SURE Project

August 25, 2022

Objective

Validate predictions using LANDSAT data for kelp area.

```
# set a directory
w.dir <- here()
d.dir <- here('data')

# load kelp area csv
area_df <- read.csv(paste(d.dir, 'NC_Landsat_kelp_area_1984_2021.csv', sep ='/')) %>%
  glimpse()
```

```
## Rows: 5,748,336
## Columns: 6
## $ lat          <dbl> 42.00981, 42.00981, 42.00981, 42.00981, 42.00981, 42.0098~
## $ lon          <dbl> -124.2244, -124.2244, -124.2244, -124.2244, -124.2244, -1~
## $ year_quarter <chr> "1984 1", "1984 2", "1984 3", "1984 4", "1985 1", "1985 2~
## $ area         <int> NA, 0, 105, 0, 0, 0, NA, 0, NA, 0, 0, 0, 0, NA, 0, NA, NA~
## $ year         <int> 1984, 1984, 1984, 1984, 1985, 1985, 1985, 1985, 1986, 198~
## $ quarter      <int> 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, ~
```

```
# vector of year
area_df$year %>% unique() # 1984 - 2021
```

```
##  [1] 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998
## [16] 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
## [31] 2014 2015 2016 2017 2018 2019 2020 2021
```

```
area_df$year %>% unique() %>% length() # 38
```

```
## [1] 38
```

```
# Landsat kelp quarters: Q1 = Winter, Q2 = Spring, Q3 = Summer, Q4 = Fall
# make an sf object from dataframe
kelp_area <- area_df %>%
  filter(quarter == 3) %>%
  mutate_at(vars(year, quarter), list(as.factor)) %>%
  st_as_sf(coords = c('lon', 'lat'))

# load kelp density predictions csv
density_df <- read.csv(paste(d.dir, 'NC_kelp_density_predictions.csv', sep ='/')) %>%
  glimpse()
```

```
## Rows: 400
## Columns: 5
## $ site_name <chr> "Caspar", "Caspar North", "Dark Gulch", "Flat Iron Rock", "F~
## $ year      <int> 2006, 2006, 2006, 2006, 2006, 2006, 2006, 2006, 2006, 2006, ~
## $ fit       <dbl> 0.5090384, 0.5003366, 0.5309656, 0.7382968, NA, 0.8228083, N~
## $ longitude <dbl> -123.8220, -123.8213, -123.7762, -124.1578, -123.2450, -123.~
## $ latitude  <dbl> 39.36173, 39.36443, 39.24030, 41.05942, 38.51060, 39.35503, ~
```

```r
# make an sf object from dataframe
kelp_density <- density_df %>%
  mutate_at(vars(year), list(as.factor)) %>%
  rename(lon = longitude, lat = latitude) %>%
  st_as_sf(coords = c('lon', 'lat'))
```

To compute a distance matrix between each site location and the kelp area for each summer, and find the closest landsat pixel to each site.

```r
# calculate the distance and match predicted kelp density to Landsat kelp area
site_name <- kelp_density[, c('site_name', 'geometry')] %>% unique()
area_1984 <- filter(kelp_area, year == 1984) # 37818
# distance matrix for 1984
mx <- st_distance(st_sfc(area_1984$geometry), st_sfc(site_name$geometry))
colnames(mx) <- site_name$site_name
head(mx)
```

```
##        Caspar Caspar North Dark Gulch Flat Iron Rock Fort Ross Frolic Cove
## [1,] 2.678480     2.675913   2.805551      0.9527113  3.633694    2.684829
## [2,] 2.678531     2.675964   2.805605      0.9527328  3.633787    2.684879
## [3,] 2.678581     2.676015   2.805659      0.9527545  3.633881    2.684929
## [4,] 2.678632     2.676066   2.805714      0.9527763  3.633975    2.684980
## [5,] 2.678212     2.675645   2.805284      0.9524415  3.633432    2.684561
## [6,] 2.677894     2.675327   2.804962      0.9521502  3.633077    2.684243
##      Gerstle Cove Glass Beach MacKerricher North Mendocino Headlands Monument
## [1,]     3.557617    2.590761           2.552208            2.735943 3.120724
## [2,]     3.557704    2.590815           2.552264            2.735994 3.120777
## [3,]     3.557792    2.590868           2.552320            2.736045 3.120830
## [4,]     3.557879    2.590922           2.552377            2.736096 3.120883
## [5,]     3.557354    2.590494           2.551941            2.735675 3.120456
## [6,]     3.557004    2.590173           2.551617            2.735357 3.120136
##      Ocean Cove Pebble Beach Point Arena Lighthouse Point Arena MPA (M2)
## [1,]   3.574772     3.409586               3.096108             3.102896
## [2,]   3.574862     3.409666               3.096160             3.102949
## [3,]   3.574951     3.409745               3.096213             3.103002
## [4,]   3.575040     3.409824               3.096265             3.103054
## [5,]   3.574510     3.409322               3.095840             3.102629
## [6,]   3.574158     3.408979               3.095520             3.102308
##      Point Arena Ref Portuguese Beach Pyramid Point Russian Gulch Salt Point
## [1,]        3.142715         2.739189    0.01674596      2.713811   3.559378
## [2,]        3.142770         2.739241    0.01690645      2.713863   3.559466
## [3,]        3.142824         2.739293    0.01707312      2.713915   3.559553
## [4,]        3.142879         2.739345    0.01724579      2.713966   3.559641
## [5,]        3.142448         2.738921    0.01650201      2.713544   3.559116
## [6,]        3.142126         2.738602    0.01610019      2.713224   3.558766
```

```
##      Stillwater Sonoma Stornetta Timber Cove  Trinidad Van Damme
## [1,]          3.593450  3.111776    3.605898 0.9585312  2.771222
## [2,]          3.593541  3.111829    3.605990 0.9585594  2.771274
## [3,]          3.593632  3.111883    3.606081 0.9585876  2.771327
## [4,]          3.593722  3.111936    3.606173 0.9586160  2.771379
## [5,]          3.593188  3.111508    3.605636 0.9582617  2.770955
## [6,]          3.592836  3.111187    3.605283 0.9579642  2.770635
```

