

# Analysis of Atlantic Halibut Data, Reproduction

Richard McBride, George Maynard, and Emilee Tholke

August 9, 2022

```
knitr::opts_chunk$set(echo = TRUE)
library(lubridate); library(readxl); library(tidyverse)
```

## Introduction

This document serves as supplemental material for a study of Atlantic Halibut, *Hippoglossus hippoglossus*, sampling in the Gulf of Maine and subsequent reproductive analyses. George Maynard started this code, Emilee Tholke audited the data, and Rich McBride added some subroutines and brought this into an Rmarkdown document.

We begin by calling up and wrangling the source data file for this study. The data is for females only. Subsequent major sections are:

- Sample sizes by month, year, and sample source
- Female size and age by sample source and time of year
- Plotting a map of the female samples
- Female oogenesis and gonad development, including reproductive seasonality
- Size and age at maturity
- Skipping and first time spawners

```
## Set directory, read and wrangle the data
setwd ('~/Enhanced_ff/Data_Poor_halibut/Data_Rcode_NWAJFS_Manuscript')
datmat <- read_excel('E_Tholke_repro_halibut_females_v4.xlsx', sheet =
'E_Tholke_repro_halibut_females_')

# Fuss with the variable type and document units of measure
datmat$CollectionDate = as.Date(datmat$CollectionDate, "%m/%d/%Y")
datmat$Length = as.numeric(datmat$Length) # Fish length is in cm
datmat$Age = as.numeric(datmat$Age) # Fish age is in years
datmat$TUNICA_THICKNESS = as.numeric(datmat$TUNICA_THICKNESS) # in microns
datmat$gsi = as.numeric(datmat$gsi) # calculated as
GonadWeight/1000)/(Weight-Gonadweight/1000)*100

# Need a day of the year later in reproductive seasonality section
datmat <- datmat %>%
  mutate (doy = yday(CollectionDate)) %>%
  mutate (month = month(CollectionDate)) %>%
  mutate (year = year(CollectionDate)) %>%
  mutate(monfac = as.factor(month))
```

## Sample sizes by month, year, and sample source

There are 315 records of females. Here is a tabulation of the sample sizes by date (Month, Year) and by source (Agency, Gear).

```
FancyTable <- function(x, y) {  
temp.table <- addmargins(table(x, y), margin = seq_along(dim(table(x, y))),  
FUN = sum, quiet = TRUE)  
return(temp.table)  
}
```

```
FancyTable(datmat$month, datmat$year)
```

```
##      y  
## x    2014 2015 2016 2017 2018 sum  
## 1      0    0    0    0   20  20  
## 2      0    0    0    0   10  10  
## 4      0    2    2    3    3  10  
## 5      8    7    3   17   17  52  
## 6      0    0    0    3   23  26  
## 7      0    0    0    2   28  30  
## 8      0    0    0   21   13  34  
## 9      0    0    0   11   13  24  
## 10     7    5    4   29   16  61  
## 11     3    4    7   20   13  47  
## 12     0    0    0    1    0   1  
## sum   18   18   16  107  156 315
```

The number of samples was greatest in the years 2017-2018, in partnership with the Cape Cod Commercial Fisherman's Alliance. This partnership also improved seasonal coverage. Ten or more fish were sampled in all months except December (n=1) and March (n=0). This partnership with CCCFA was forged as a result of funding to The Nature Conservancy by a NOAA Staltonstall-Kennedy grant.

Fish were collected by a variety of private, state, and federal groups using different sampling gears.

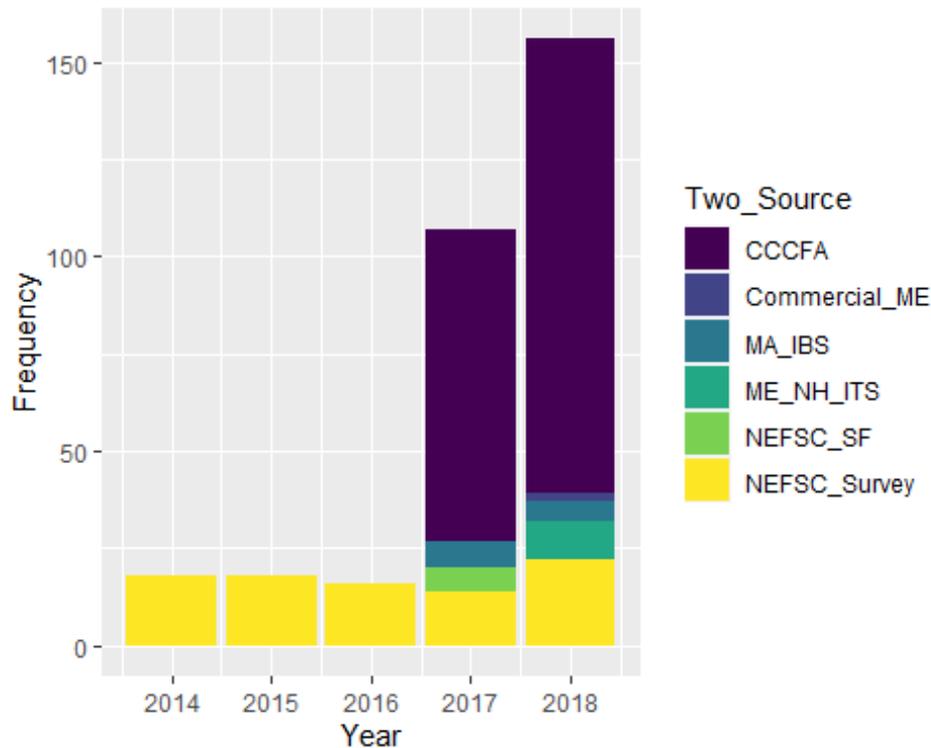
```
FancyTable(datmat$Two_Source, datmat$GEAR)
```

```
##      y  
## x    Benthic Longline Bottom Trawl Gillnet Rod Reel sum  
## CCCFA      1      43     153      0  197  
## Commercial_ME 2       0       0       0   2  
## MA_IBS        0      12       0       0  12  
## ME_NH_ITS     0      10       0       0  10  
## NEFSC_SF      0       5       0       1   6  
## NEFSC_Survey 40      48       0       0  88  
## sum          43     118     153      1 315
```

We dress these summaries up as colored figures, including summaries of size and age of the female halibut.

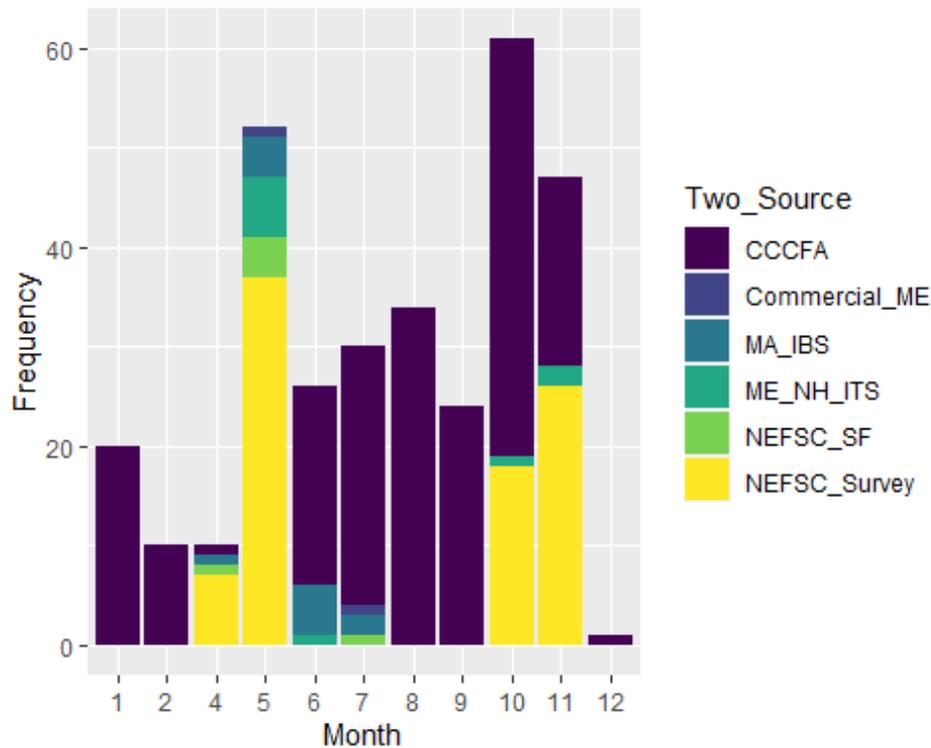
```
library (viridis)

ggplot (datmat) +
  geom_bar(mapping = aes (x = year, fill = Two_Source)) +
  #theme (Legend.position = 'top') +
  labs (x = 'Year', y = "Frequency") +
  scale_fill_viridis_d ()
```



```
ggsave ("HalFigYEARxGEAR.png")

ggplot (datmat) +
  geom_bar(mapping = aes (x = factor(month), fill = Two_Source)) +
  #theme (Legend.position = 'top') +
  labs (x = 'Month', y = "Frequency") +
  scale_fill_viridis_d ()
```



```
ggsave ("HalfFigMONTHxGEAR.png")
```

