

# nwfsc\_combo\_survey

Joshua Zahner

1/31/2023

## NWFSC Slope Survey

```
library(nwfscSurvey)

## Loading required package: dplyr
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##   filter, lag
## The following objects are masked from 'package:base':
##   intersect, setdiff, setequal, union
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.4.0      v purrr    1.0.1
## v tibble   3.1.8      v stringr  1.5.0
## v tidyr    1.3.0      vforcats  0.5.1
## v readr    2.1.2

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()

species <- "shortspine thornyhead"
survey.name <- "NWFSC.Slope"
survey.name.path <- str_replace(tolower(survey.name), "[.]", "_")

data.dir <- here::here("data/") # location of data directory
outputs.dir <- here::here("outputs/")

raw.data.dir <- file.path(data.dir, "raw")
catch.fname <- file.path(raw.data.dir, paste0(survey.name.path, "_survey_catch.csv")) # raw survey catch
bio.fname <- file.path(raw.data.dir, paste0(survey.name.path, "_survey_bio.csv"))      # raw survey bio
out.dir <- file.path(outputs.dir, "surveys", survey.name.path) # send all plot/data outputs here

dir.create(raw.data.dir, recursive=TRUE)

## Warning in dir.create(raw.data.dir, recursive = TRUE):
## '/Users/jzahner/Desktop/shortspine_thornyhead_2023/data//raw' already exists
```

```

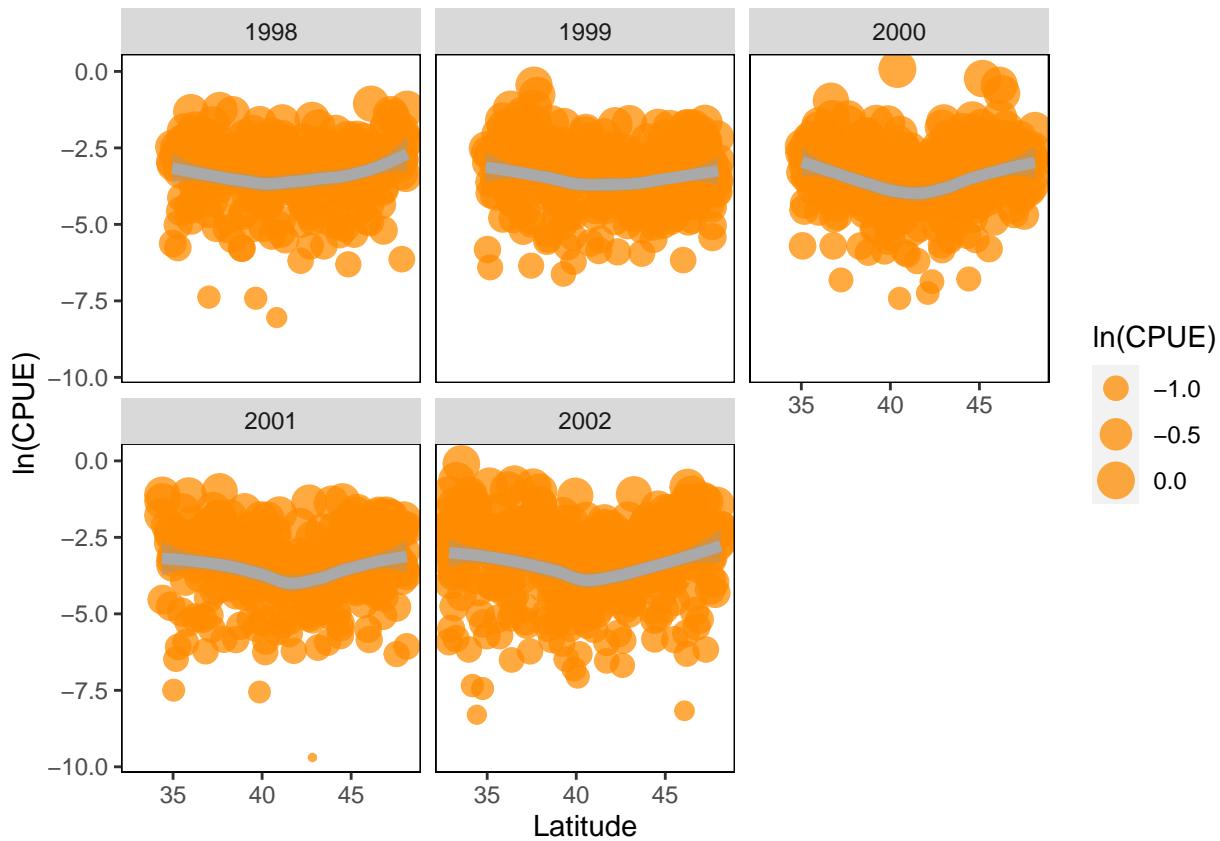
dir.create(out.dir, recursive=TRUE)

## Warning in dir.create(out.dir, recursive = TRUE):
## '/Users/jzahner/Desktop/shortspine_thornyhead_2023/outputs//surveys/nwfsc_slope'
## already exists

plot_cpue(
  dir = out.dir,
  catch = catch
)

## `geom_smooth()` using formula = 'y ~ x'

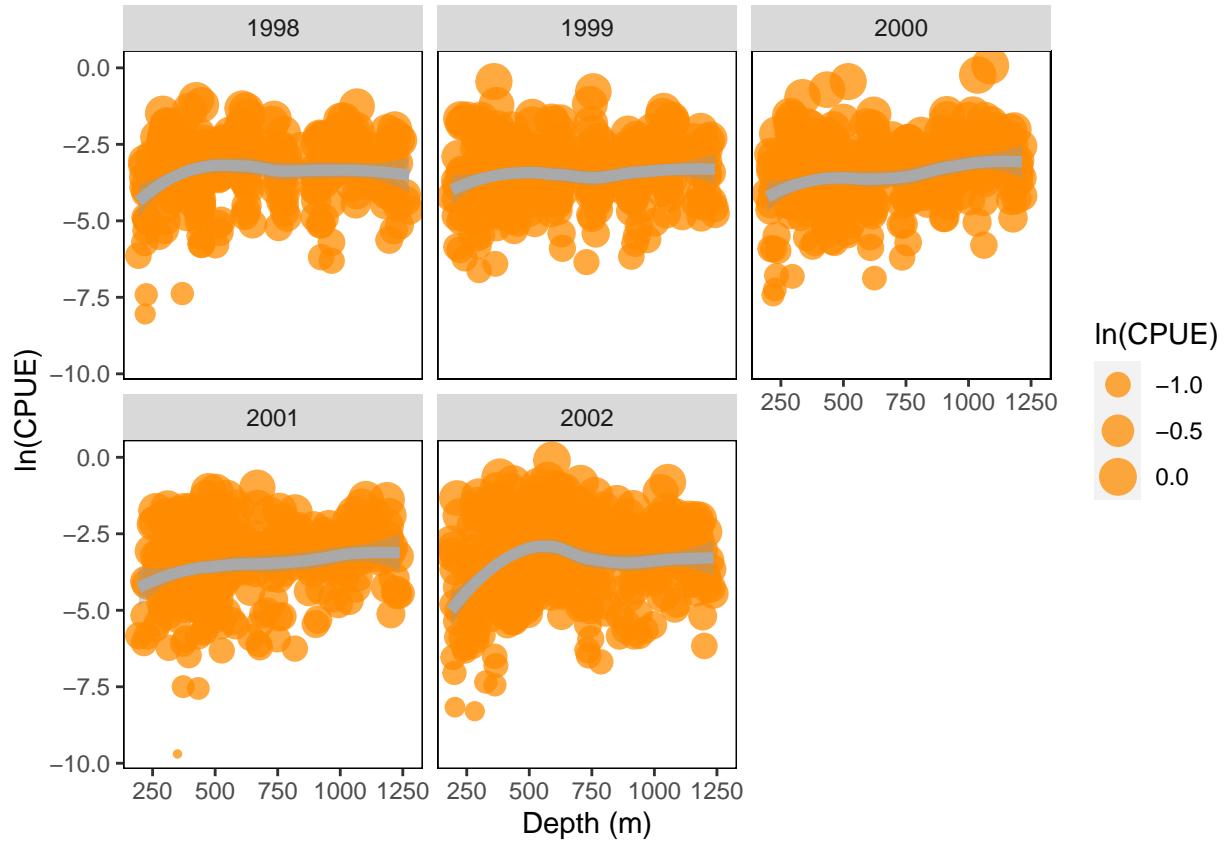
```



```

## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'