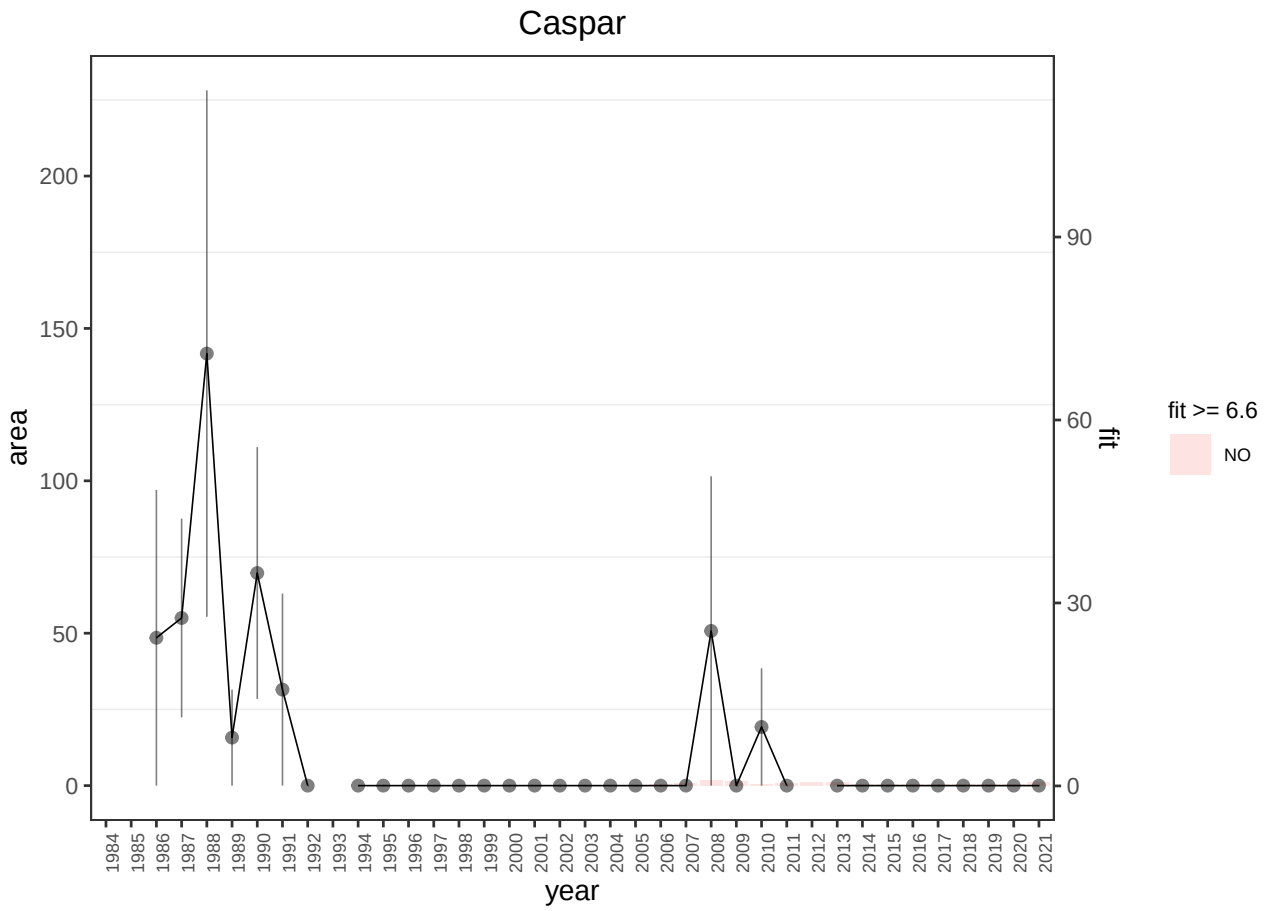
```
area_1984 <- area_df %>%
  filter(quarter == 3 & year == 1984) %>%
  select(-year_quarter) %>%
  mutate(
    site_name = apply(mx, 1, FUN = function(x) names(x) [which.min(x)]),
    .before = lat
  )
  # group_by(site_name) %>%
  # summarise(mean_area = mean(area, na.rm = TRUE),
  #           se_area = std.error(area, na.rm = TRUE))
head(area_1984)
```

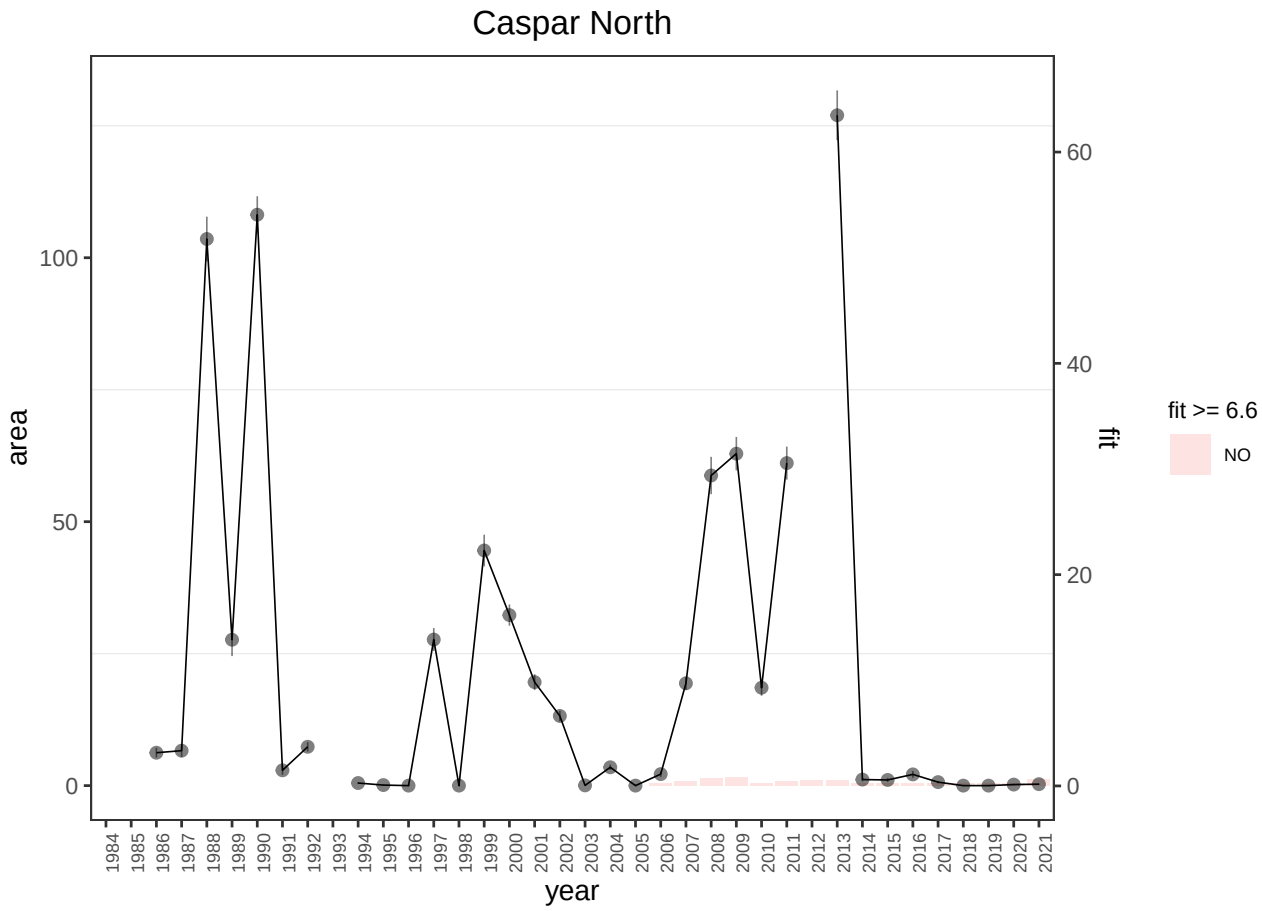
```
##       site_name      lat       lon area year quarter
## 1 Pyramid Point 42.00981 -124.2244  105 1984       3
## 2 Pyramid Point 42.00980 -124.2248    0 1984       3
## 3 Pyramid Point 42.00980 -124.2252    0 1984       3
## 4 Pyramid Point 42.00979 -124.2255    0 1984       3
## 5 Pyramid Point 42.00954 -124.2244    0 1984       3
## 6 Pyramid Point 42.00927 -124.2241   93 1984       3
```