## Female size and age by sample source and time of year

Here is a summary of female size (Fork length, cm) and age (years) by sampling gear and month. We focus, in particular on female size as compared by a non-parametric Kruskal-Wallis test among the three dominant collecting gears.

```
library(ggpubr); library(psych)

describeBy(datmat$Length, datmat$GEAR, mat=T, skew=F, digits=3)[2:10]

##           group1 vars  n  mean   sd  min max range  se
## X11 Benthic Longline    1  43 103.233 21.349  42.5 146 103.5 3.256
## X12   Bottom Trawl     1 118  84.930 27.346  30.0 171 141.0 2.517
## X13      Gillnet       1 151  95.748 14.709  70.0 159  89.0 1.197
## X14      Rod Reel       1   1 130.000    NA 130.0 130   0.0  NA

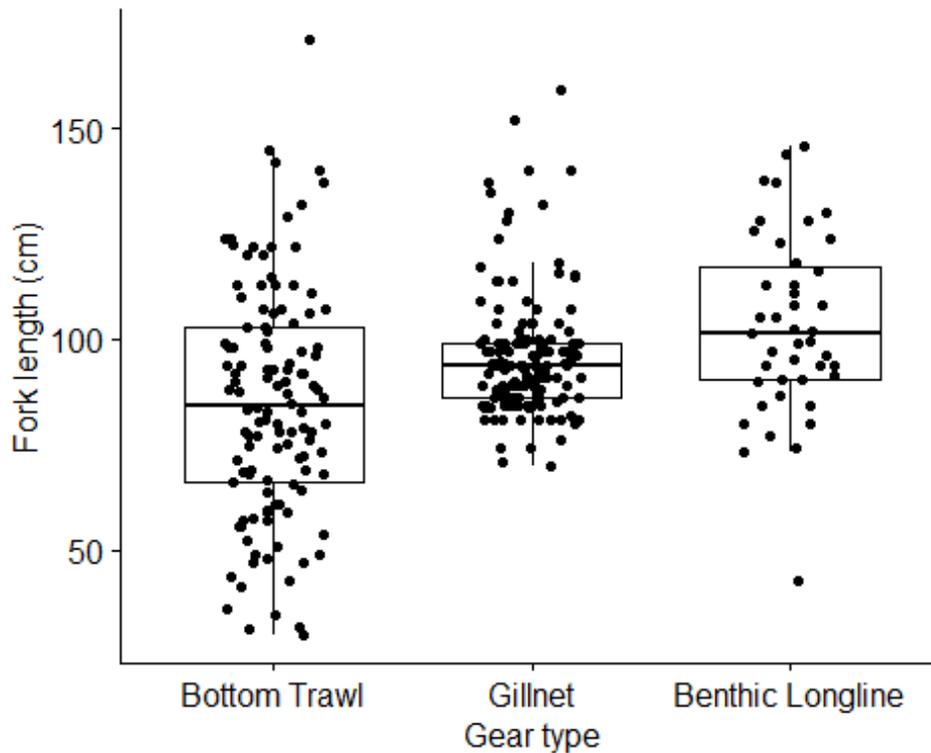
#In the following plot and its corresponding KW test,
#the single fish (130 cm) female collected by rod and reel has been removed.
datmat.no.RR <- datmat %>%
  filter (GEAR != 'Rod Reel')

kruskal.test(Length ~ GEAR, data = datmat.no.RR)

##
## Kruskal-Wallis rank sum test
```

```
##
## data: Length by GEAR
## Kruskal-Wallis chi-squared = 25.863, df = 2, p-value = 2.421e-06

ggboxplot (datmat.no.RR, x = "GEAR", y = "Length", add = "jitter") +
  labs (x = "Gear type", y = "Fork length (cm)")
```

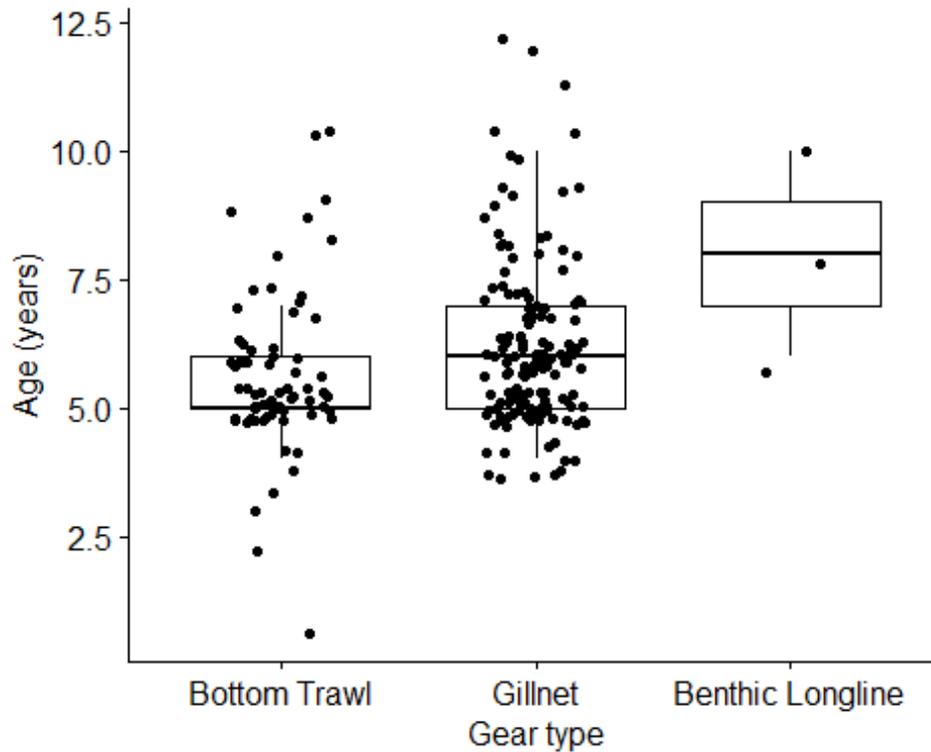


```
ggsave ("HalFig2Size.png")

kruskal.test(Age ~ GEAR, data = datmat.no.RR)

##
## Kruskal-Wallis rank sum test
##
## data: Age by GEAR
## Kruskal-Wallis chi-squared = 10.223, df = 2, p-value = 0.006028

ggboxplot (datmat.no.RR, x = "GEAR", y = "Age", add = "jitter") +
  labs (x = "Gear type", y = "Age (years)")
```



```
ggsave ("HalFig2Age.png")
```

Not surprising, the size range of female halibut varied significantly by gear, as a consequence of difference in mesh sizes, fishing regulations and permitted exceptions, and gear selectivity. In total, these gears sampled a wide range of fishes sizes.

We also include a quick check of ages (yr) by gear, as well as sizes and ages by sampling month.

```
describeBy (datmat$Age, datmat$GEAR, mat=T, skew=F, digits=3)[2:10]
```

```
##           group1 vars  n mean   sd min  max range  se
## X11 Benthic Longline  1   3 8.000 2.000  6  10    4 1.155
## X12 Bottom Trawl     1  69 5.594 1.584  1  10    9 0.191
## X13 Gillnet         1 149 6.168 1.583  4  12    8 0.130
## X14 Rod Reel        1   0  NaN   NA Inf -Inf -Inf  NA
```

```
describeBy (datmat$Length, datmat$month, mat=T, skew=F, digits=3)[2:10]
```

```
##      group1 vars  n   mean   sd min  max range  se
## X11     1     1  20 106.200 16.276 92.0 159  67.0 3.639
## X12     2     1  10  98.500 10.113 84.0 118  34.0 3.198
## X13     4     1  10 105.580 21.706 59.0 123  64.0 6.864
## X14     5     1  52  87.725 28.215 36.0 144 108.0 3.913
## X15     6     1  26  96.031 18.656 72.0 140  68.0 3.659
## X16     7     1  30  96.733 18.840 77.2 152  74.8 3.440
## X17     8     1  34  94.059 15.439 69.0 140  71.0 2.648
## X18     9     1  24  92.500 10.538 70.0 109  39.0 2.151
```

```
## X19      10      1 59  94.534 24.585 31.5 171 139.5 3.201
## X110     11      1 47  81.679 23.231 30.0 126  96.0 3.389
## X111     12      1  1  89.000      NA 89.0  89   0.0   NA

describeBy (datmat$Age, datmat$month, mat=T, skew=F, digits=3)[2:10]

##      group1 vars  n mean    sd min  max range  se
## X11      1     1 19 7.526 1.467  5  11    6 0.337
## X12      2     1 10 6.900 0.994  6   9    3 0.314
## X13      4     1  4 5.500 0.577  5   6    1 0.289
## X14      5     1 16 5.375 1.586  3   9    6 0.397
## X15      6     1 24 6.500 1.865  4  12    8 0.381
## X16      7     1 29 5.966 1.546  4  10    6 0.287
## X17      8     1 32 5.906 1.766  4  12    8 0.312
## X18      9     1 23 5.304 0.876  4   7    3 0.183
## X19     10     1 43 6.140 1.424  4  10    6 0.217
## X110    11     1 21 5.000 1.517  1   8    7 0.331
## X111    12     1  0  NaN    NA Inf -Inf -Inf   NA
```

## Plotting a map of the female samples

Finally, we plot the collection locations of female halibut by source and gear.

```
library(ggspatial)
library(mapdata); library(marmap)
library(raster); library(rgdal); library(rgeos)
library(rnaturalearth); library(rnaturalearthdata)
library(sf)

namerica <- ne_countries(scale = 'medium', type = 'countries', continent =
'north america', returnclass = "sf")

datmat.sf <- st_as_sf(datmat, coords =
c('CollectionLongitude','CollectionLatitude'), crs = 4326)

Trawl.sf <- datmat.sf %>% filter (GEAR == "Bottom Trawl")
Gillnet.sf <- datmat.sf %>% filter (GEAR == "Gillnet")
Longline.sf <- datmat.sf %>% filter (GEAR == "Benthic Longline")
RodReel.sf <- datmat.sf %>% filter (GEAR == "Rod Reel")

b <- getNOAA.bathy (lon1 = -71.5, lon2 = -64, lat1 = 39, lat2 = 45.5, res=1)
# 1 is the finest resolution; this requires internet connection
bathyLat = as.numeric(colnames(b)); bathyLon = as.numeric(rownames(b))
bathyZ = as.numeric(b); dim(bathyZ) = dim(b)
bf = fortify.bathy(b) # explixedly spatial?

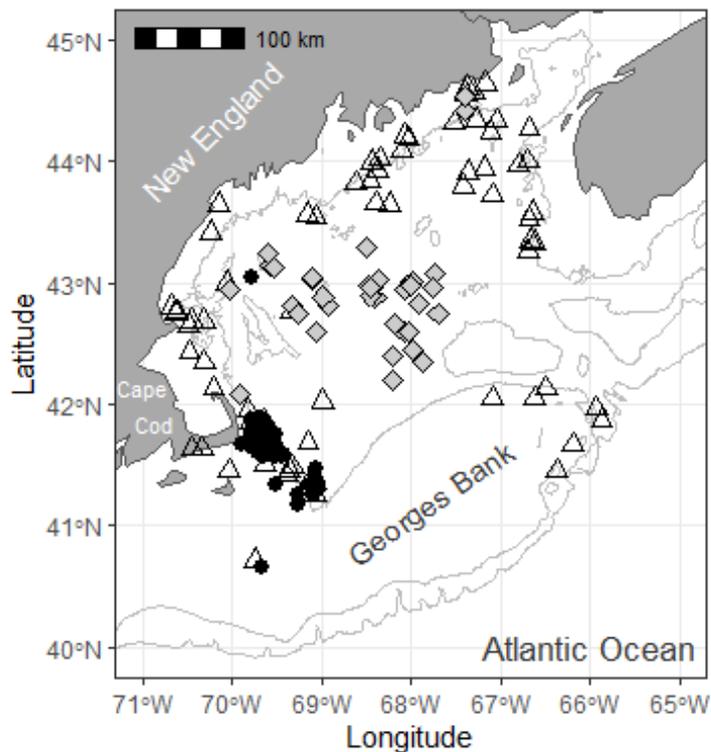
ggplot(namerica) + geom_contour(data=bf, aes(x=x, y=y, z=z), breaks=c(-100),
size=c(0.3), colour="grey") + # add 100m contour

  geom_sf(fill = "darkgray") +
```

```

geom_contour(data=bf, aes(x=x, y=y, z=z), breaks=c(-250), size=c(0.6),
colour="grey") + # add 250m contour
geom_sf(data=Trawl.sf, x = 'CollectionLongitude', y = 'CollectionLatitude',
pch = 2, size = 2.5) +
geom_sf(data=Gillnet.sf, x = 'CollectionLongitude', y =
'CollectionLatitude', pch = 19, size = 2.5) +
geom_sf(data=Longline.sf, x = 'CollectionLongitude', y =
'CollectionLatitude', pch = 23, fill="grey80", size = 2.5) +
# geom_sf(data=RodReel.sf, x = 'CollectionLongitude', y =
'CollectionLatitude', pch = 4, size = 2.5) +
coord_sf( xlim = c(-71, -65), ylim = c(40, 45) ) +
labs(x="Longitude", y="Latitude") +
theme_bw( ) +
annotate(geom = "text", x = -67.8, y = 41.2, label = "Georges Bank", angle
= 35, color = "grey22", size = 4) +
annotate(geom = "text", x = -70.25, y = 44.25, label = "New England",
angle = 45, color = "white", size = 4) +
annotate(geom = "text", x = -71, y = 42.15, label = "Cape", color =
"white", size = 3) +
annotate(geom = "text", x = -70.9, y = 41.85, label = " Cod", color =
"white", size = 3) +
annotate(geom = "text", x = -66, y = 40, label = "Atlantic Ocean", color
= "grey22", size = 4.5) +
ggspatial::annotation_scale(location = 'tl')