```



```
plot_bio_patterns(
```

```
  dir = out.dir,
```

```
  bio = bio,
```

```
  col_name = "Length_cm")
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_summary()`).
```

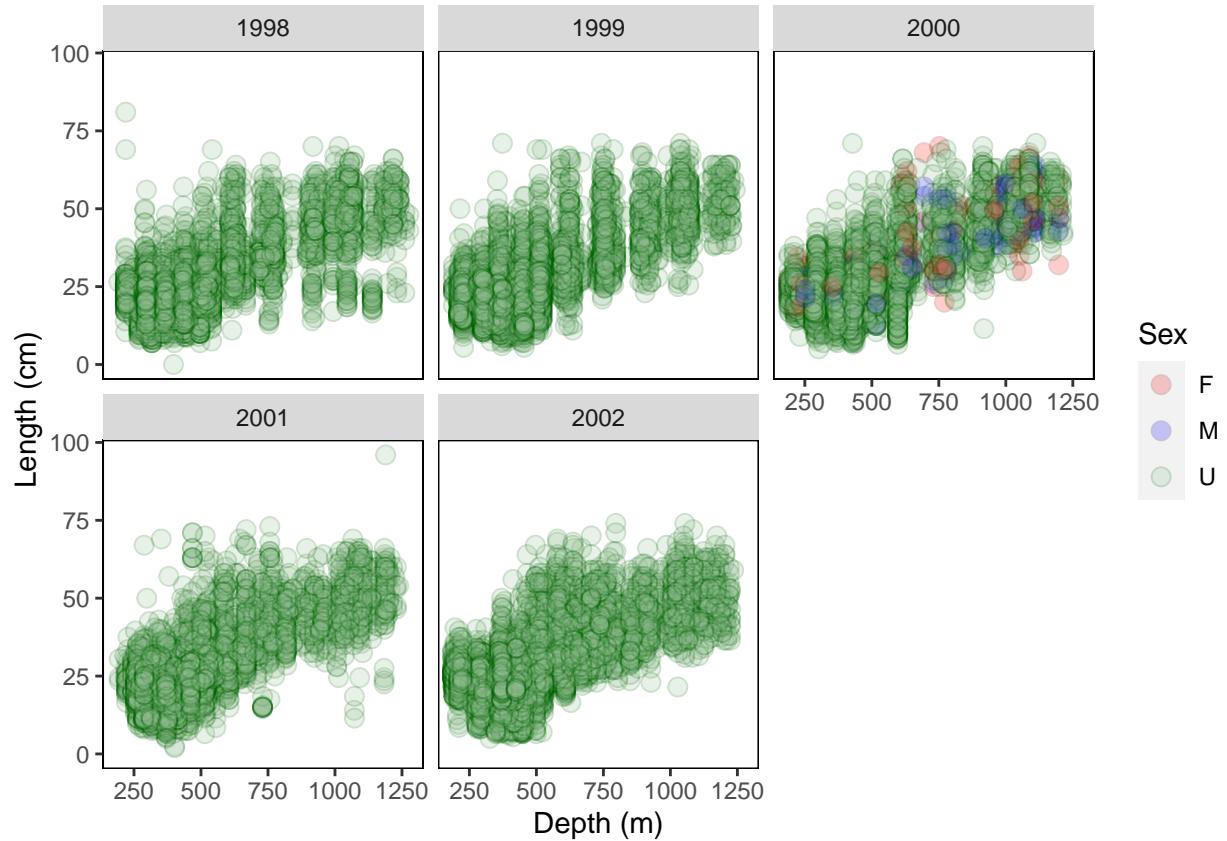
```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_summary()`).
```

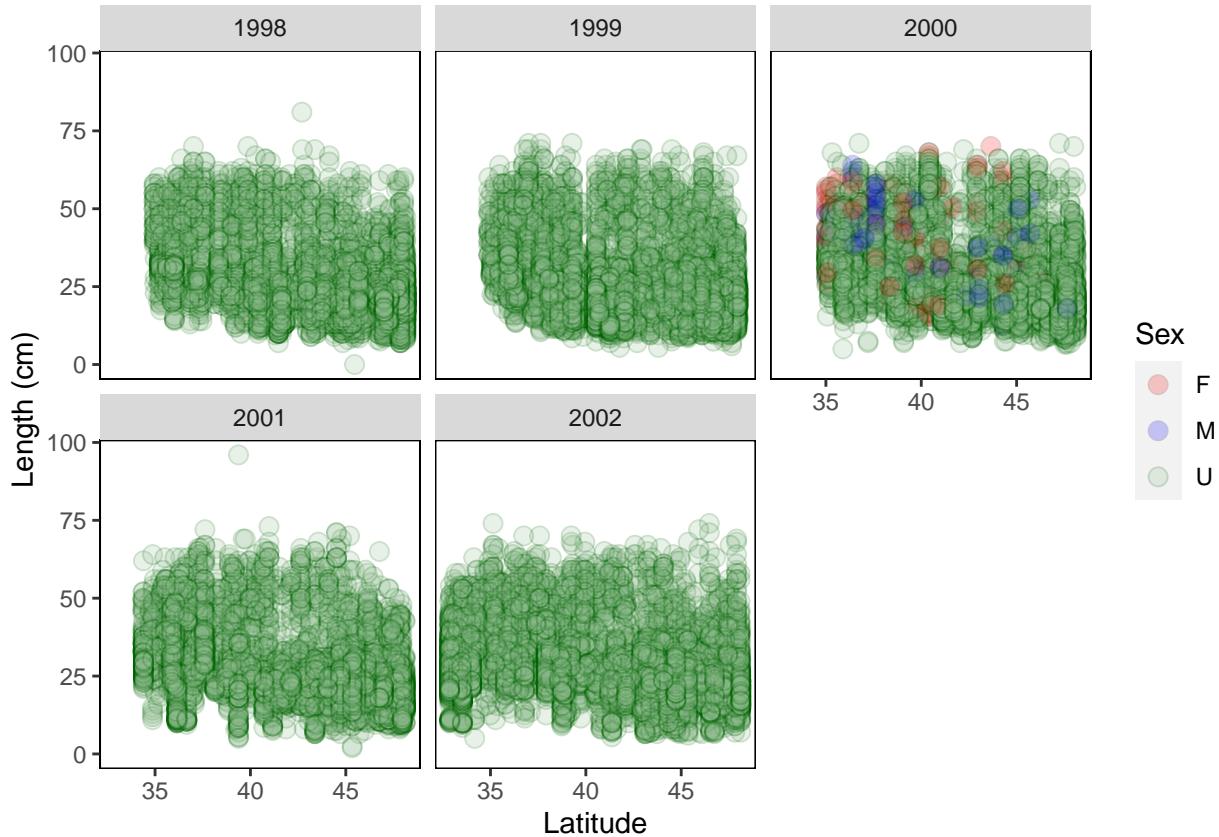
```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```