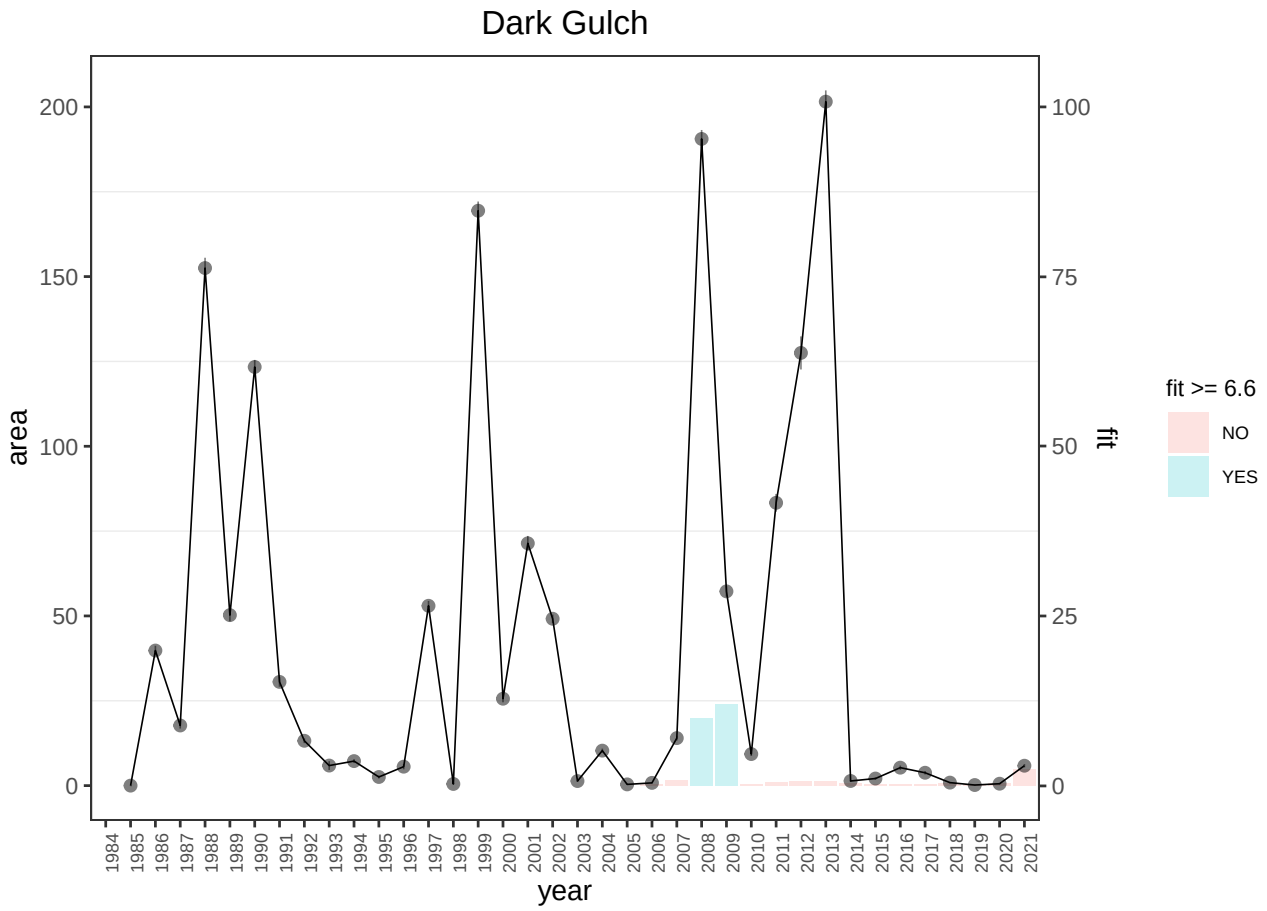
```
# using for loop
# declaring an empty data frame
# kelp_area2 <- data.frame(site_name = character(),
#                          lat = numeric(),
#                          lon = numeric(),
#                          area = numeric(),
#                          year = numeric(),
#                          quarter = numeric())
#
# for (i in c(1984:2021)) {
#   area_year <- filter(kelp_area, year == i)
#
#   # distance matrix for year i
#   mx <- st_distance(st_sfc(area_year$geometry), st_sfc(site_name$geometry))
#   colnames(mx) <- site_name$site_name
#
#   area_year <- area_df %>%
#     filter(quarter == 3 & year == i) %>%
#     select(-year_quarter) %>%
#     mutate(
#       site_name = apply(mx, 1, FUN = function(x) names(x) [which.min(x)]),
#       .before = lat
#     )
#
#   kelp_area2 <- rbind(kelp_area2, area_year)
```