```



```

ggsave ("HalFig1Map.png")

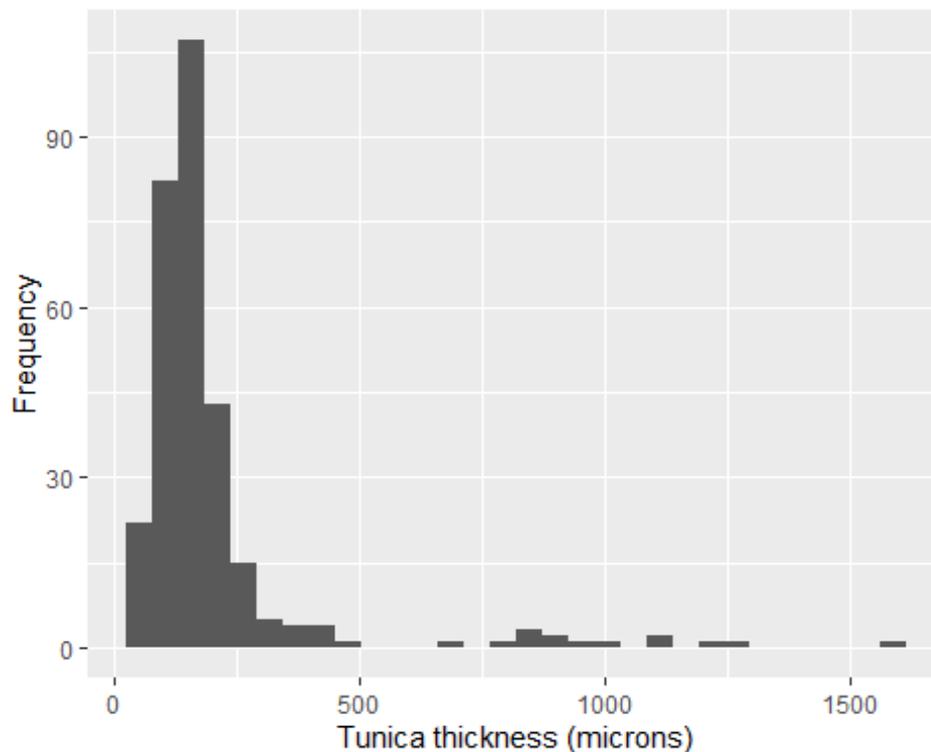
```

## Female Atlantic Halibut oogenesis and gonad development

The focus here is on halibut reproduction. We need to wrangle and examine some variables a bit more.

One of the criteria for maturity was the thickness of the gonad wall, also called the tunica. We assume a bimodal distribution in the tunica thickness is a result of immature fish that have a relatively thin gonad wall, whereas fish that have spawned in the past, including skippers (by definition), have a thick gonad wall as a result of the expansion, collapse, and reworking on the gonad from the process of spawning. While we expect a bimodal distribution, the specific measurements of what constitutes thin versus thick varies by species. We can identify that here because we actually measured the tunica thickness for most histology slides.

```
ggplot(data = datmat) +  
  geom_histogram(mapping = aes (x = TUNICA_THICKNESS)) +  
  labs (x = 'Tunica thickness (microns)', y = 'Frequency')
```



```
ggsave("HalFig5tunica.png")
```

```
describeBy (datmat$TUNICA_THICKNESS, datmat$TUNICA, mat=T, skew=F,  
digits=3)[2:10]
```

```
##      group1 vars   n  mean    sd  min  max range    se  
## X11      K    1  14 997.386 238.10 685.0 1576 891.0 63.635  
## X12     NA    1   0    NaN    NA  Inf -Inf -Inf    NA  
## X13      T    1 283 157.260  71.72  40.5  476 435.5  4.263
```

```
# T= thin, K = thick
```

These data suggest to set a maturity threshold for tunica thickness at 500 microns.

We are also curious about the progression of oogenesis in relation to fish size, fish age, and seasonality. Let's take a closer look before we apply that to assign a binary maturity value.

We are specifically curious about fish size and age thresholds relative to the timing of the major stages of oogenesis. Here MAOS is an acronym for "most advanced oocyte stage." We tabulate and illustrate the MAOS's by major stages observed in these samples of female halibut (below).

```
datmat <- datmat %>%
  transform(MAOS=factor(MAOS,levels=c("PG", "EC", "LC", "V1", "V2")))

describeBy (datmat$Length, datmat$MAOS, mat=T, skew=F, digits=3)[2:10]

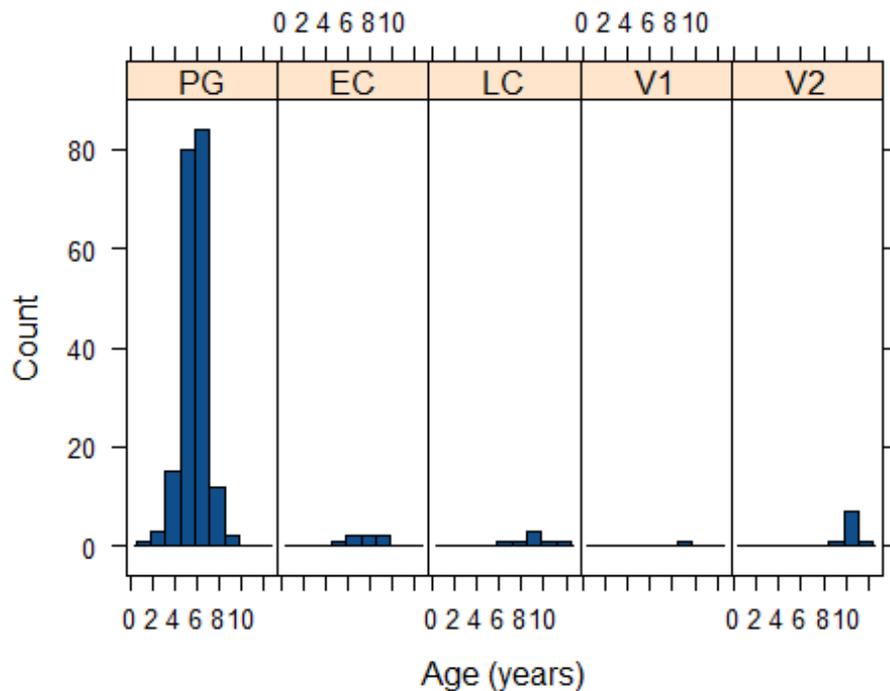
##      group1 vars   n   mean    sd  min  max range   se
## X11     PG    1 261  86.686 17.128  30.0 124.0  94.0 1.060
## X12     EC    1  14 104.250 17.834   63.5 130.0  66.5 4.766
## X13     LC    1  15 122.260  9.283 102.4 135.0  32.6 2.397
## X14     V1    1   6 127.650  8.500 117.0 137.9  20.9 3.470
## X15     V2    1  17 139.076 14.779 105.3 171.0  65.7 3.584

describeBy (datmat$Age, datmat$MAOS, mat=T, skew=F, digits=3)[2:10]

##      group1 vars   n   mean    sd min max range   se
## X11     PG    1 197   5.645 1.163   1  9     8 0.083
## X12     EC    1   7   7.571 1.397   5  9     4 0.528
## X13     LC    1   7   9.000 1.826   6 12     6 0.690
## X14     V1    1   1   9.000  NA    9  9     0  NA
## X15     V2    1   9 10.222 0.833   9 12     3 0.278

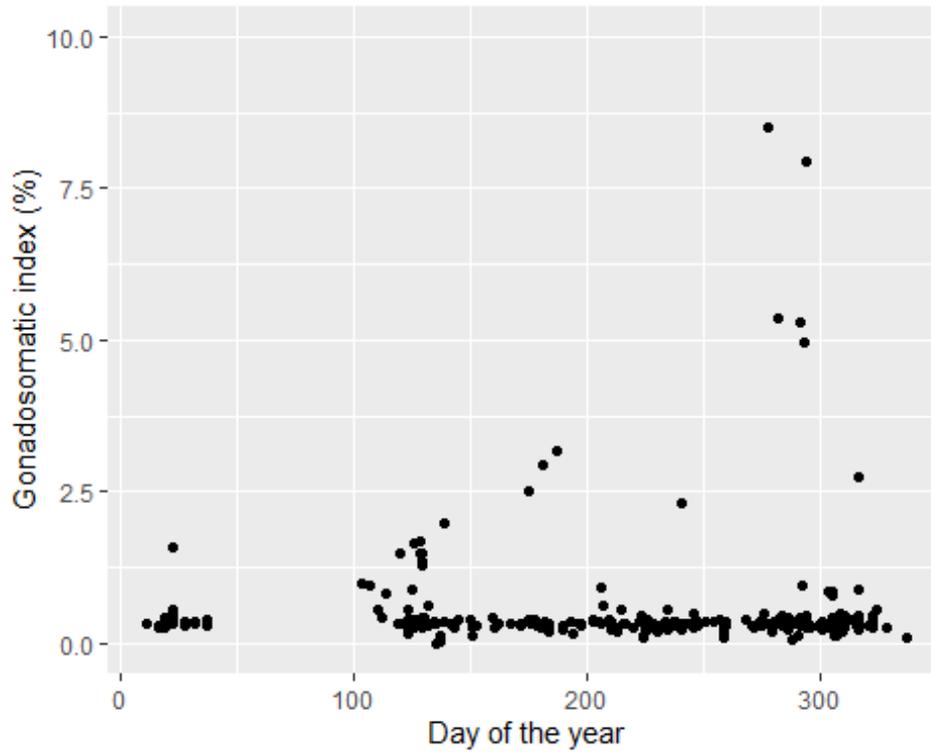
library (lattice)

histogram(~Age|MAOS, data=datmat, # breaks=seq(0,12,5),
          type="count", layout=c(5,1), col="dodgerblue4",
          xlab="Age (years)")
```