```
## Removed 2 rows containing missing values (`geom_point()`).
```

```
## Removed 2 rows containing missing values (`geom_point()`).
```



```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
## Removed 2 rows containing missing values (`geom_point()`).
```



CPUE seems to be pretty flat across latitude over time, though there is definitely a decrease in middle latitudes ( $\sim 38^\circ - 42^\circ$ ). Maybe a slight positive trend in CPUE with depth (higher CPUE as greater depth). Definitely something going on in 2002.

Clear positive trend in length with depth, and no obvious trends between length and latitude. Indicates we need to be thoughtful as to where to set depth strata, but latitudinal strata can be set pretty much anywhere.

NOTE that almost all samples in this survey are unsexed. Will need to be VERY careful how we deal with length compositions form this survey if we assume sex-specific growth parameters.

```
# Made a small change to the default wh_plot_proportion() function in nufscSurvey so
# that the plots are shown in line as well as saved to files.
wh_plot_proportion <- function (data_catch, data_bio, dir = file.path(getwd(), "plots"),
                                bar_width = c("n", "equal"))
{
  stopifnot(any(c(!missing(data_catch), !missing(data_bio))))
  files_all <- file.path(dir, t(outer(X = c(if (!missing(data_catch)) {
    "presence-absence"
  }, if (!missing(data_bio)) {
    "sex"
  }), Y = paste0("_by_", c("depth", "latitude"), ".png"), FUN = paste0)))
  data <- c(if (!missing(data_catch)) {
    dplyr::mutate(data_catch, the_factor = factor(cpue_kg_km2 <=
      0, levels = c(FALSE, TRUE), labels = c("Present",
      "Absent"))) %>% purrr::rerun(.n = 2)
  }, if (!missing(data_bio)) {
    dplyr::mutate(data_bio, the_factor = codify_sex(Sex)) %>%
      purrr::rerun(.n = 2)
  })
}
```

```

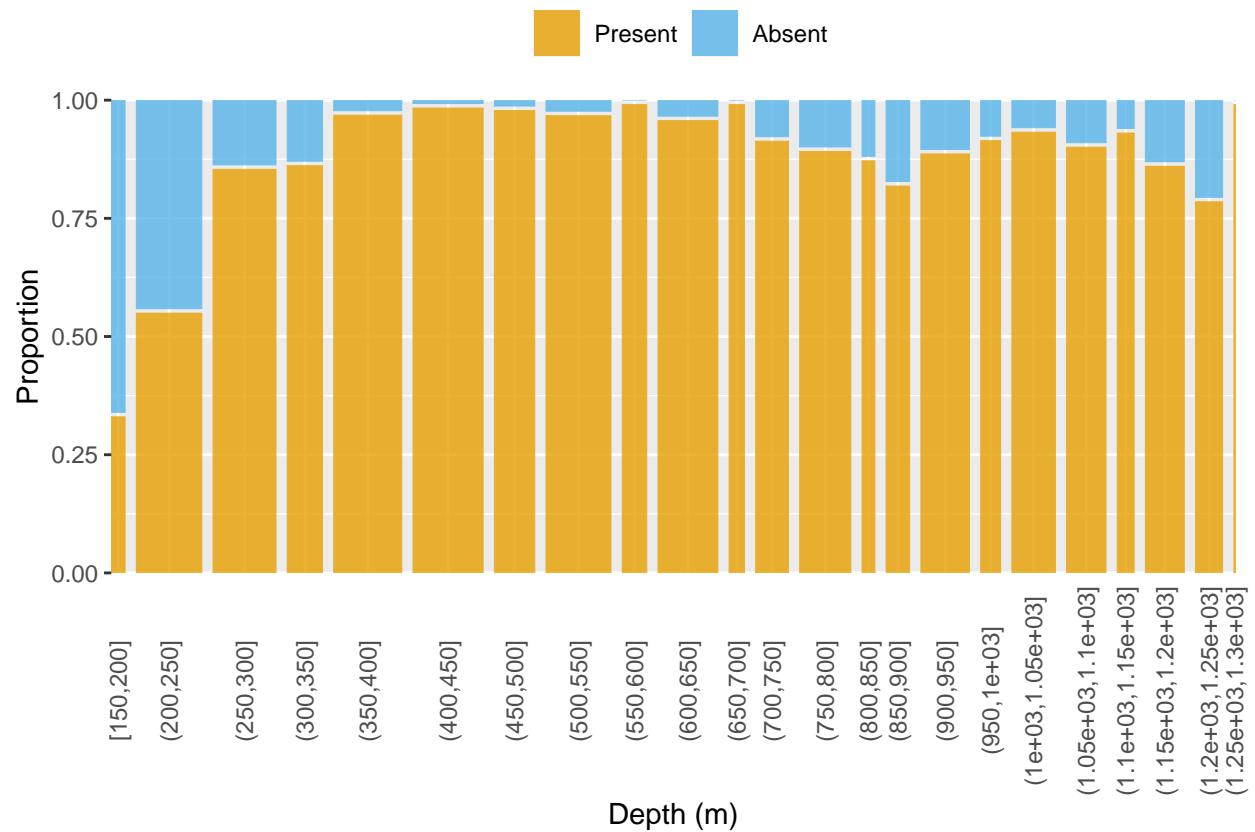
  })
gg_all <- purrr::pmap(.l = list(pdata = data, x = rep(ggplot2::quos(Depth_m,
  Latitude_dd), length(data)/2), width = rep(c(50, 1),
  length(data)/2)), .f = function(pdata, x, width, bar_width) {
  gg <- plot_proportion(data = pdata, column_factor = the_factor,
    column_bin = !!x, width = width, bar_width = bar_width,
    boundary = 0)
}, bar_width = match.arg(bar_width))
files_out <- purrr::map2_chr(.x = files_all, .y = gg_all,
  .f = ggplot2::ggsave, height = 7, width = 7)
return(list(files_out, gg_all))
}

wh_plot_proportion(
  dir = file.path(out.dir, "plots"),
  data_catch = catch,
  data_bio = bio
)