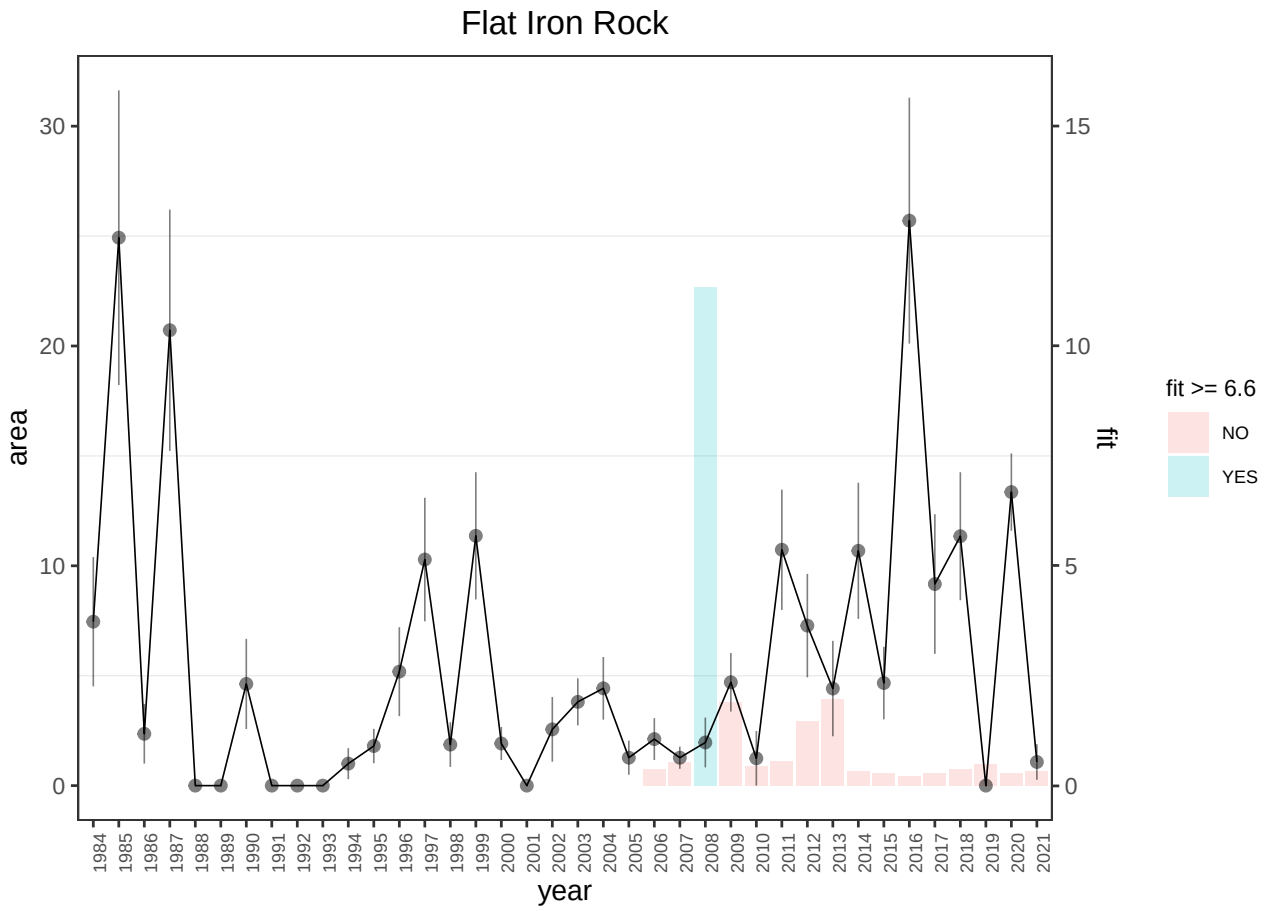
```
# }
# dim(kelp_area2) # 38 * 37818 = 1437084
#
# head(kelp_area2)
#
# write to file
# write.csv(kelp_area2, file.path(d.dir, 'NC_kelp_area_1984_2021_site_name.csv'), row.names = FALSE)
```

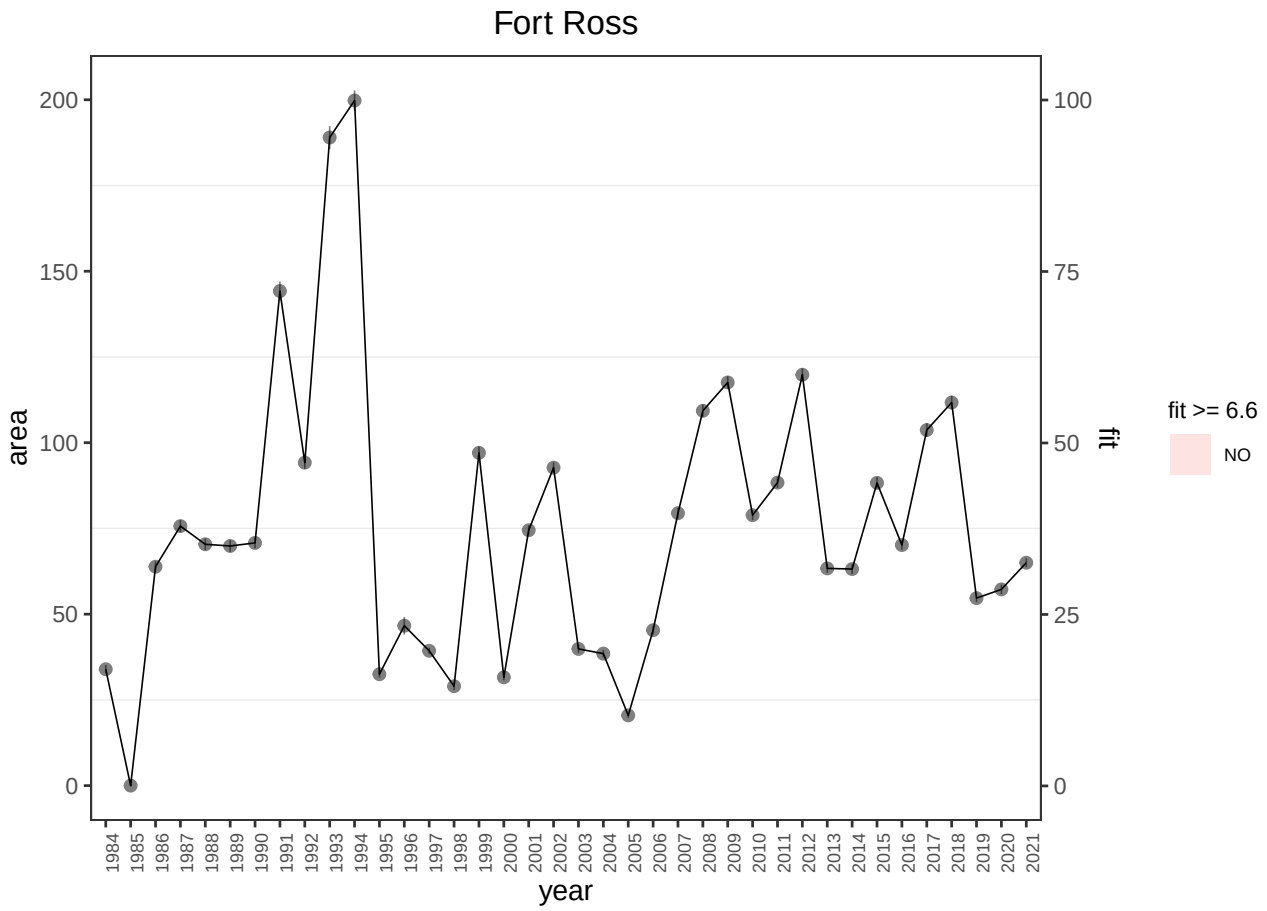
```
# plotting
kelp_area <- read.csv(paste(d.dir, 'NC_kelp_area_1984_2021_site_name.csv', sep ='/')) %>%
  mutate_at(vars(year), list(as.factor))

kelp_area3 <- kelp_area %>%
  group_by(site_name, year) %>%
  summarise(mean_area = mean(area, na.rm = TRUE),
            se_area = std.error(area, na.rm = TRUE))

sites <- kelp_area3$site_name %>% unique()

for (i in sites) {
  plot <- ggplot() +
    geom_bar(data = filter(kelp_density, site_name == i),
             aes(x = year, y = fit,
                 fill = ifelse(!is.na(fit) & fit >= 6.6, 'YES', 'NO')),
             stat = 'identity', position = 'dodge', alpha = 0.2) +
    geom_pointrange(data = filter(kelp_area3, site_name == i),
                    aes(x = year, y = mean_area,
                        ymin = mean_area - se_area,
                        ymax = mean_area + se_area),
                    alpha = 0.5, size = 0.3) +
    geom_line(data = filter(kelp_area3, site_name == i),
              aes(x = year, y = mean_area, group = 1), size = 0.3) +
    scale_y_continuous(name = 'area',
                       sec.axis = sec_axis(~./2, name = 'fit')) +
    labs(fill = 'fit >= 6.6', title = i) +
    theme_bw() +
    theme(axis.text.x = element_text(angle = 90, size = 7),
          plot.title = element_text(hjust = 0.5),
          panel.grid.major = element_blank(),
          legend.title = element_text(size = 9),
          legend.text = element_text(size = 7))
  print(plot)
}
```

