file='YFSole_Larvae_Catch.csv',header=TRUE,check.names=TRUE)
```

```
####Generalized Additive Analyses: Yellowfin Sole
#the following code creates generalized additive models for eggs and larvae of yellowfin sole.
#these analyses form the basis of my MS thesis.
#egg data uses an averaged sea surface temperature for the month of March in the Southeastern Bering Sea
#May index was chosen because it is two months before the peak of yellowfin sole CPUE, and thus May
#conditions are likely more relevant to spawning behavior than temperatures in later months.
#load egg and larval data:

yfsub<-read.csv(file='./Ichthyo Data/Cleaned_Cut_YfEggs.csv',header=TRUE,check.names=TRUE)

yflarv.ctd<-read.csv(file='./Ichthyo Data/Cleaned_Cut_YfLarv_wCTD.csv',header=TRUE,check.names=TRUE)

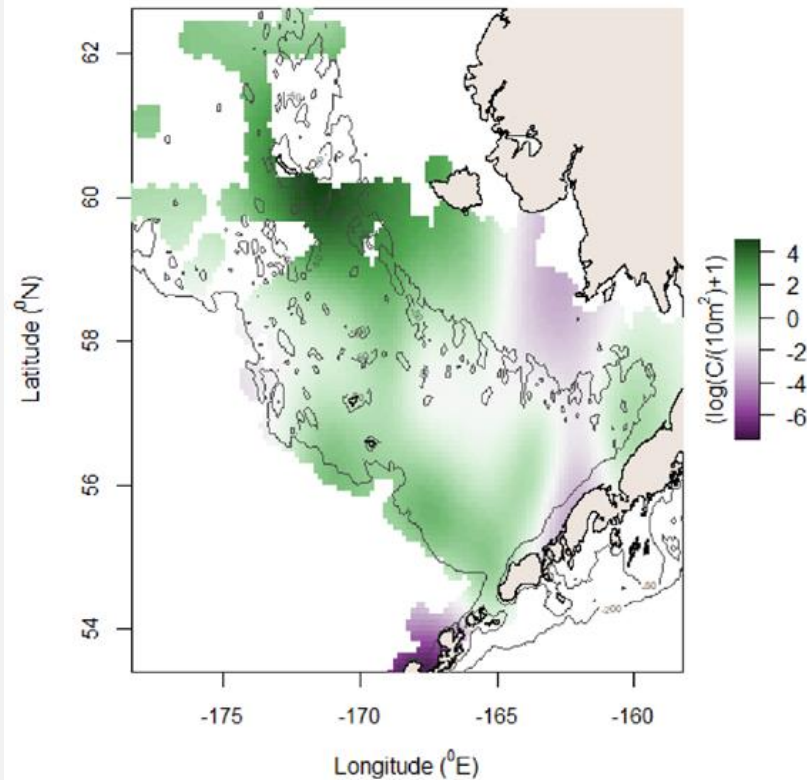
###EGGS: Spawning Behavior