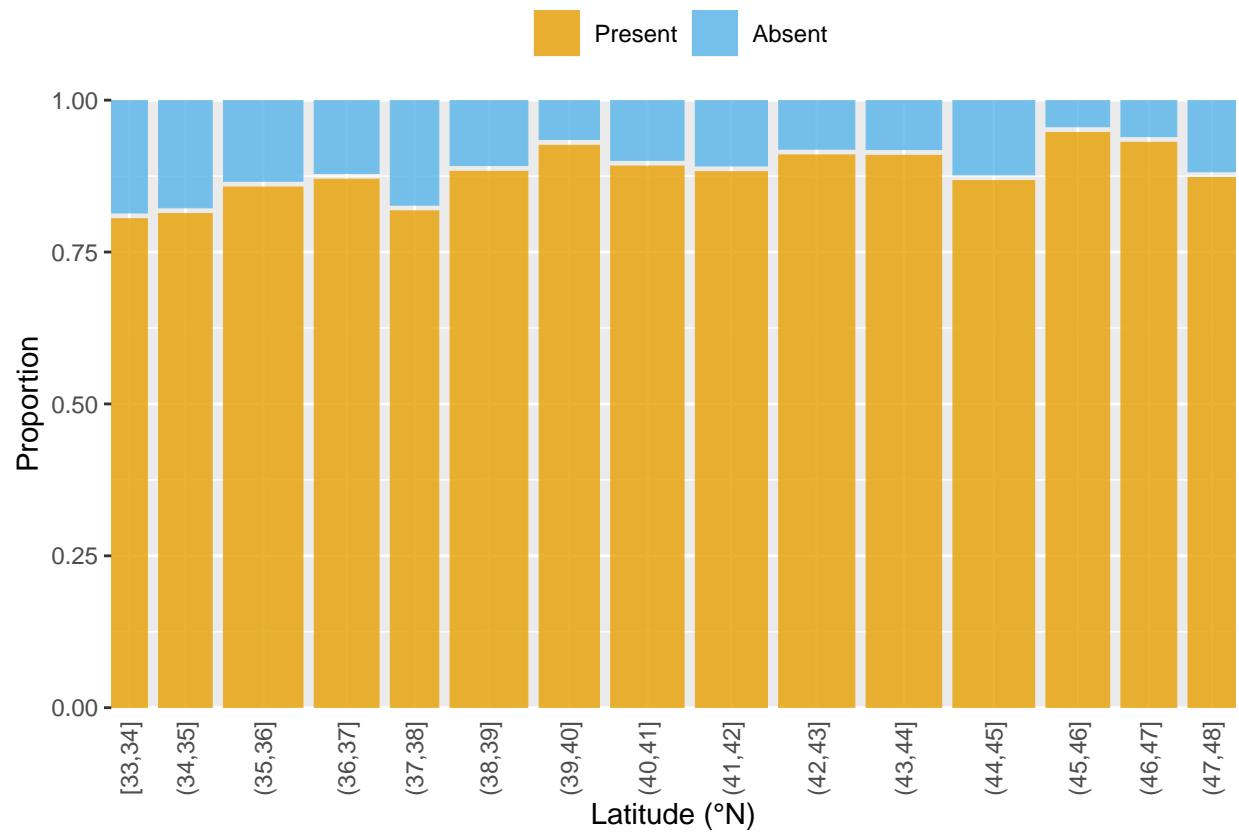
## Warning: `rerun()` was deprecated in purrr 1.0.0.
## i Please use `map()` instead.
##   # Previously
##   rerun(2, .)
##
##   # Now
##   map(1:2, ~.)

## [[1]]
## [1] "/Users/jzahner/Desktop/shortspine_thornyhead_2023/outputs//surveys/nwfsc_slope/plots/presence-absence_by_depth"
## [2] "/Users/jzahner/Desktop/shortspine_thornyhead_2023/outputs//surveys/nwfsc_slope/plots/presence-absence_by_latitude"
## [3] "/Users/jzahner/Desktop/shortspine_thornyhead_2023/outputs//surveys/nwfsc_slope/plots/sex_by_depth"
## [4] "/Users/jzahner/Desktop/shortspine_thornyhead_2023/outputs//surveys/nwfsc_slope/plots/sex_by_latitude"
##
## [[2]]
## [[2]][[1]]

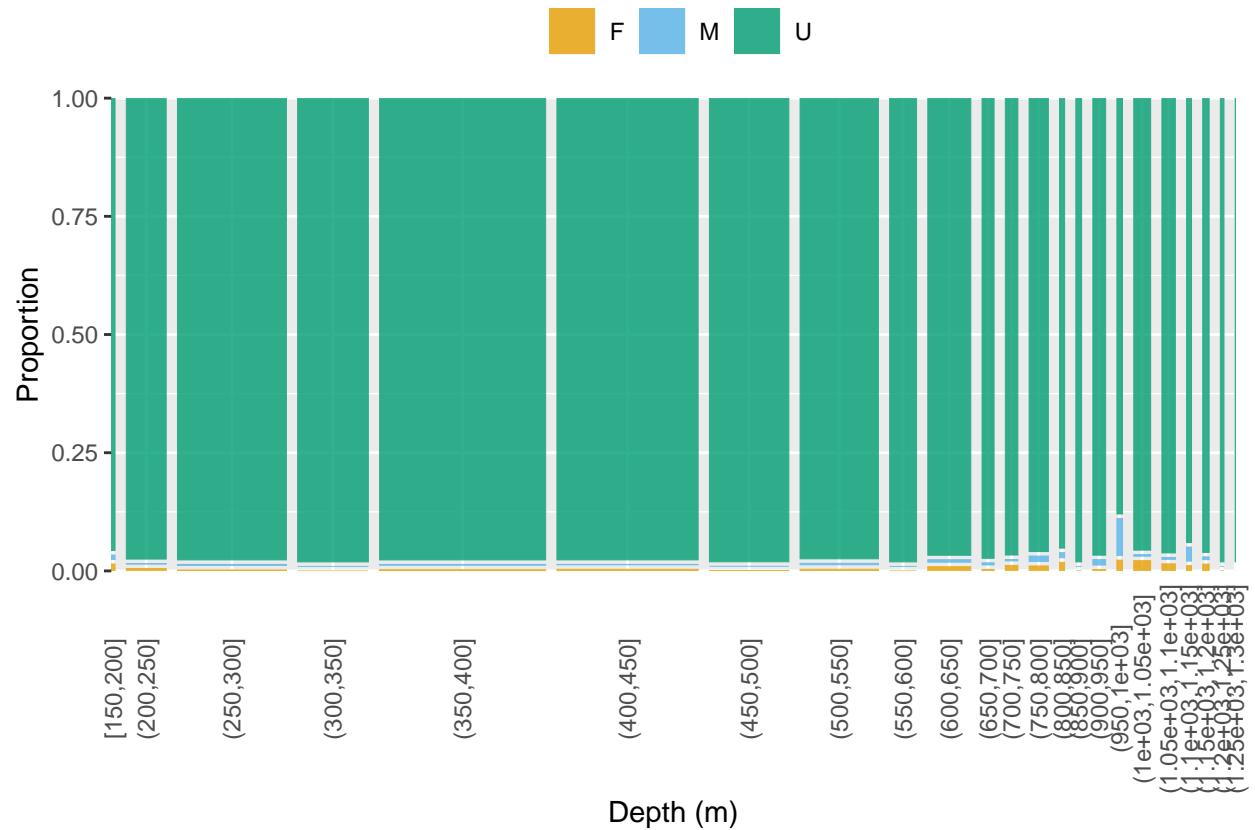
```



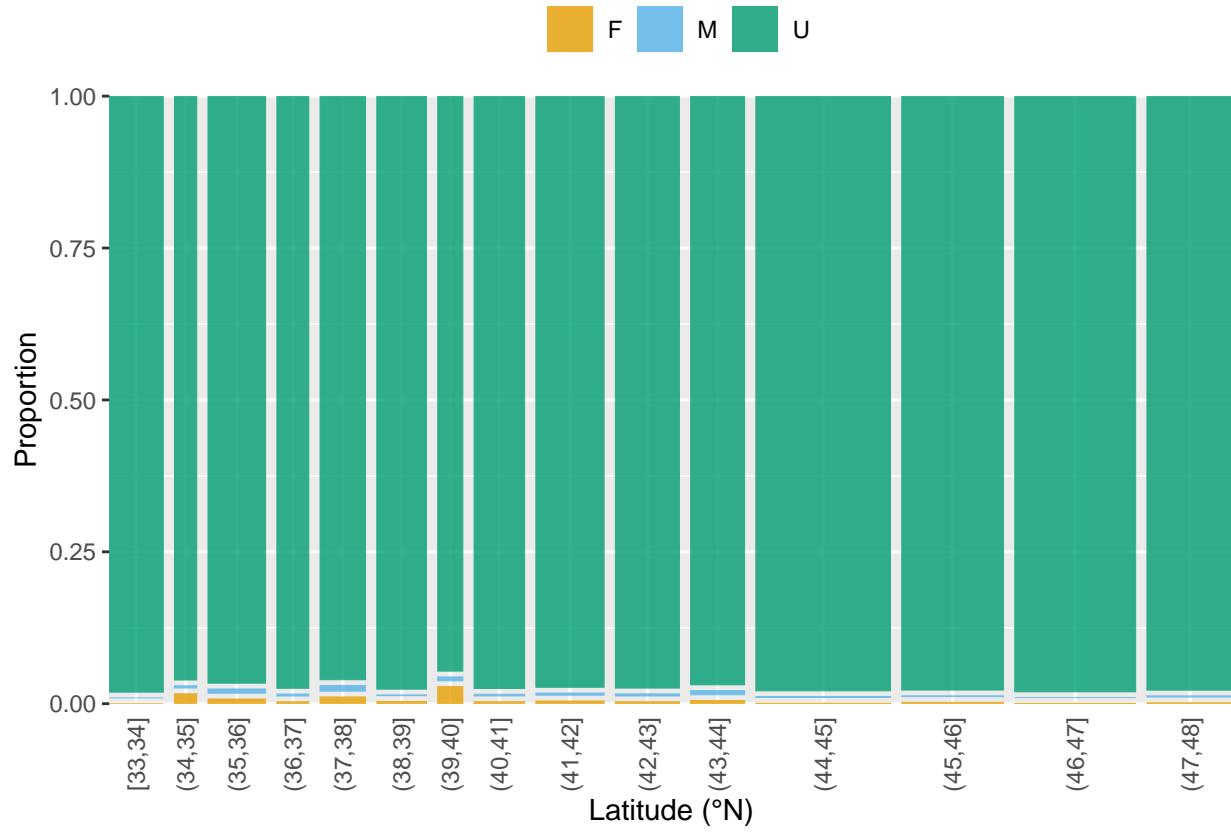
```
##  
## [[2]] [[2]]
```