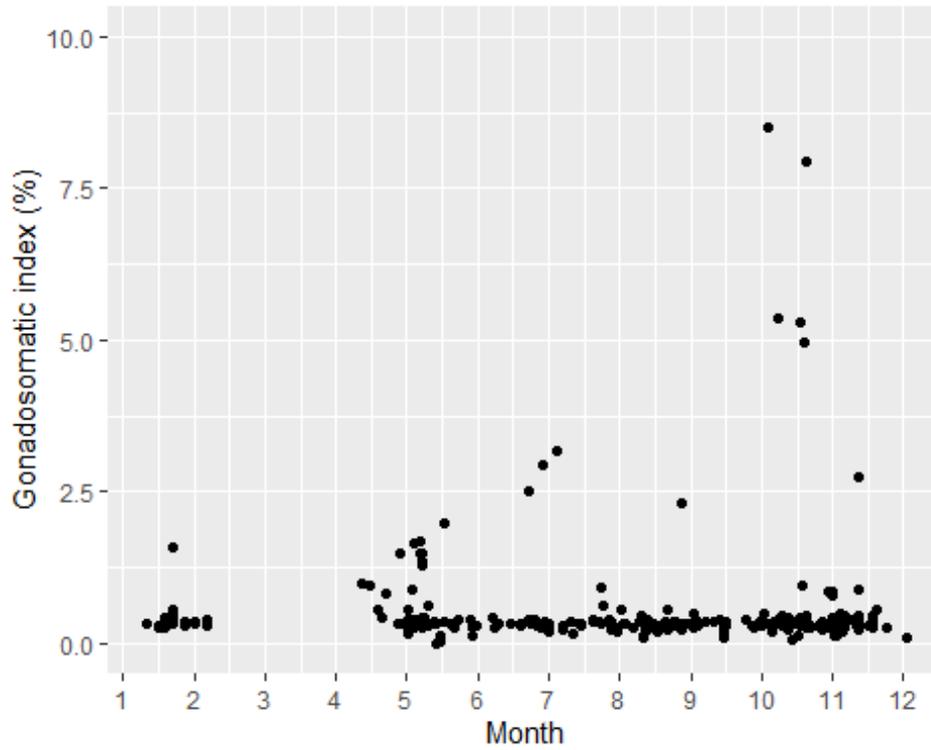


It appears that we did not get fish in spawning condition, which would be evident by individual females with mature oocytes (i.e., nucleus migration or breakdown not observed). Still, the gonadosomatic index ranged up to 8.51 and we can use this index to illustrate a general pattern of reproductive seasonality, and compare that to a table of most advance oocyte stage.

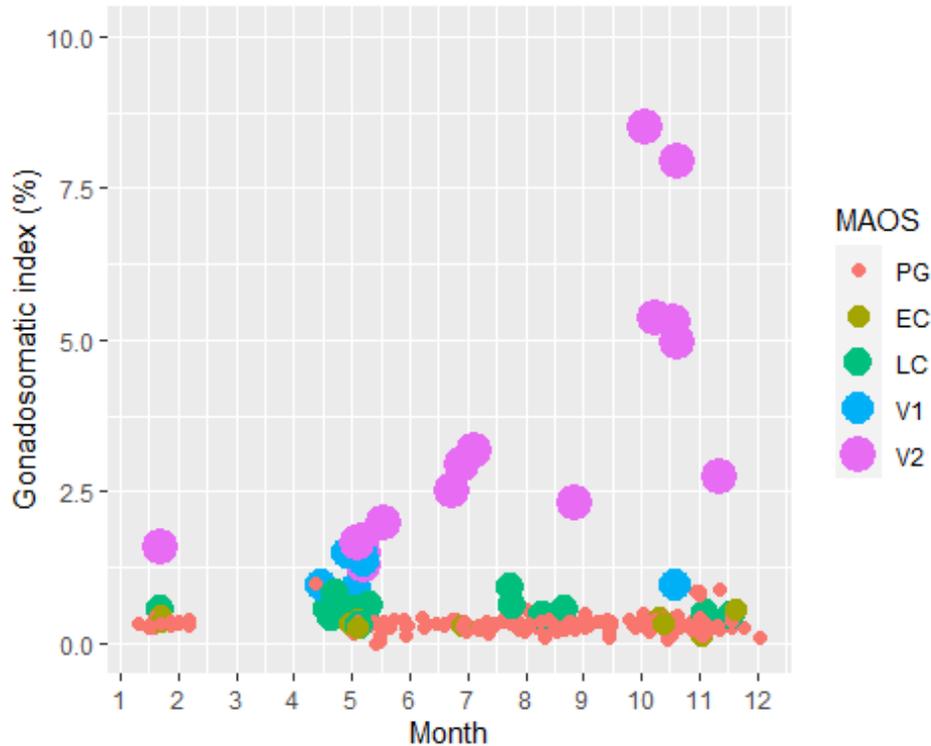
```
ggplot (datmat) +
  geom_point (aes(x = doy, y = gsi)) +
  labs (x = "Day of the year", y = 'Gonadosomatic index (%)') +
  ylim ( c(0,10))
```



```
ggplot (datmat) +  
  geom_point (aes(x = doy, y = gsi)) +  
  labs (x = "Month", y = 'Gonadosomatic index (%)') +  
  scale_x_continuous(breaks = seq(1, 366, by=30.4), labels =1:13) +  
  ylim ( c(0,10))
```



```
ggplot (data = datmat, mapping = aes(x = doy, y = gsi, group=MAOS)) +  
  geom_point (mapping = aes(color = MAOS, size = MAOS)) +  
  labs (x = "Month", y = 'Gonadosomatic index (%)') +  
  scale_x_continuous(breaks = seq(1, 366, by=30.4), labels =1:13) +  
  ylim ( c(0,10))
```



```
describeBy (datmat$gsi, datmat$MAOS, mat=T, skew=F, digits=3)[2:10]
```

##	group1	vars	n	mean	sd	min	max	range	se	
##	X11	PG	1	248	0.321	0.111	0.000	0.984	0.984	0.007
##	X12	EC	1	13	0.358	0.101	0.142	0.574	0.432	0.028
##	X13	LC	1	14	0.556	0.163	0.305	0.938	0.633	0.044
##	X14	V1	1	6	1.187	0.288	0.882	1.490	0.607	0.117
##	X15	V2	1	16	3.467	2.292	1.288	8.514	7.226	0.573

```
ggsave("HalFig8GSI2.png")
```

You can see that the index increases from spring to winter for a subset of fish, while the rest of the fish have a low index, < 1. Across all months, most fish have a MAOS of only PG, and there is no evidence these PG females are mature. The few 'EC' females have a similarly low gsi as PG females. There are, however, notable increases in gsi for LC females and certainly for V1 females, the mean gsi is > 1.

Will females with a MAOS of LC advance a cohort to ovulation in one year? To help us address this, let's look at the seasonality of the oocyte stages.

```
MAOSxMonth <- FancyTable(datmat$MAOS, datmat$month)
MAOSxMonth
```

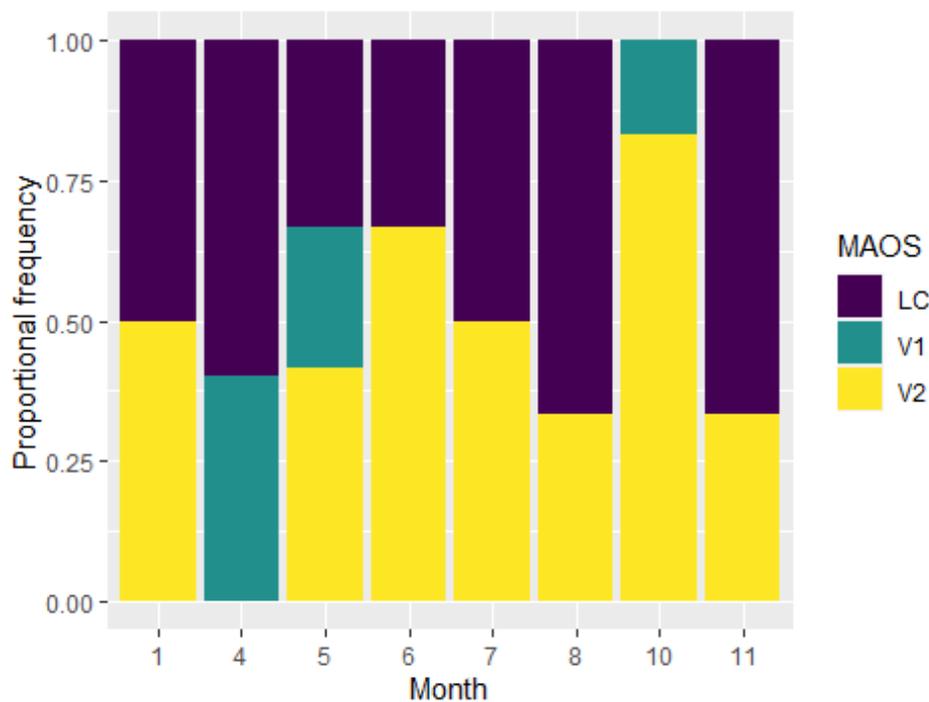
##		y												
##	x	1	2	4	5	6	7	8	9	10	11	12	sum	
##	PG	16	10	4	34	22	26	31	24	53	42	1	263	
##	EC	2	0	1	6	1	0	0	0	2	2	0	14	

```
## LC 1 0 3 4 1 2 2 0 0 2 0 15
## V1 0 0 2 3 0 0 0 0 1 0 0 6
## V2 1 0 0 5 2 2 1 0 5 1 0 17
## sum 20 10 10 52 26 30 34 24 61 47 1 315
```

```
#write.csv(MAOSxMonth, "MAOSxMonth.csv", row.names = TRUE)
```

```
datmat.noPG <- datmat %>% filter (MAOS != 'PG' & MAOS != 'EC')
```

```
ggplot (datmat.noPG) +
  geom_bar(mapping = aes (x = monfac, fill = MAOS), position='fill') +
  labs (x = 'Month', y = "Proportional frequency",
  title = '') +
  scale_fill_viridis_d ()
```



The seasonality of vitellogenesis appears to begin in spring, as monthly modes of V1 and V2 are both in May in the table, above. However the sample sizes of LC, V1 and V2 females are modest, and there are no females with mature oocytes, so the seasonal evolution of an advancing, vitellogenic cohort of oocytes is ambiguous.