##Load in local and regional temperature index for May (2 mos before peak in egg CPUE in July)
##eggs<-read.csv(file='./Environmental Data/May SST_Bering_Tideg_NCEP_R5.csv',header=TRUE,check.names=TRUE)
```

ACCOMPLISHED GOAL:
COMMENTED CODE & CODE FOR CLEANING

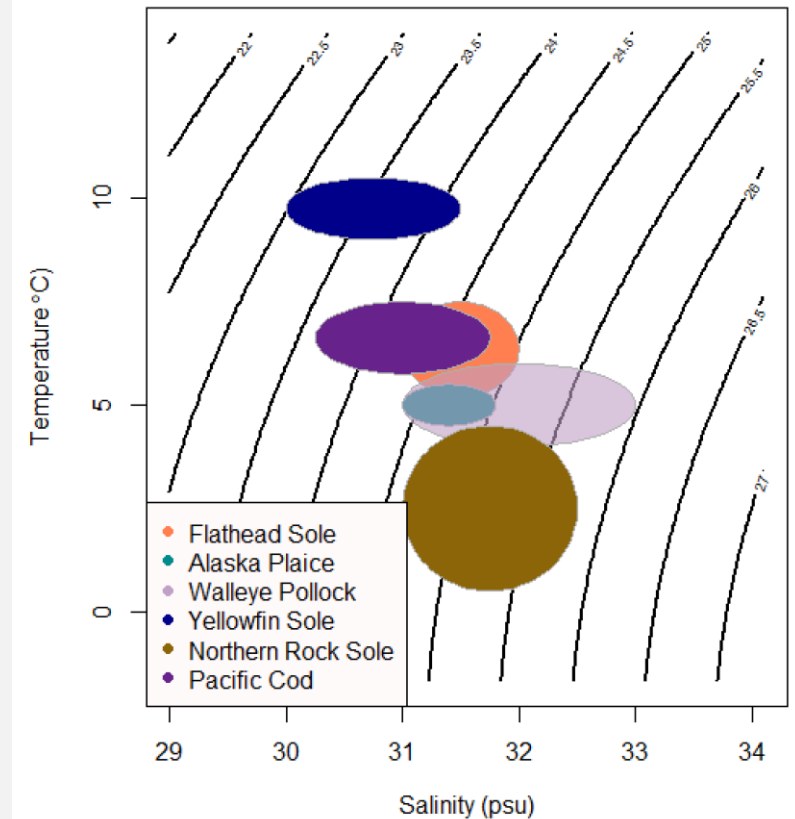
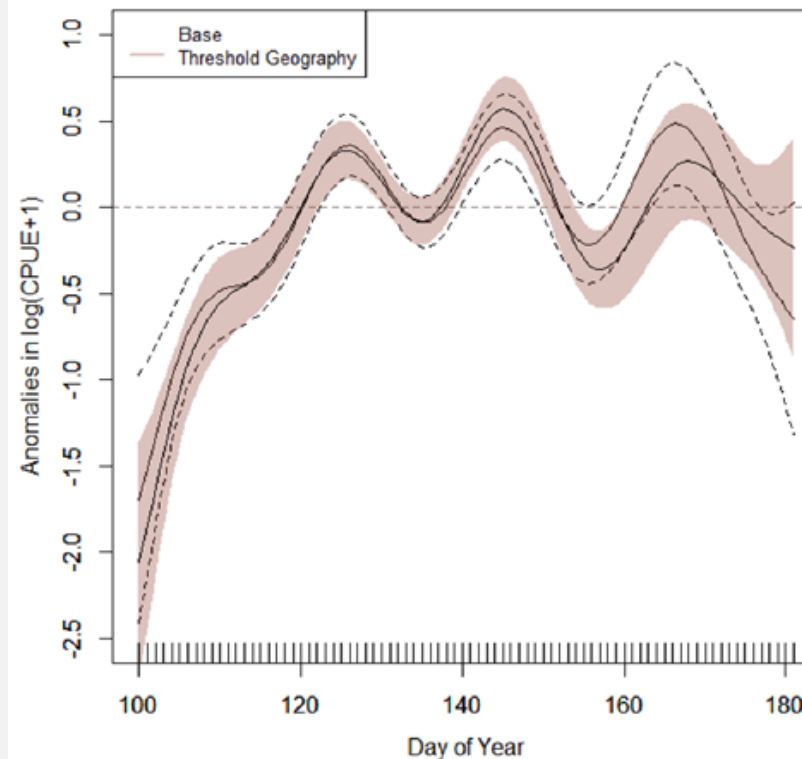
ACCOMPLISHED* GOAL: CLEAN FIGURES, MODEL OUTPUT TABLE

Species	Life Stage	Best Model	Δ MSE	Δ AIC	Deviance Explained
Alaska Plaice	Egg	$(\text{Log}(\text{Catch per } 10\text{m}^2)+1) = \text{factor}(\text{year}) + s(\text{doy}) + s(\text{bottom depth, } k=5) + g(\text{longitude, latitude, by} = \text{threshold regional SST}) + e_{\text{doy, year, (lon, lat)}}$	0.149	463.75	75%

Change in AP(E) Distribution w Threshold Temperature Effect



Alaska Plaice Phenology, Eggs



GOALS IN PROGRESS

Generate a
streamlined figure
process (that saves
figures too)

Write a style guide
for figures and code