```
##  
## [[2]] [[3]]
```



```
##  
## [[2]][[4]]
```



Presence-absence plots indicate the species is most prevalent in ~350-1000m of water, and is much prevalent in shallower waters. There are no obvious presence-absence trends across latitude. These both again imply that depth strata should be set thoughtfully, but that latitudinal strata can be set at convenience.

No trends in sex-specific presence-absence across depth or latitude is discernible due to the large proportion of unsexed fish.

```
# Set spatial strata
strata = CreateStrataDF.fn(
  names      = c("shallow_south", "deep_south", "shallow_cen", "deep_cen", "shallow_north", "deep_north"),
  depths.shallow = c(55,      500,      55,      500,      55,      500),
  depths.deep   = c(500,     1280,     500,     1280,     500,     1280),
  lats.south    = c(32,      32,      40.5,    40.5,    43,      43),
  lats.north    = c(40.5,    40.5,    43,      43,      49,      49)
)
```

Strata are set to the same as they were in 2014 (p. 27). I dont see any particularly strong reason to change the latitudinal bins, since CPUE and presence-absence are very consistent over latitude. Only using two depth bins seems like it could be revised though. Maybe use 55-350 m, 350-750m, 750+ m?

## Design Based Indices

```
## Calculate design-based index

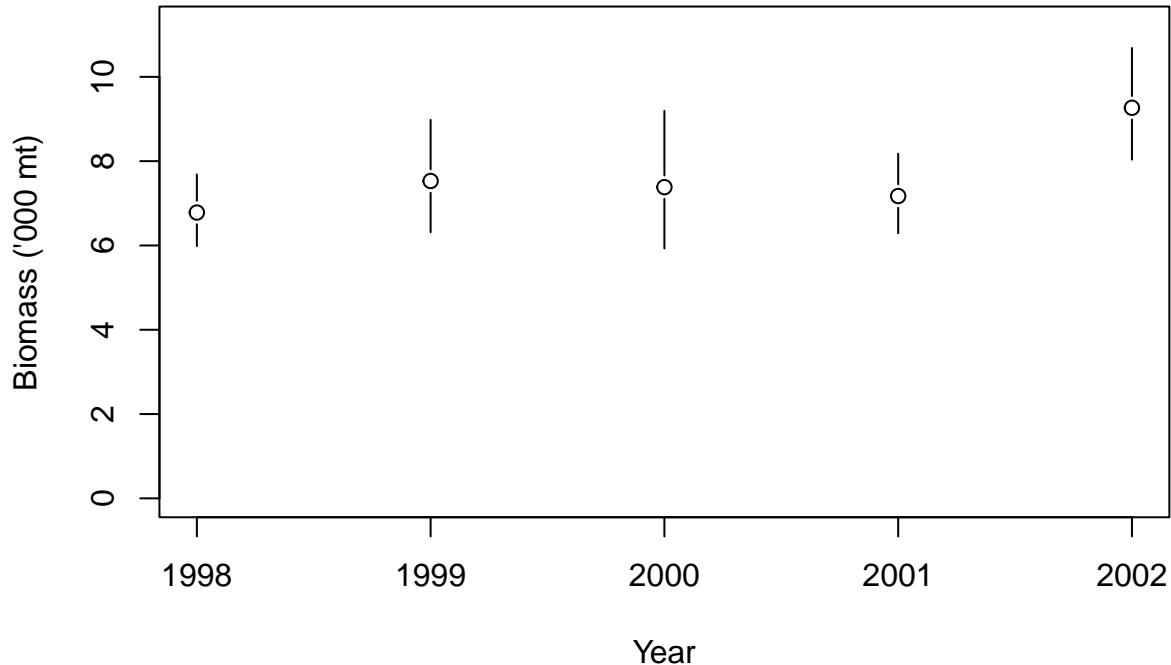
biomass.idx = Biomass.fn(dir = out.dir,
                         dat = catch,
                         strat.df = strata)
```

```

PlotBio.fn(
  dir = NULL, # set to `out.dir`to save as file
  dat = biomass.idx,
  scalar = 1000, # convert to kg like csv file
) # produces '_designed_based_index.png'