We can confirm that no females with MAOS=early cortical alveoli stage (EC) are associated with either a thick tunica or POFs. Therefore, in terms of oogenesis, maturity is triggered by fish with a MAOS as either LC or V1 advancing to oocyte maturation and ovulation. It may be the period that this requires is 1 or 2 years.

```

datmat %>%
  filter(TUNICA == 'K' | POF_PRESENCE == "Y") %>%
  dplyr::select (HistologyID, MAOS, POF_PRESENCE, TUNICA, TUNICA_THICKNESS)

```

##	HistologyID	MAOS	POF_PRESENCE	TUNICA	TUNICA_THICKNESS
## 1	HHI043	V2	N	K	1000.0
## 2	HHSK006	V2	NA	K	822.0
## 3	HHSK085	V1	NA	K	904.0
## 4	HHSK104	V2	Y	K	822.0
## 5	HHSK137	V2	NA	K	863.0
## 6	HHSK188	V2	NA	K	685.0
## 7	HHSK205	V2	NA	K	1274.0
## 8	HHSK225	V2	N	K	932.0
## 9	HHSK211	V2	NA	K	1096.0
## 10	HHSK231	LC	N	K	1233.0
## 11	LLHHI3886	V1	Y	K	1576.0
## 12	LLHHI3888	V1	N	K	1096.0
## 13	LLHHI4709	V2	N	K	782.0
## 14	LLHHI6985	V2	N	K	878.4

Although there are 15 females with a MAOS of 'LC,' only one of them (7%) has an indication of past maturity (i.e., a thick gonad); as we will see in the last section, this fish is a skip spawner. In contrast, there are 6 females with a MAOS of 'V1,' and 3 of them (50%) had an indication of past maturity. A limited number of females were examined for oocyte diameters (see McGuire et al., 2020; Final report to NOAA Fisheries 2016 Saltonstall Kennedy Award Number NA16NMF4270248), but these plots did not clarify if fish with a MAOS of either LC or V1 would spawn within one calendar year.

As a standing hypothesis, a cohort of LC oocytes in spring will require 2 years to spawn and a cohort of V1 oocytes in spring will require 1 year to spawn. This has been demonstrated for Greenland halibut (*Reinhardtius hippoglossoides*) by Kennedy et al. (2011; <https://doi.org/10.1139/F10-149>).

To test the sensitivity of this, we will compare maturity schedules between assignment of a binary maturity value according to this standing hypothesis, versus an alternative hypothesis that cohorts of either LC or V1 will spawn within one year, which would likely lead to a smaller and younger median maturity estimate. Here we select both data sets.

```

datmat <- datmat %>%
  mutate (MAT01 = NA) %>% # The decision on binary maturity assignment
  mutate (MAT01 = ifelse( MAOS == 'V1' | MAOS == "V2" |
    POF_PRESENCE == "Y" | TUNICA == 'K', 1, 0))

LC.mat.check <- datmat %>% # save for later to check sensitivity
  mutate (MAT01 = NA) %>%
  mutate (MAT01 = ifelse(MAOS == 'LC' | MAOS == 'V1' | MAOS == "V2" |
    POF_PRESENCE == "Y" | TUNICA == 'K', 1, 0)) %>%
  transform(MAOS=factor(MAOS,levels=c("PG", "EC", "LC", "V1", "V2")))

looksee1 <- datmat %>%

```

```

  filter (MAT01 == 1) %>%
  dplyr::select (HistologyID, MAOS, POF_PRESENCE, TUNICA, MAT01)
nrow(looksee1); head(looksee1)

## [1] 24

##   HistologyID MAOS POF_PRESENCE TUNICA MAT01
## 1   HHI034   V1          N      NA     1
## 2   HHI039   V2          N      T     1
## 3   HHI043   V2          N      K     1
## 4   HHI090   V1          N      T     1
## 5   HHSK006  V2          NA     K     1
## 6   HHSK085  V1          NA     K     1

looksee2 <- LC.mat.check %>%
  filter (MAT01 == 1) %>%
  dplyr::select (HistologyID, MAOS, POF_PRESENCE, TUNICA, MAT01)
nrow(looksee2)

## [1] 38

```

We find that there are 24 mature fish in total using the base threshold of the advancing cohort needs to be in vitellogenesis.

### Size and age at maturity

We will examine size and age at maturity of the females twice, once with a canned program (sizeMat), which does bootstrapping and has some very nice diagnostics, and a second way to plot a final figure for the publication using some organically written code.

```

library(MASS)
library(car)
library (sizeMat)

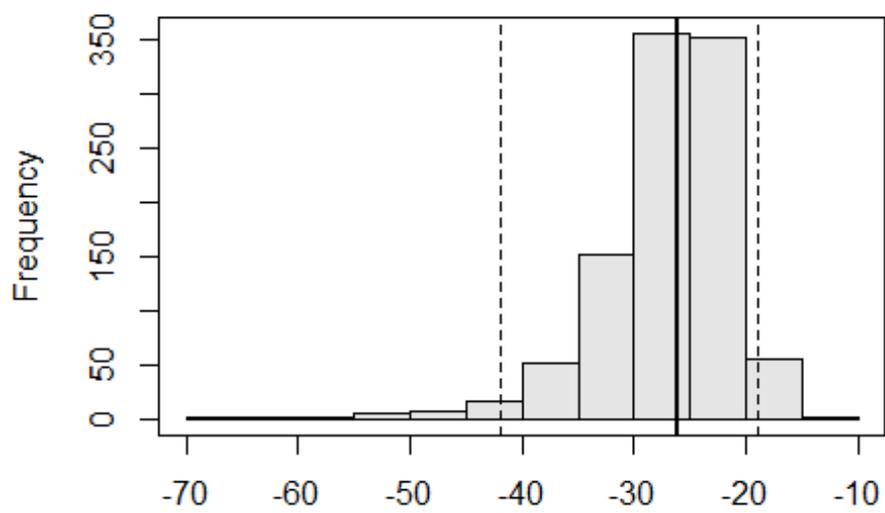
# set NA free matrices for both length, age, and maturity status

datmat_L <- datmat %>%
  drop_na(Length) %>%
  drop_na(MAT01)

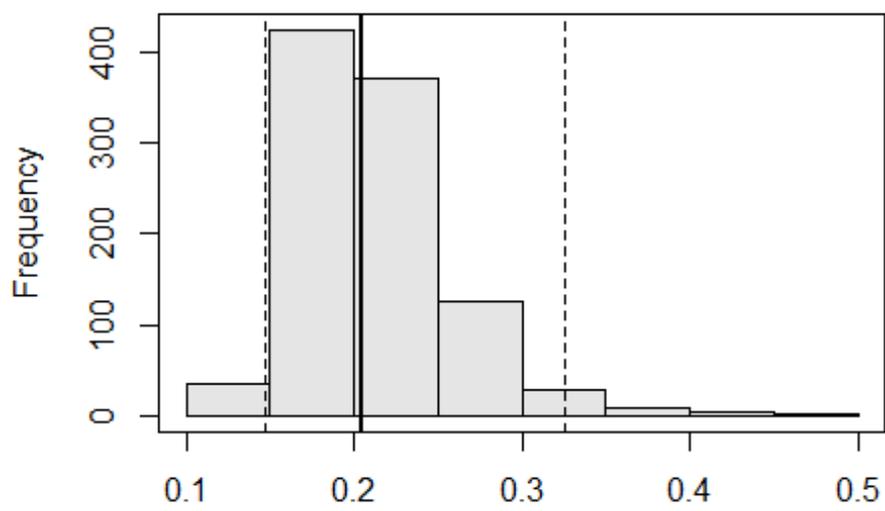
# quick check with sizeMat, use bootstrap intervals

L_ogive_fq <- gonad_mature(datmat_L, varNames = c("Length", "MAT01"), inmName
= 0, matName = 1, method="fq", niter = 1000)
plot(L_ogive_fq)

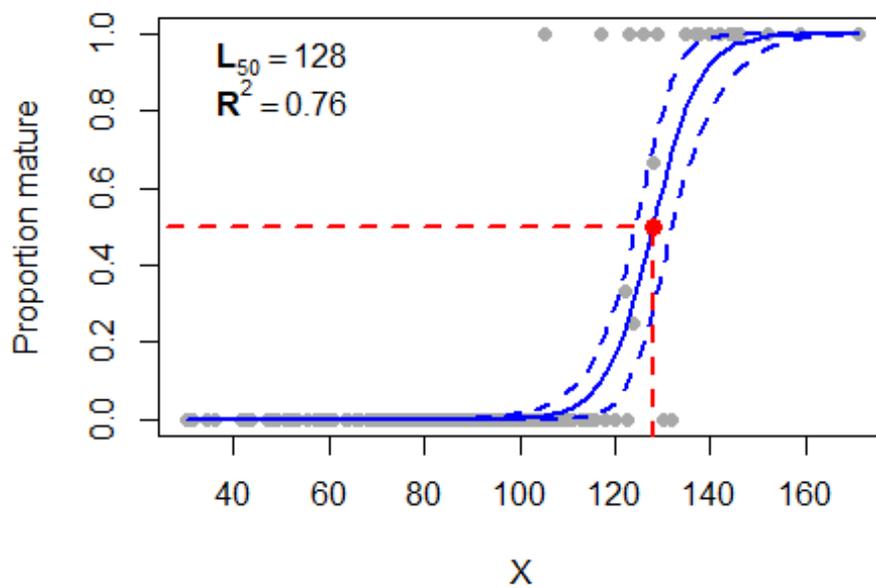
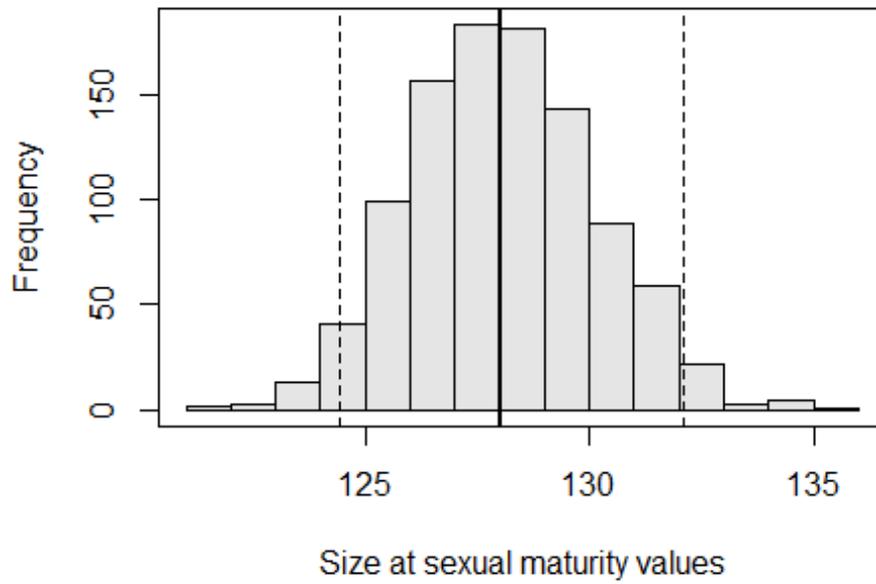
```