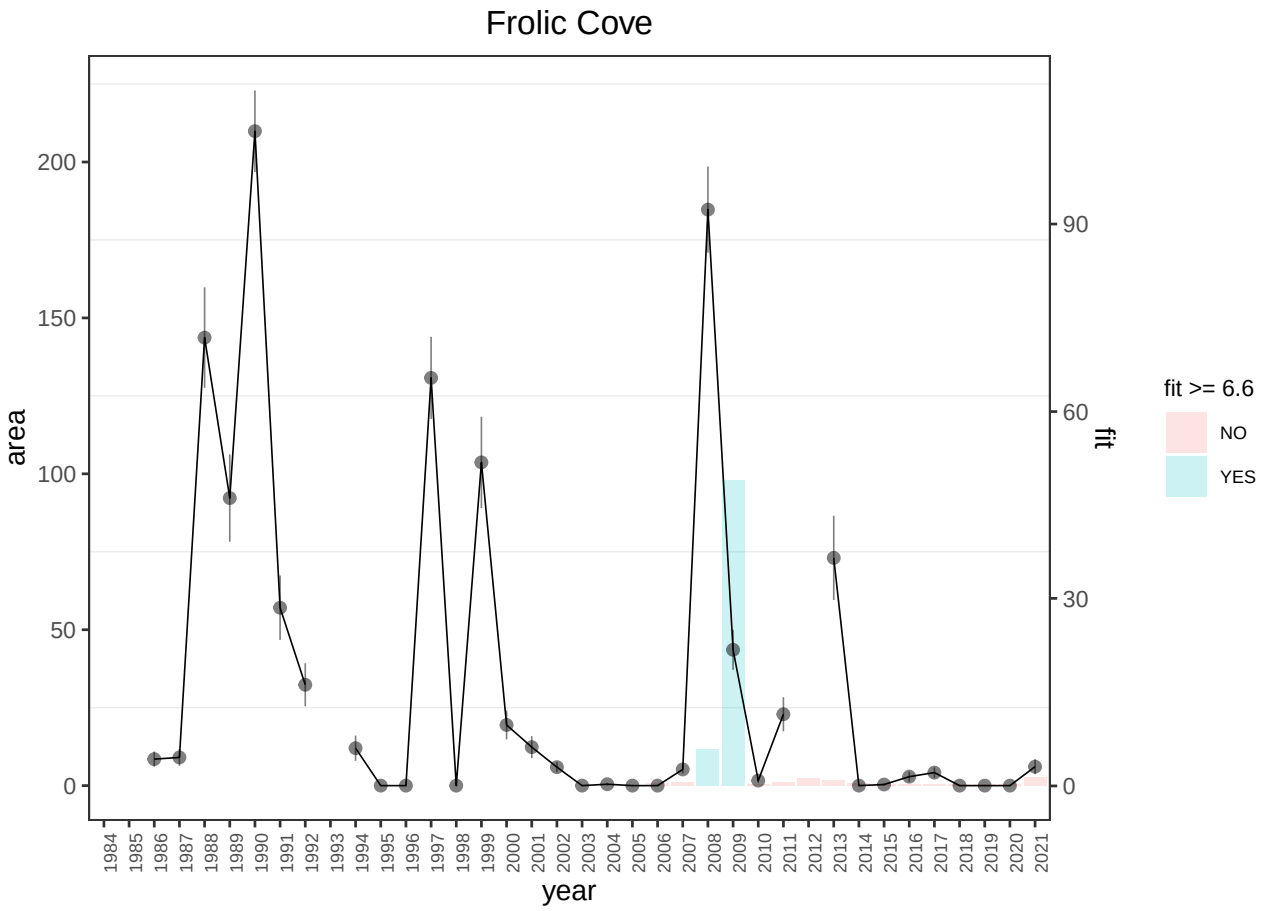

Caspar
area
fit
year
fit >= 6.6
NO
0
50
100
150
200
0
30
60
90
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