## Output will not be saved in `dir` because dir = NULL.

```

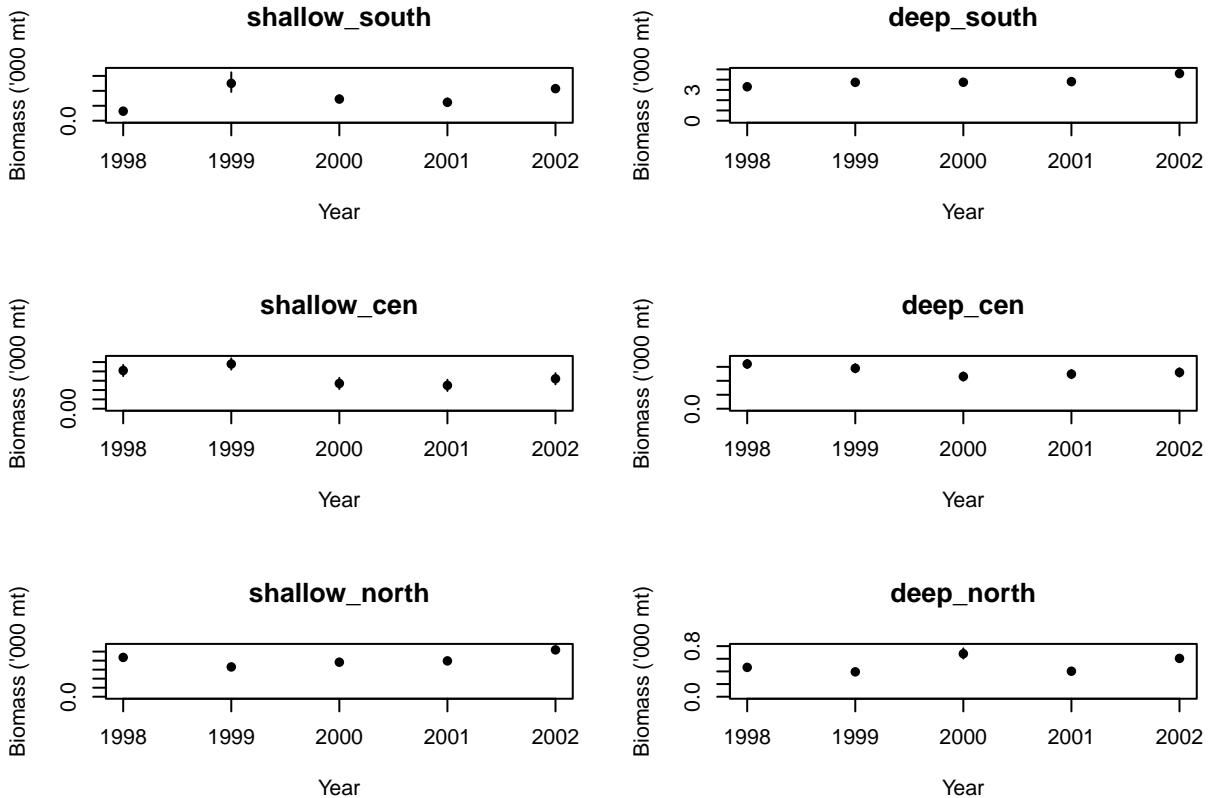


```

PlotBioStrata.fn(
  dir = NULL, # set to `out.dir`to save as file
  dat = biomass.idx,
  scalar = 1000, # convert to kg like csv file
) # produces '_designed_based_by_strata_index.png'

## Output will not be saved in `dir` because dir = NULL.

```



## Length Compositions

Lengths only available for the year 2000. Male length comps look generally consistent with other surveys, though female comps look less similar. Could be due to low sample sizes or lack of sexed data from this survey. Length comps from this survey may not be of sufficient quality to warrant use in the assessment.

```
n.len.raw <- nrow(bio[!is.na(bio$Length_cm),]) # check if data available

if(n.len.raw > 0){
  n <- GetN.fn(dir = out.dir,
    dat = bio,
    type = "length",
    species = "thorny")

  len_bins <- seq(6, 72, 2) # from last assessment (p. 125)

  Length_Freq <- SurveyLFs.fn(dir = out.dir,
    datL = bio,
    datTows = catch,
    strat.df = strata,
    lgthBins = len_bins)

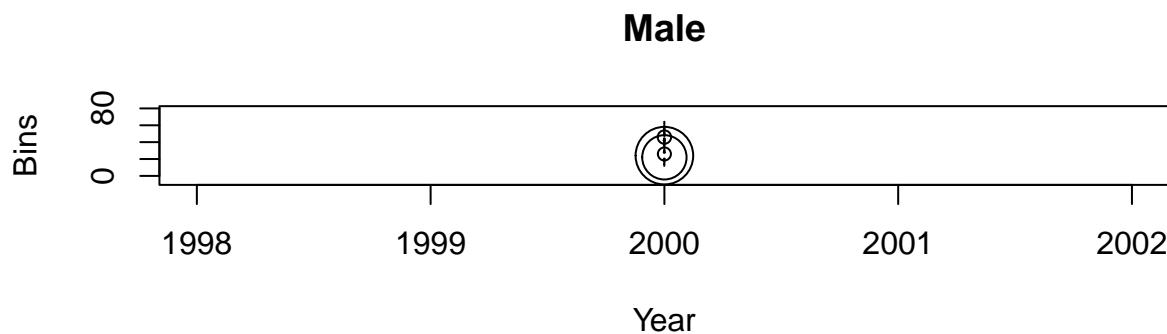
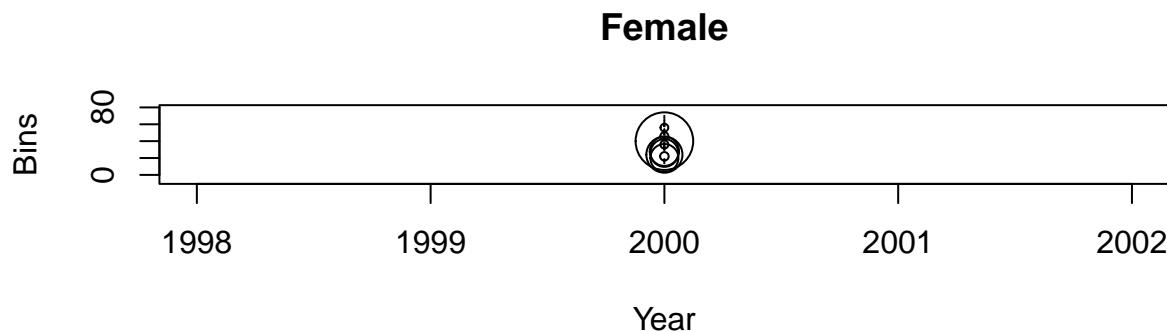
  PlotFreqData.fn(dir = NULL, # set to `out.dir` to save as file
    dat = Length_Freq)
```

```

    plot_comps(dir = out.dir, data = Length_Freq) # uses version in R/utils/
} else{
  print(n.len.raw)
}

##
## The effN sample size is calculated using the thorny multiplier of 6.91. This number is multiplied by
##
## There are 8 tows where fish were observed but no lengths/ages taken.
## These tows contain 53 lengths/ages that comprise 0.1 percent of total sampled fish.
## There are 46988 records kept out of 46990 records after removing missing records.
##
## NOTE: Files have been saved the the printfolder directory.
## The first file has the 999 column showing fish smaller or younger than the initial bin.
## Check to make sure there is not a large number of fish smaller or younger than the initial bin.
## The second file has combined the 999 with the first bin and is ready for use in SS.
##
## Output will not be saved in `dir` because dir = NULL.