A

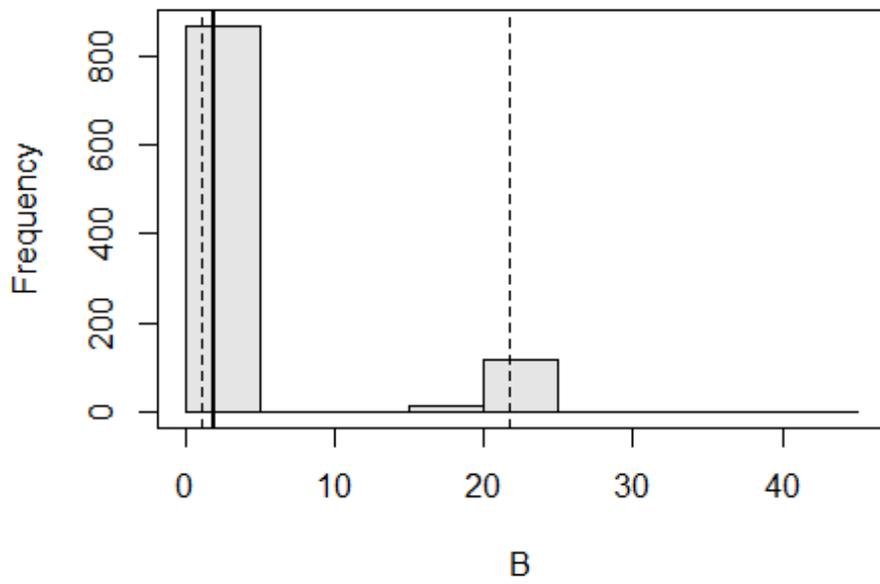
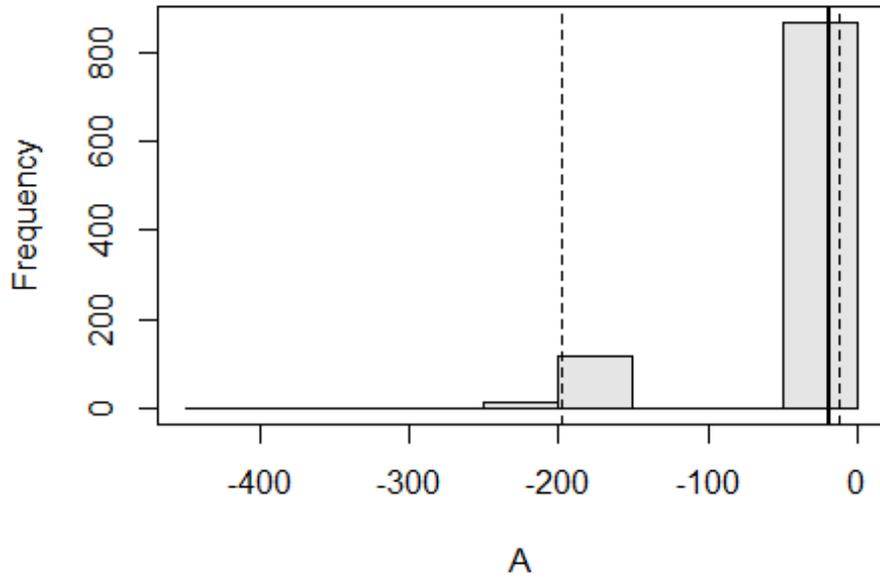


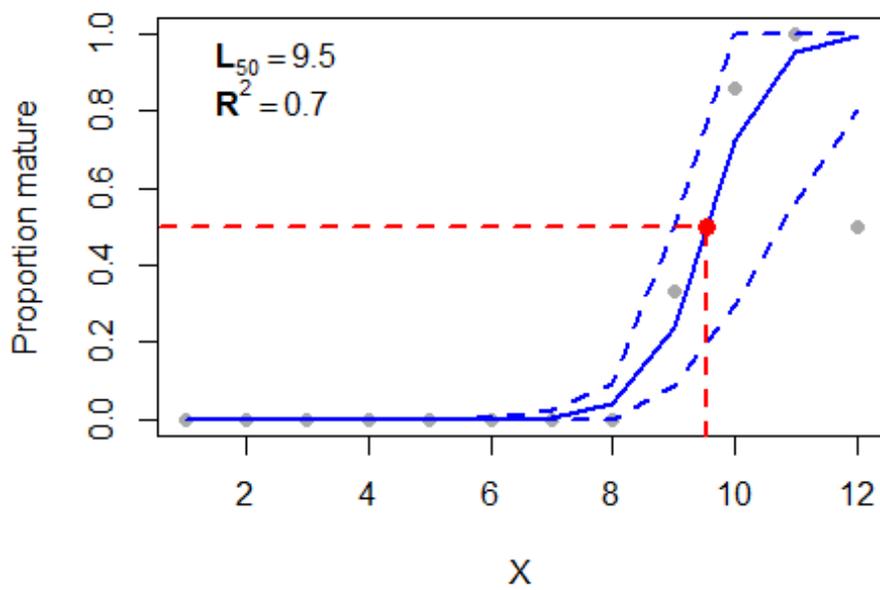
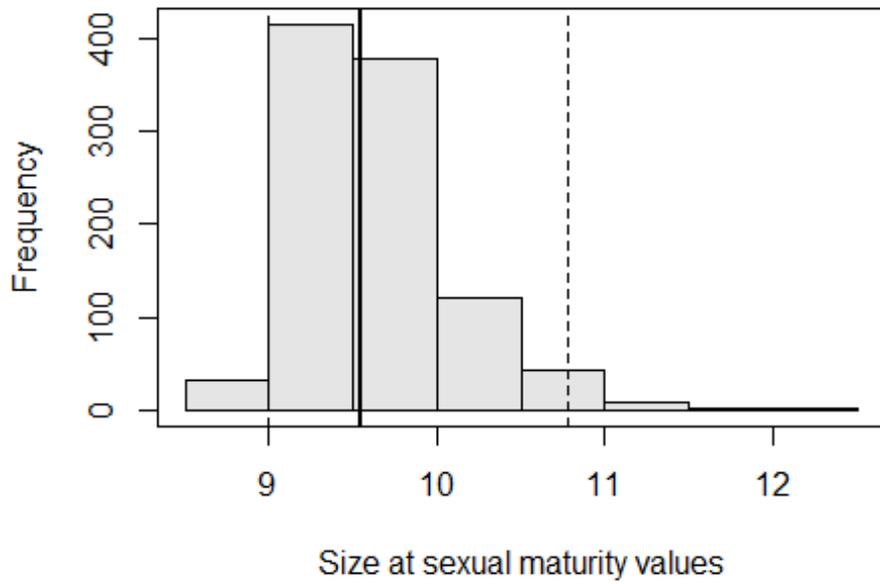
B



```
## Size at gonad maturity = 128
## Confidence intervals = 124.4 - 132.1
## Rsquare = 0.76
```

```
datmat_A <- datmat %>%  
  drop_na(Age) %>%  
  drop_na(MAT01)  
  
A_ogive_fq <- gonad_mature(datmat_A, varNames = c("Age", "MAT01"), inmName =  
0, matName = 1, method="fq", niter = 1000)  
plot(A_ogive_fq)
```





```
## Size at gonad maturity = 9.5
## Confidence intervals = 9 - 10.8
## Rsquare = 0.7
```

Here, I want to completely control the format of the image for the publication, so I have redone the plotting and bootstrapping.

*#NB: you need to run the previous check to get the boot strap parameters to roll over to this chunk*

```
par (mfcol=c(2,1))
par(mai = c(0.5, 1.5, 0, 0))
par (las=1)

# Length first, use all data (all months)
maxL <- 180
plotL1 <- 30
plotL2 <- 3
predL <- function(cf,x) exp(cf[1]+cf[2]*x)/(1+exp(cf[1]+cf[2]*x))
lens <- seq(1,maxL,1)

#The Logit model
#fitting and bootstrapping
logitETt <- glm (MAT01 ~ Length, data =datmat_L, family=binomial);
summary (logitETt)

##
## Call:
## glm(formula = MAT01 ~ Length, family = binomial, data = datmat_L)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.51500  -0.08236  -0.03102  -0.00979   2.98824
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -25.03595    4.72152  -5.303 1.14e-07 ***
## Length       0.19547    0.03784   5.166 2.39e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 169.382  on 312  degrees of freedom
## Residual deviance:  49.363  on 311  degrees of freedom
## AIC: 53.363
##
## Number of Fisher Scoring iterations: 9

BLETt <- matrix (NaN, maxL, 2); bcE <- bootCase(logitETt, B=1000)
for (i in 1:maxL){BLETt[i,] <- as.numeric(quantile(apply(bcE,1,predL,x=i),
c(0.025, 0.975))))}

#   insert plotting for Length here
```

```

pETt <- predict(logitETt,data.frame(Length=lens),type="response")
plot (lens, pETt, type = 'l', col = 1, xlab = '', ylab = '', xlim = c(0,
maxL), ylim = c(0,1), main = "")
  lines (lens, BLETt[,1], type = 'l', col = 4, lty=2, xlab = '', ylab = '',
xlim = c(0, maxL), ylim = c(0,1), main = "")
  lines (lens, BLETt[,2], type = 'l', col = 4, lty=2, xlab = '', ylab =
'', xlim = c(0, maxL), ylim = c(0,1), main = "")

rug ((datmat_L$Length[datmat_L$MAT01==0]))
rug ((datmat_L$Length[datmat_L$MAT01==1]), side=3)
# arrows(67,.85,67,1, col="red")

min (datmat_L$Length)
## [1] 30

max (datmat_L$Length)
## [1] 171

MatL.N <- length (datmat_L$Length)
MatL.025 <- signif(quantile(L_ogive_fq$L50_boot, probs = 0.025, na.rm =
TRUE), 3)
MatL.50 <- signif(quantile(L_ogive_fq$L50_boot, probs = 0.5, na.rm =
TRUE), 4)
MatL.975 <- signif(quantile(L_ogive_fq$L50_boot, probs = 0.975, na.rm =
TRUE), 3)

text (plotL1, 0.95, "Maturity at Length (cm)")
text(plotL1, 0.80, bquote(italic(n)==.(MatL.N)))
text (plotL1, 0.3, bquote(italic(L)[50]==.(MatL.50)~cm))
text (plotL1, 0.15, bquote(. (MatL.025)~" - "~.(MatL.975)~ "cm 95%
c.l.))
mtext ("Proportional mature", 2, at = -0.1, line = 3, las = 3)

# Age next, use all data (all months)

par (las=1)
predL <- function(cf,x) exp(cf[1]+cf[2]*x)/(1+exp(cf[1]+cf[2]*x))
lens <- seq(1,20,1)

#The logit model
#fitting and bootstrapping
logitETt <- glm (MAT01 ~ Age, data =datmat_A, family=binomial); summary
(logitETt)

##
## Call:
## glm(formula = MAT01 ~ Age, family = binomial, data = datmat_A)
##