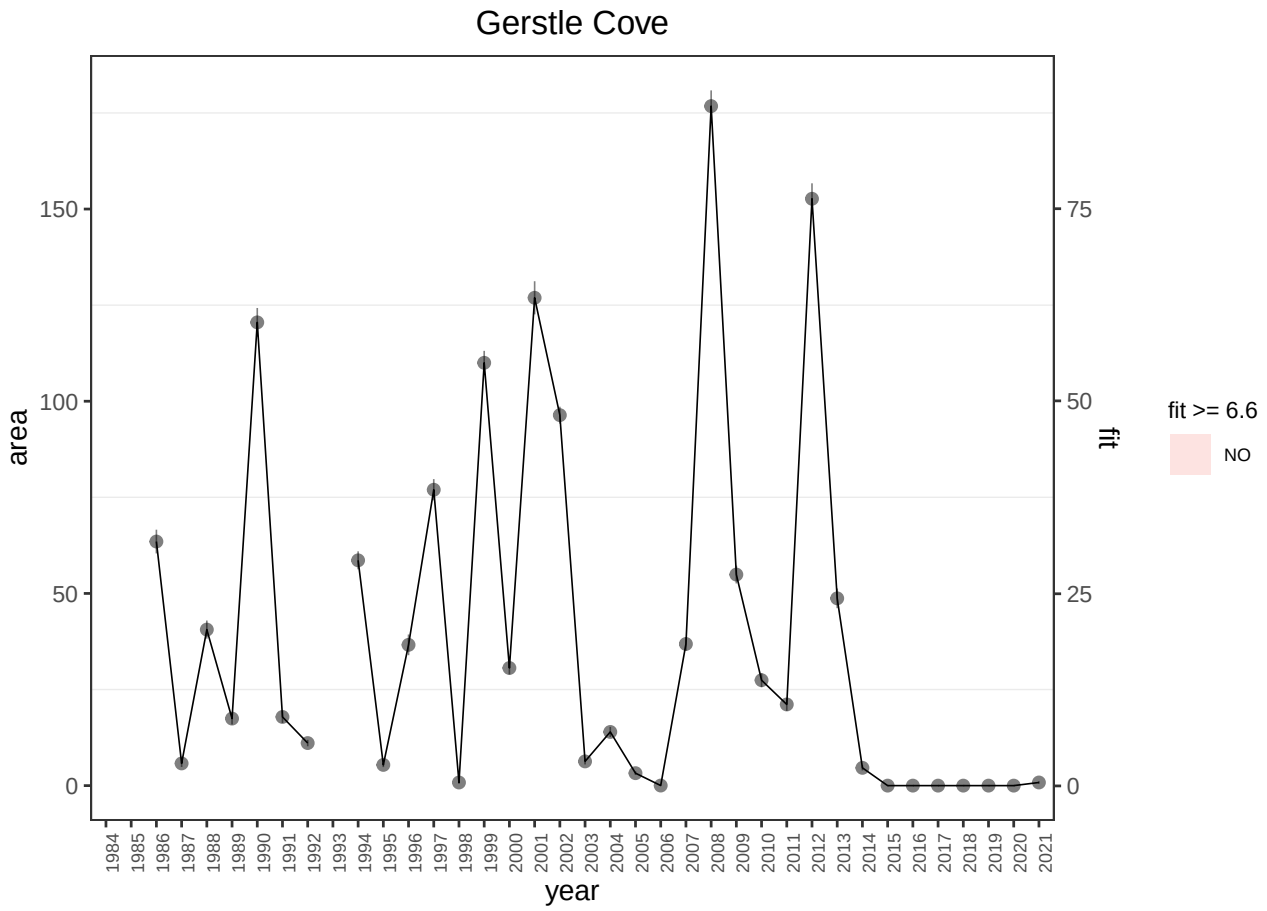

Caspar North
area
fit
year
0
50
100
0
20
40
60
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
fit >= 6.6
NO

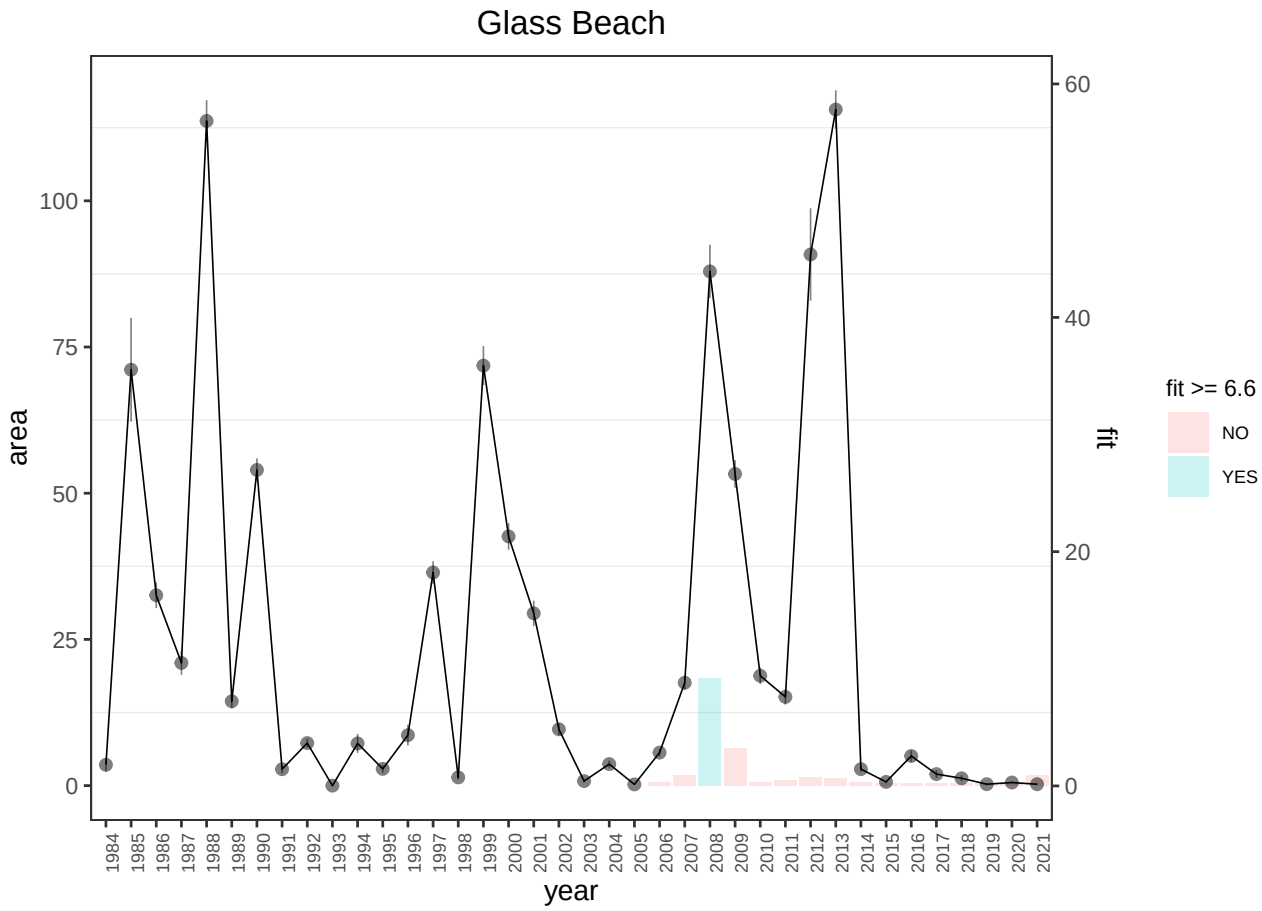

Dark Gulch
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
200
0
25
50
75
100
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