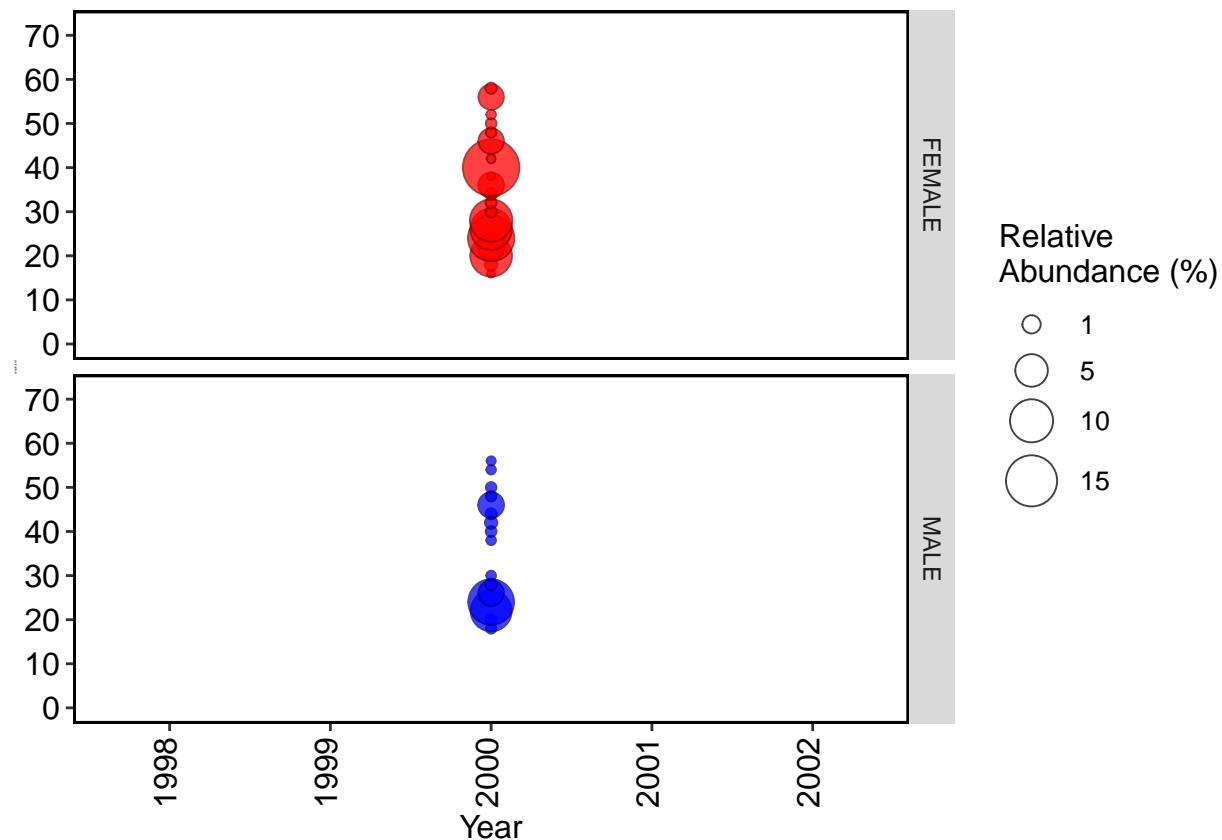
```

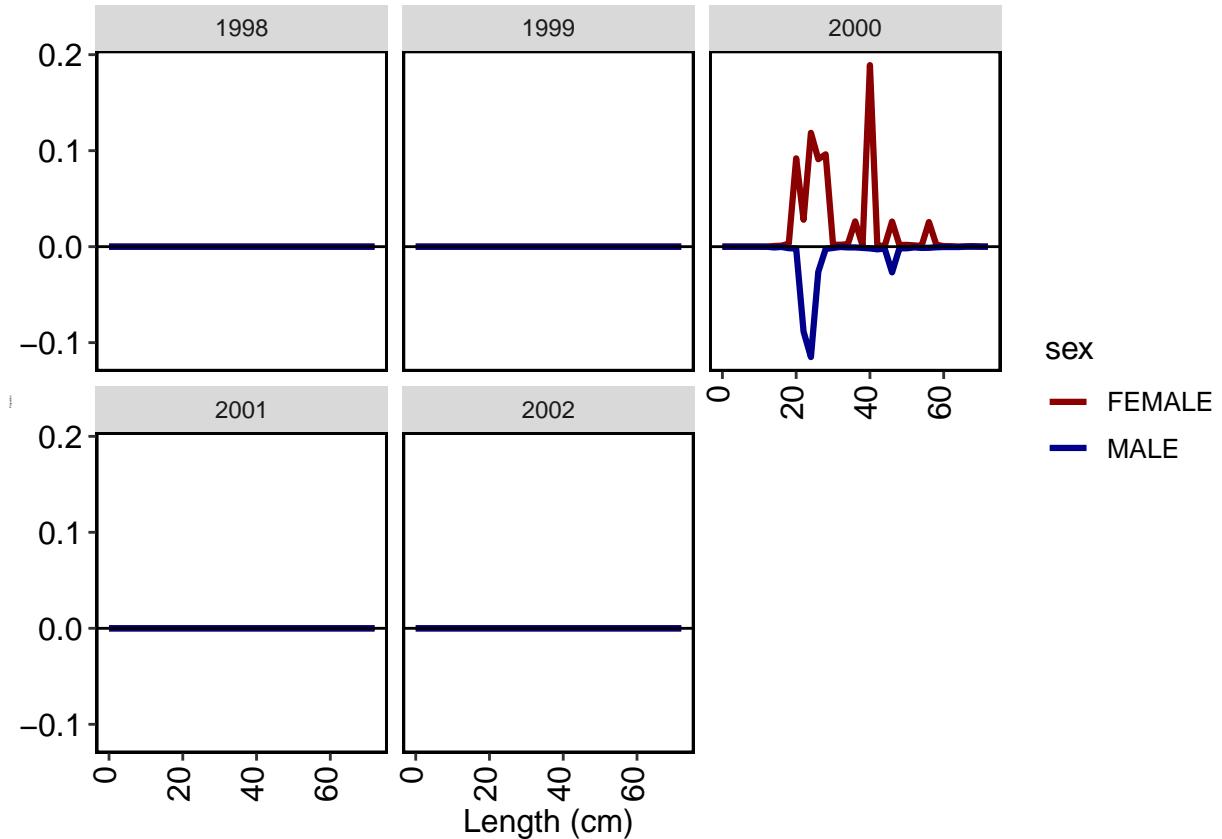


```

## Warning: Removed 314 rows containing missing values (`geom_point()`).
## Removed 314 rows containing missing values (`geom_point()`).

```





## Age Compositions

Small samples sizes available for ages in 2000. Quality of data is unknown, but is likely suffers from similar sample size and sexing issues as the length composition data.

```
n.age.raw <- nrow(bio[!is.na(bio$Age),]) # check is data available

if(n.age.raw > 0){

  n <- GetN.fn(dir = out.dir,
    dat = bio,
    type = "age",
    species = "thorny")

  age_bins <- 1:22 # arbitrarily set right now as not used in previous assessment

  Ages <- SurveyAFs.fn(dir = out.dir,
    datA = bio,
    datTows = catch,
    strat.df = strata,
    ageBins = age_bins,
    nSamps = n)