```

```

## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.97653 -0.05207 -0.02079 -0.02079  1.66233
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -17.6234     4.2312  -4.165 3.11e-05 ***
## Age          1.8368     0.4643   3.956 7.62e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 87.449  on 220  degrees of freedom
## Residual deviance: 29.939  on 219  degrees of freedom
## AIC: 33.939
##
## Number of Fisher Scoring iterations: 9

      BLEtT <- matrix (NaN, 20, 2); bcE <- bootCase(logitETt, B=1000)

      for (i in 1:20){BLEtT[i,] <-
as.numeric(quantile(apply(bcE,1,predL,x=i), c(0.025, 0.975))))}

      #      insert plotting for length here
      pETt <- predict(logitETt,data.frame(Age=lens),type="response")
      plot (lens, pETt, type = 'l', col = 1, xlab = '', ylab = '', xlim =
c(0, 20), ylim = c(0,1), main = "")
      lines (lens, BLEtT[,1], type = 'l', col = 4, lty=2, xlab = '', ylab =
'', xlim = c(0, 20), ylim = c(0,1), main = "")
      lines (lens, BLEtT[,2], type = 'l', col = 4, lty=2, xlab = '', ylab =
'', xlim = c(0, 20), ylim = c(0,1), main = "")

      rug (jitter(datmat_A$Age[datmat_A$MAT01==0]))
      rug (jitter(datmat_A$Age[datmat_A$MAT01==1]), side=3)
      # arrows(67,.85,67,1, col="red")

      min (datmat_A$Age)

## [1] 1

      max (datmat_A$Age)

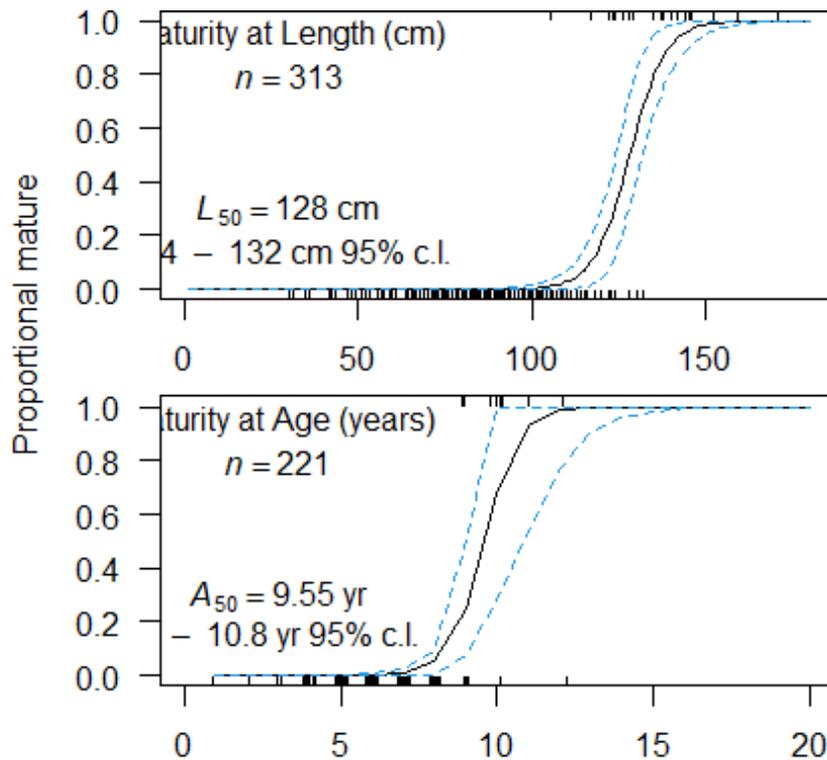
## [1] 12

      MatA.N <- length (datmat_A$Age)
      MatA.025 <- signif(quantile(A_ogive_fq$L50_boot, probs = 0.025, na.rm =
TRUE), 2)
      MatA.50 <- signif(quantile(A_ogive_fq$L50_boot, probs = 0.5, na.rm =
TRUE), 3)
      MatA.975 <- signif(quantile(A_ogive_fq$L50_boot, probs = 0.975, na.rm =

```

```
TRUE), 3)
```

```
text(plotL2, 0.95, "Maturity at Age (years)")
text(plotL2, 0.80, bquote(italic(n)==.(MatA.N)))
text(plotL2, 0.3, bquote(italic(A)[50]==.(MatA.50)~yr))
text(plotL2, 0.15, bquote(.(MatA.025)~" - "~.(MatA.975)~ "yr 95%
c.l."))
```



*#need to save figure manually, HalFig10.0give.png*

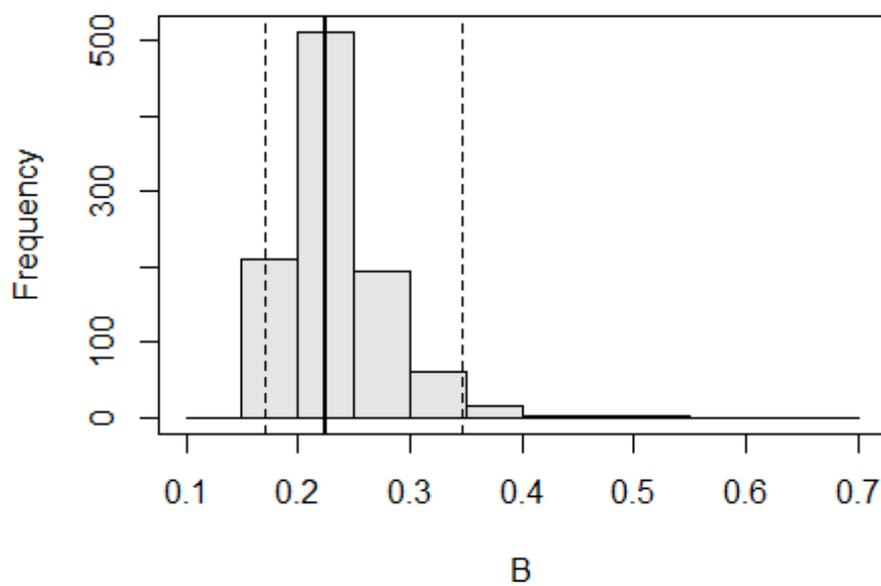
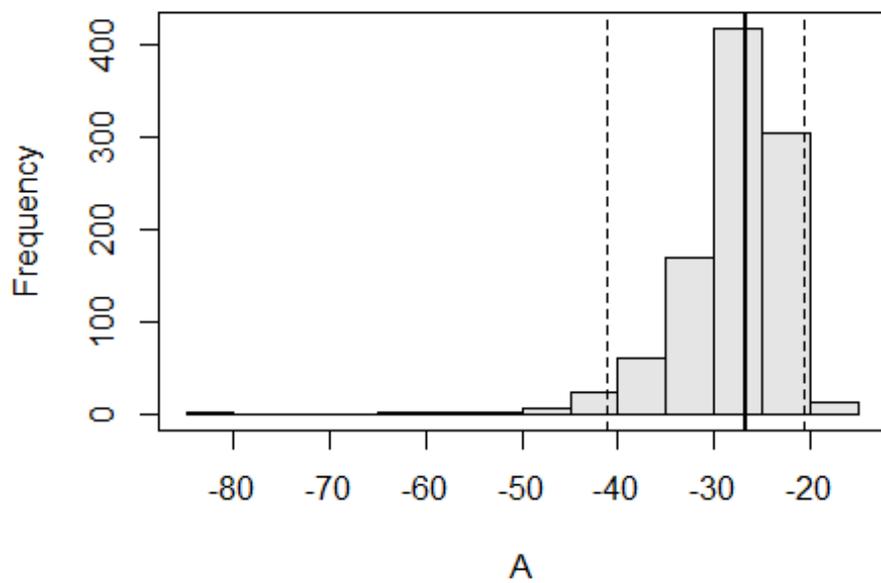
I want to also check for sensitivity of choosing V1 as the lower threshold for assigning maturity by MAOS. Here I will rerun using the alternative maturity assignments (i.e., a cohort of MAOS=LC will advance to spawn within 1 year).

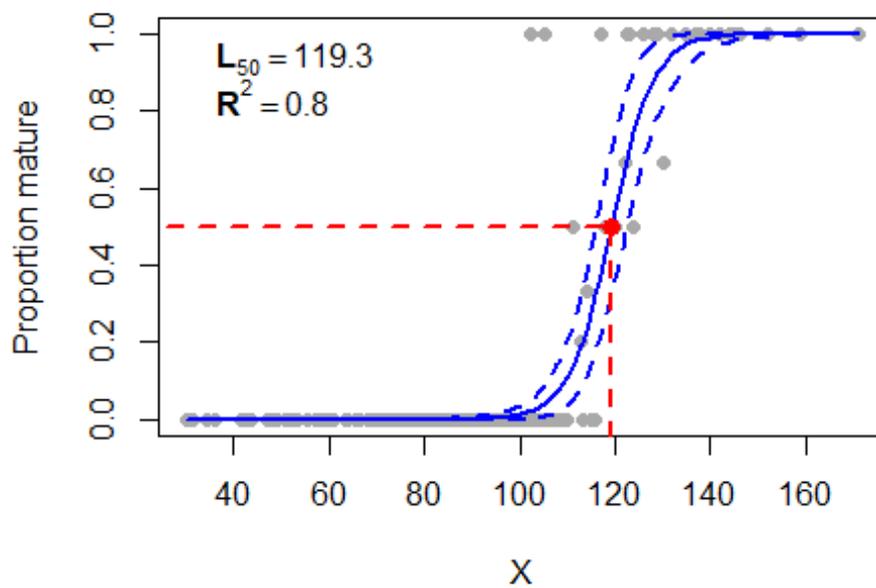
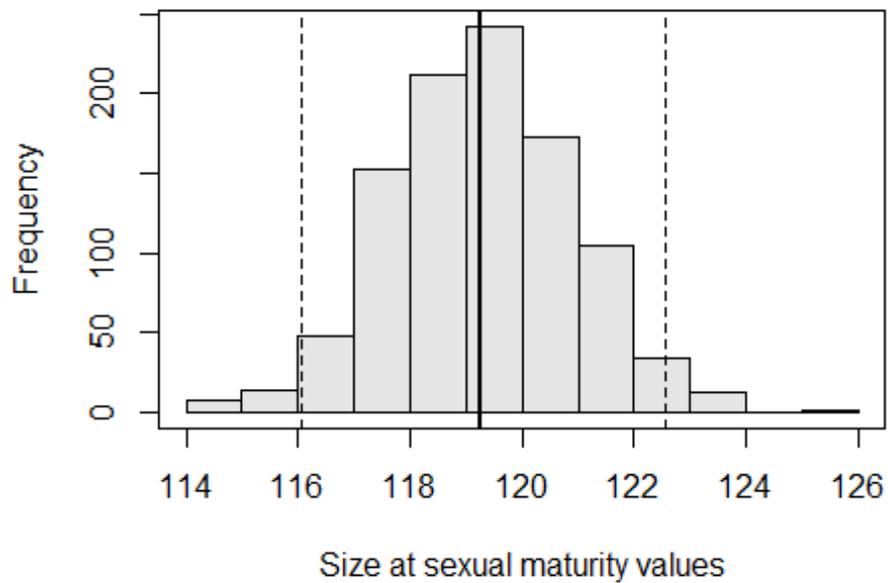
```
# set NA free matrices for both length, age, and maturity status
```

```
datmat_L <- LC.mat.check %>%
  drop_na(Length) %>%
  drop_na(MAT01)
```

```
# quick check with sizeMat, use bootstrap intervals
```