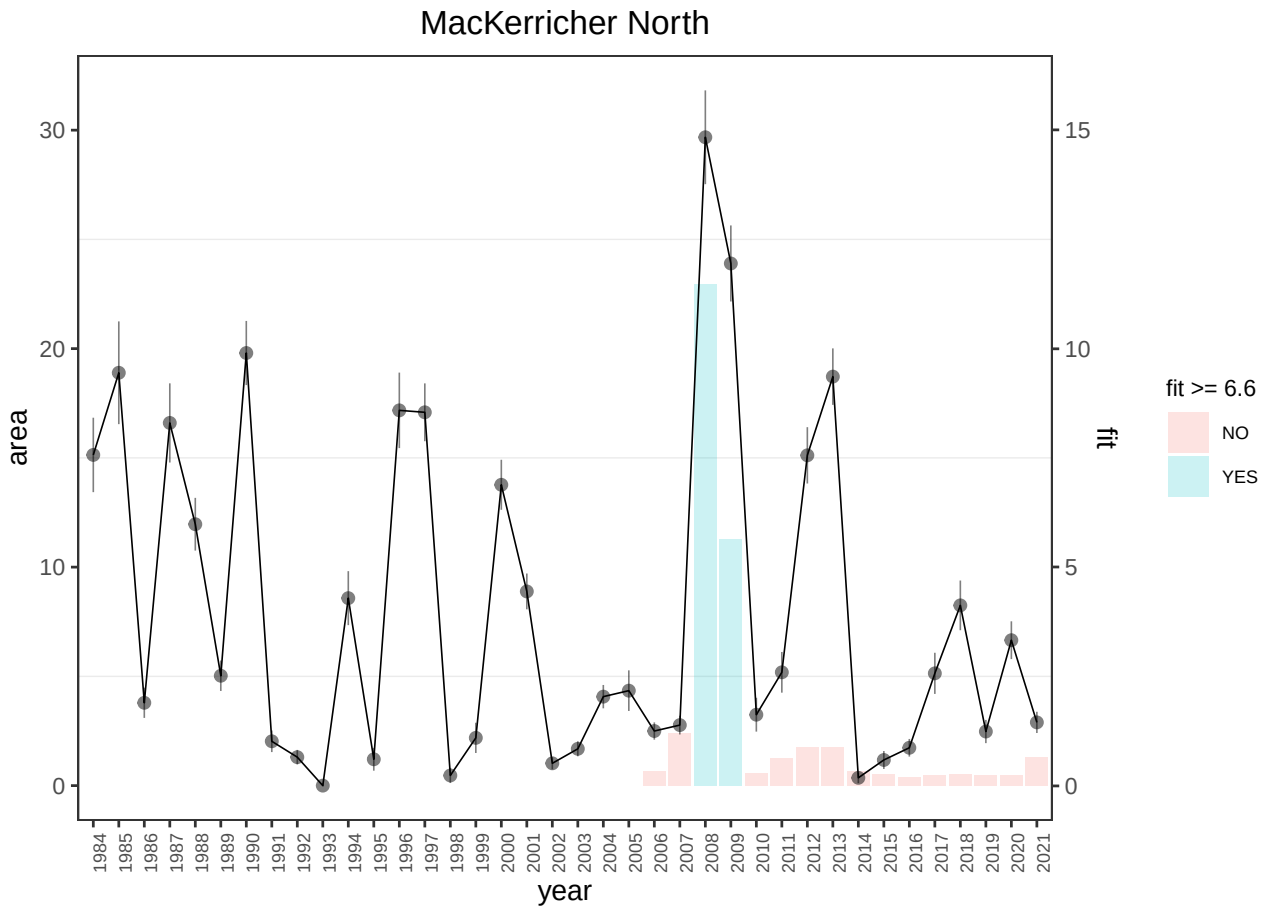

Flat Iron Rock
area
fit
year
fit >= 6.6
NO
YES
0
10
20
30
0
5
10
15
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

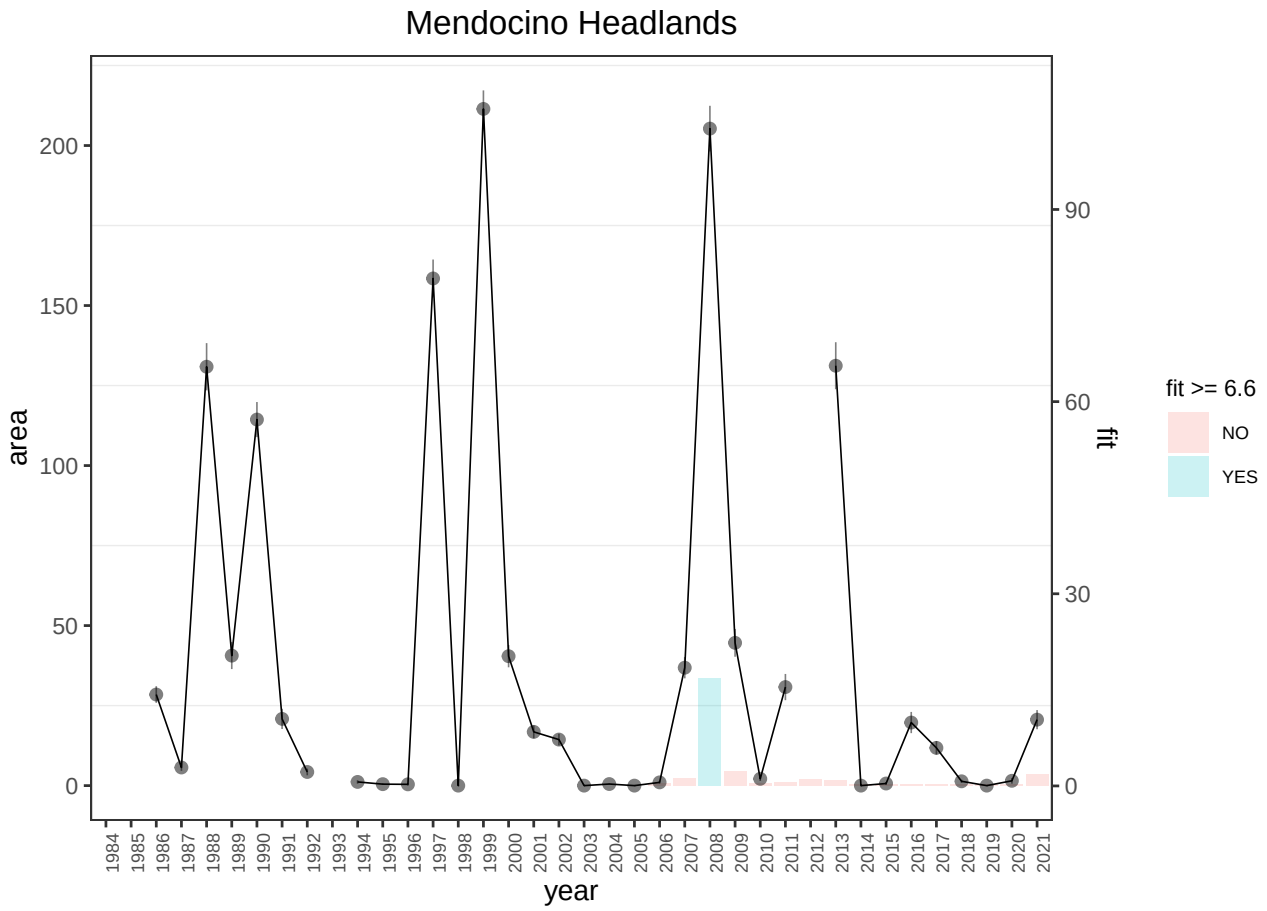

Fort Ross
area
fit
year
fit >= 6.6
NO
200
150
100
50
0
100
75
50
25
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