  PlotFreqData.fn(
    dir = out.dir,
    dat = Ages
```

```

)
plot_comps(
  dir = out.dir,
  data = Ages
) # uses the version in R/utils/

} else{
  print(n.age.raw)
}

##  

## The effN sample size is calculated using the thorny multiplier of 6.91. This number is multiplied by  

##  

## Input sample size exceeded the number of fish for 1999 2000 and has been capped equal to number of fish.  

##  

## There are 8 tows where fish were observed but no lengths/ages taken.  

##           These tows contain 53 lengths/ages that comprise 0.1 percent of total sampled fish.  

## There are 820 records kept out of 46990 records after removing missing records.  

##  

## NOTE: Files have been saved the the printfolder directory.  

##           The first file has the 999 column showing fish smaller or younger than the initial bin.  

##           Check to make sure there is not a large number of fish smaller or younger than the initial bin.  

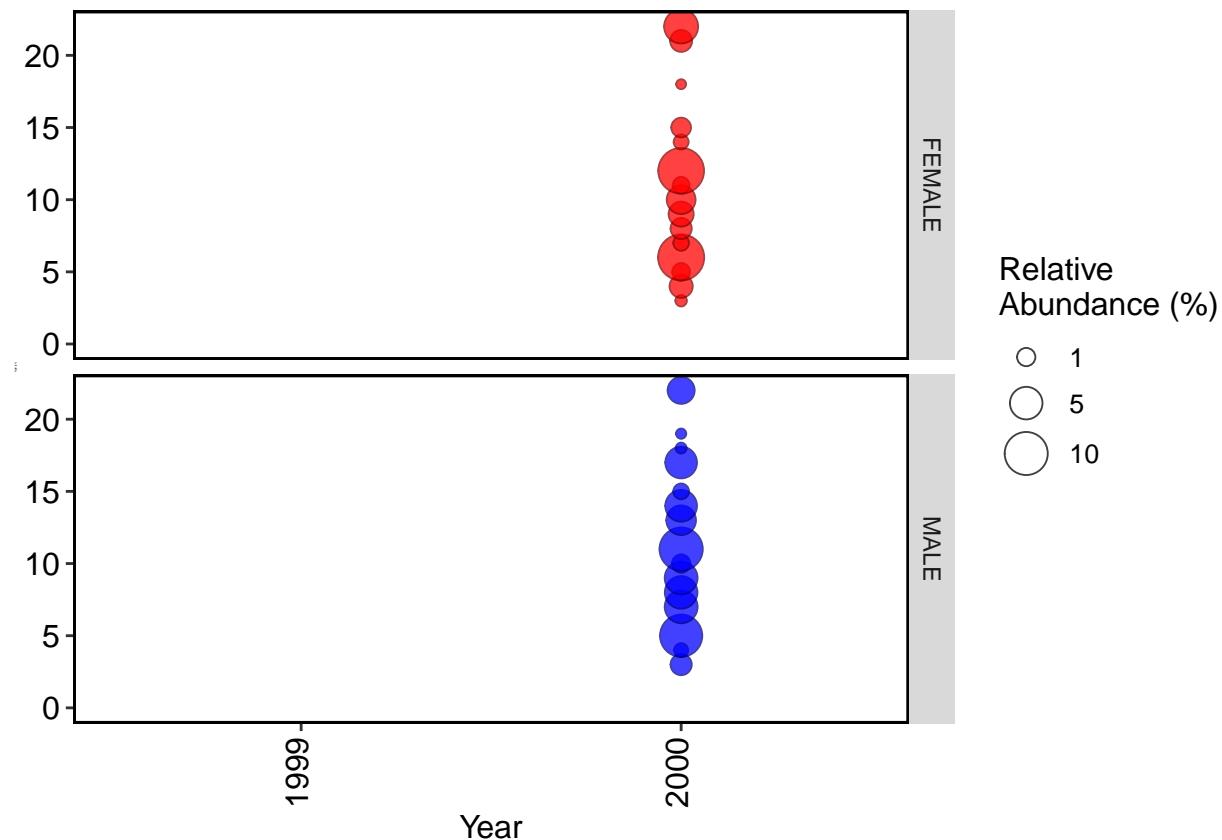
##           The second file has combined the 999 with the first bin and is ready for use in SS.  

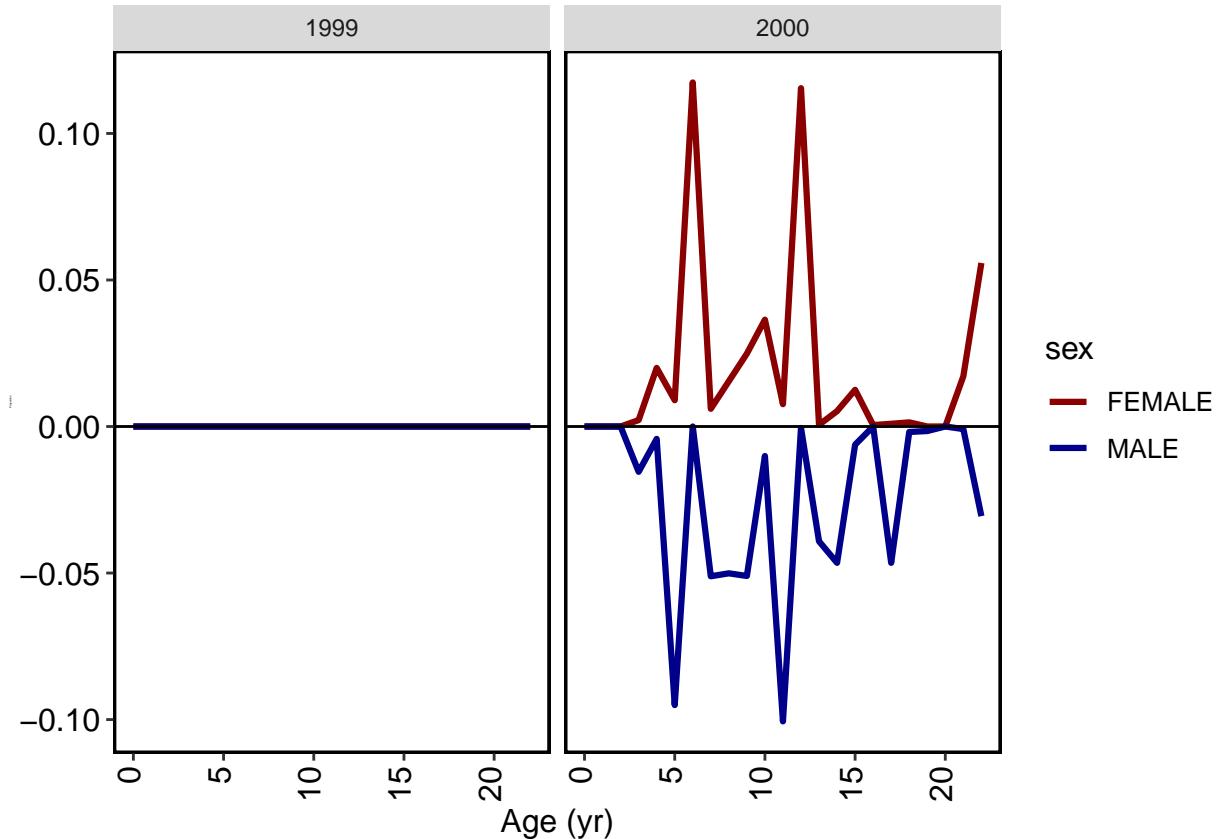
##  

## Warning: Removed 62 rows containing missing values (`geom_point()`).  

## Removed 62 rows containing missing values (`geom_point()`).

```





### Conditional Age-at-Length

```

caal <- SurveyAgeAtLen.fn(dir = out.dir,
                           datAL = bio,
                           datTows = catch,
                           strat.df = strata,
                           lgthBins = len_bins,
                           ageBins = age_bins)

## There are 820 records kept out of 46990 records after removing missing records.
##
## Effective sample size is based on number of fish.
##
##
## Using raw numbers of age-at-length
##
## There are 0 females age 0 to age 1 minus group that were added into the first age bin
## There are 0 males age 0 to age 1 minus group that were added into the first age bin
plot_age_length_sampling(dir=out.dir, caal)

## pdf
## 2

```