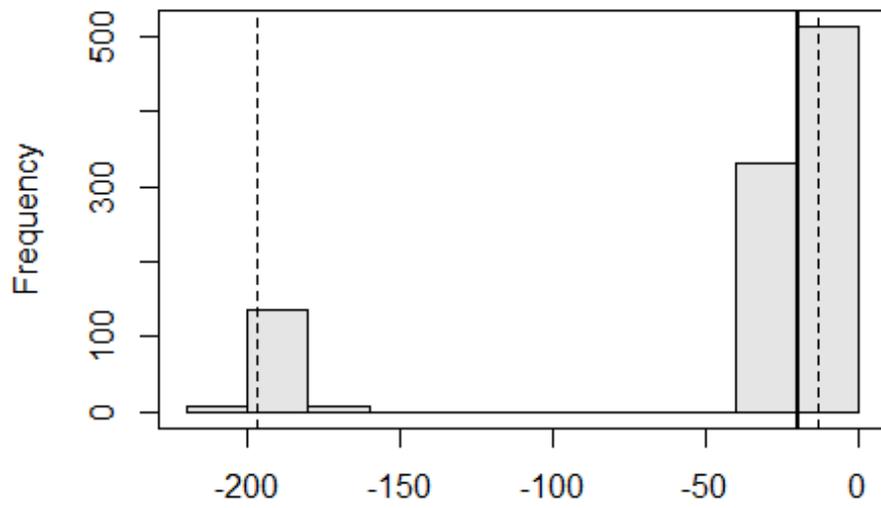
```
L_ogive_fq <- gonad_mature(datmat_L, varNames = c("Length", "MAT01"), inmName
= 0, matName = 1, method="fq", niter = 1000)
plot(L_ogive_fq)
```



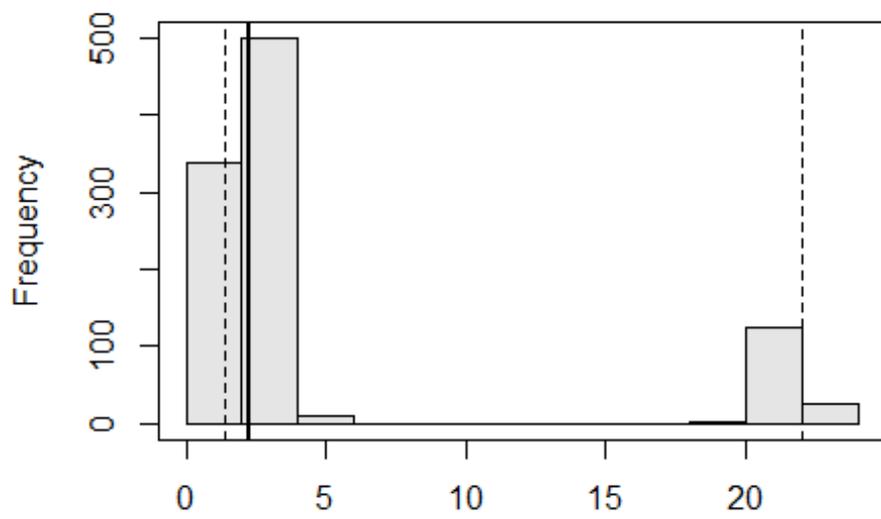


```
## Size at gonad maturity = 119.3
## Confidence intervals = 116 - 122.6
## Rsquare = 0.8
```

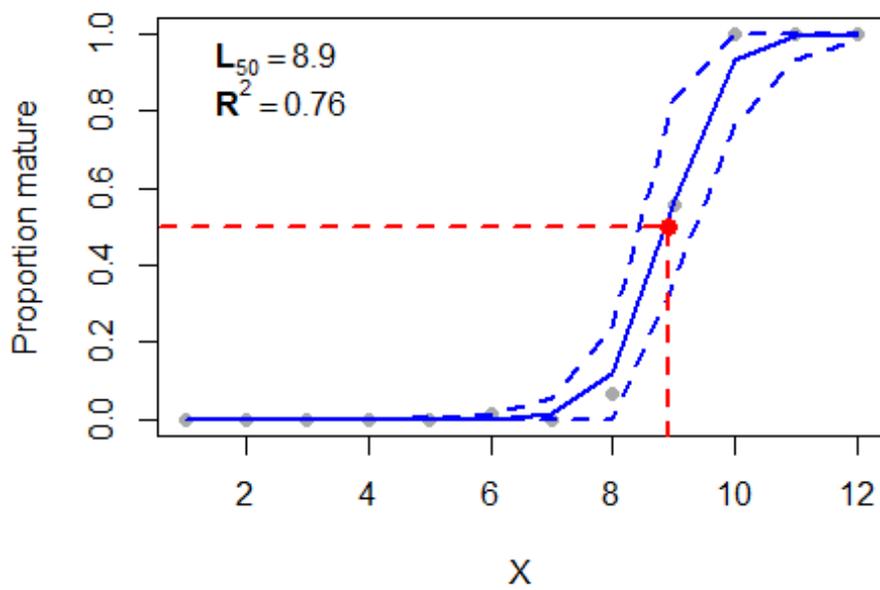
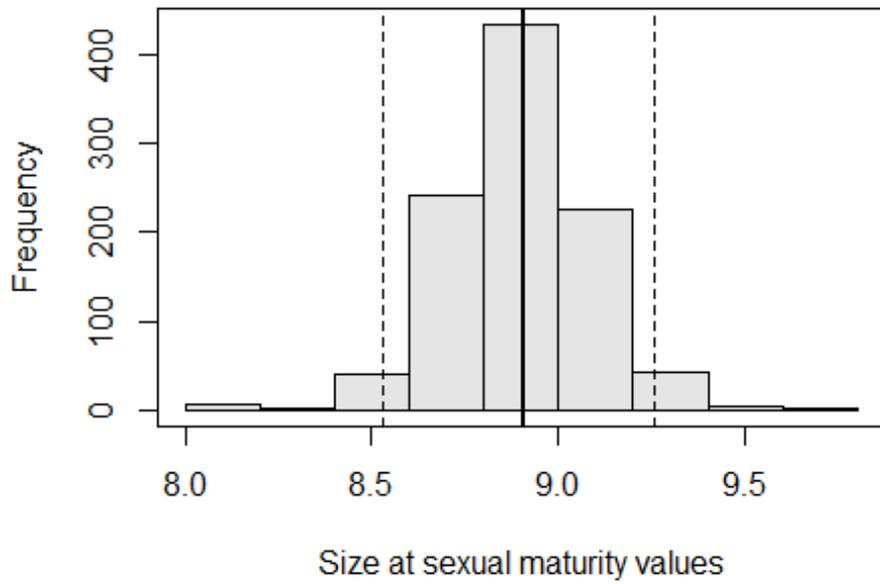
```
datmat_A <- LC.mat.check %>%  
  drop_na(Age) %>%  
  drop_na(MAT01)  
  
A_ogive_fq <- gonad_mature(datmat_A, varNames = c("Age", "MAT01"), inmName =  
0, matName = 1, method="fq", niter = 1000)  
plot(A_ogive_fq)
```



A



B



```
## Size at gonad maturity = 8.9
## Confidence intervals = 8.5 - 9.3
## Rsquare = 0.76
```

Using an alternative assignment drops both the  $L_{50}$  and  $A_{50}$  about 7%, from 128 (124 - 132 95% cl) to 119 (116 - 123) cm, and from 9.5 (9 - 10.8) to 8.9 (8.5-9.3) yr. This alternative does not seem likely but here we have accessed its effect on the estimates if it was true.

## Skipping and first-time spawners

To check for skippers, we pull the larger complete data set of females lengths and maturity and examine those fish with a thick gonad wall in the month of October, which precedes the presumed spawning period.

```
datmat_L %>%
  filter(TUNICA=='K') %>%
  filter (month == 10) %>%
  dplyr::select(HistologyID, CollectionDate, TUNICA_THICKNESS, Length,
MAOS, gsi, Age, MAT01)
```

##	HistologyID	CollectionDate	TUNICA_THICKNESS	Length	MAOS	gsi	Age
MAT01							
## 1	HHI043	2015-10-20	1000	145	V2	4.9731006	NA
1							
## 2	HHSK085	2017-10-19	904	117	V1	0.9471065	9
1							
## 3	HHSK137	2017-10-21	863	140	V2	7.9393096	9
1							
## 4	HHSK188	2018-10-04	685	137	V2	8.5139055	10
1							

Note that, in addition to a thick gonad wall (indicates past maturity) 3 of these fish have advanced yolk stage (V2) and high gsi ( $\geq 5$ ). A fourth fish, the smallest female, appears to be not spawning ready (to spawn in the imminent spawning period) but is getting ready to spawn again (in next year's spawning period). This is a suitable manner to check for skippers but the sample size was rather small.

An alternative method to find skippers at other times of year would be to check for a thick gonad wall (evidence of past spawning) but with out evidence of spawning within the year. Here, we found a likely skipper, based on its thick gonad wall, collected during July. This fish is 0.5 year from the spawning season but it does not show evidence of being spawning ready, as it has a low gsi ( $\sim 1$ ) and no evidence of vitellogenesis.

```
datmat_L %>%
  filter(TUNICA=='K') %>%
  filter (MAOS == 'LC') %>%
  dplyr::select(HistologyID, CollectionDate, TUNICA_THICKNESS, Length,
MAOS, gsi, Age, MAT01)
```

##	HistologyID	CollectionDate	TUNICA_THICKNESS	Length	MAOS	gsi	Age
MAT01							
## 1	HHSK231	2018-07-25	1233	135	LC	0.9380291	9
1							

Among fish selected here, without a seasonal restriction, this suggests skipping rate of 1 of 24 fish or 4.2% (using the number of mature females identified in a previous section as the denominator).

Some people have asked about evidence of spawning for the first time. Here is a fish that looks ready to spawn for the first time (primiparous), as evident from a low gsi (~1) but it has started vitellogenesis in April.

```
datmat_L %>%
  filter(TUNICA=='T') %>%
  filter (MAOS == 'V1') %>%
  dplyr::select(CollectionDate, TUNICA_THICKNESS, Length, MAOS, gsi, Age,
MAT01)

##   CollectionDate TUNICA_THICKNESS Length MAOS      gsi Age MAT01
## 1      2015-04-17                374   122   V1 0.9560923 NA     1
```

This is only a single, isolated example, but as it is among the same group of 24 mature females, it represents 4.2% of the mature fish.

**End**