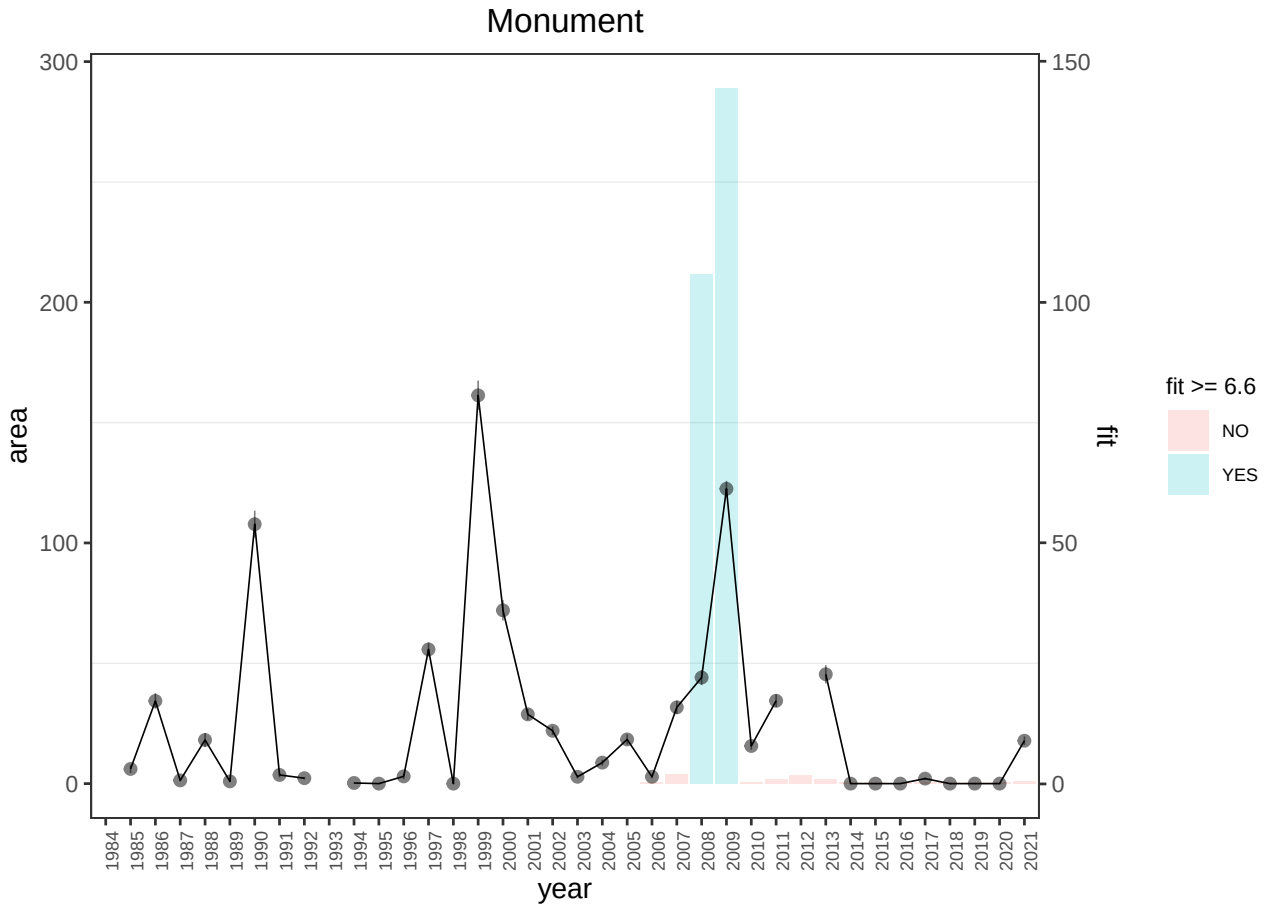

Frolic Cove
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
200
30
60
90
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

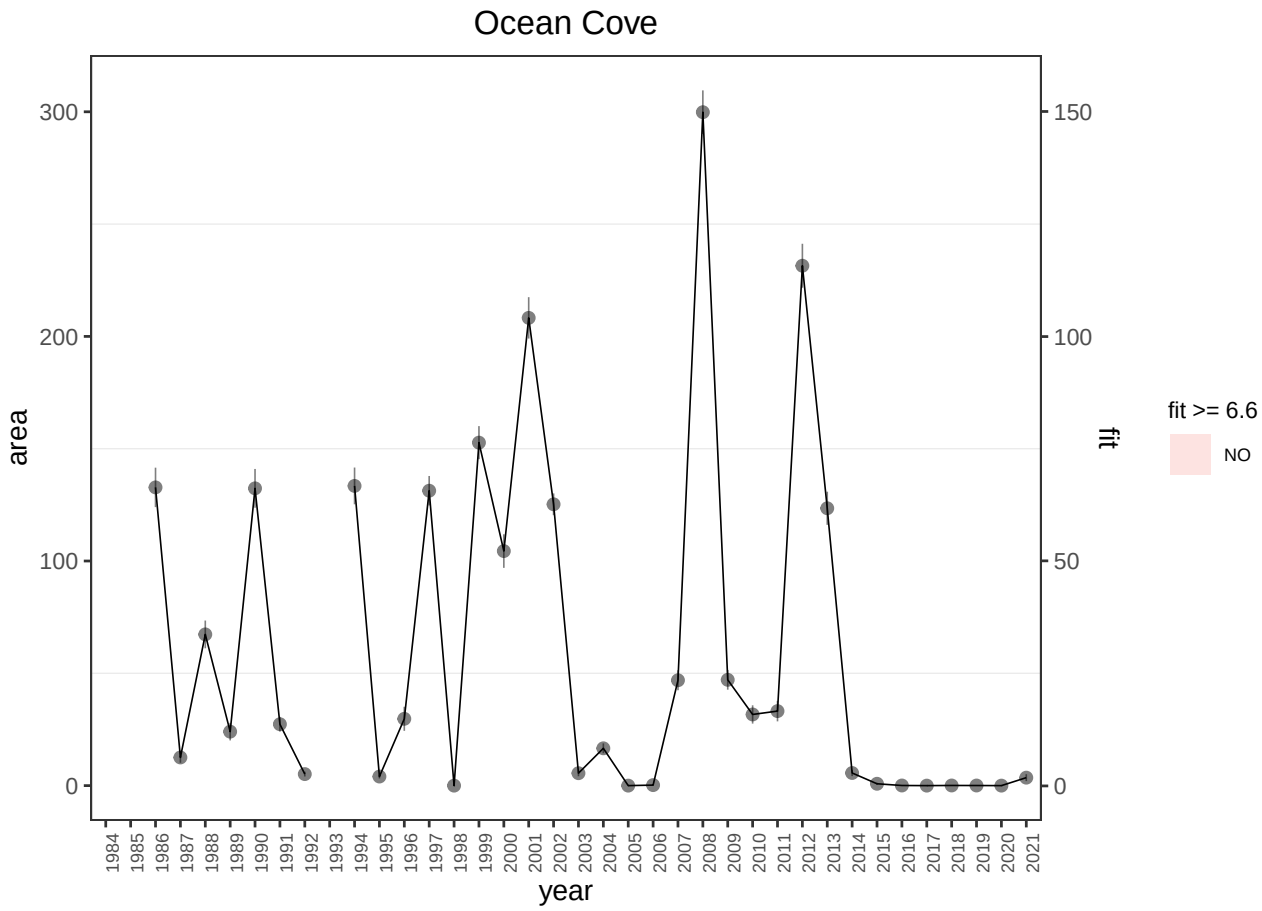

Gerstle Cove
area
fit
year
150
100
50
0
75
50
25
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
fit >= 6.6
NO

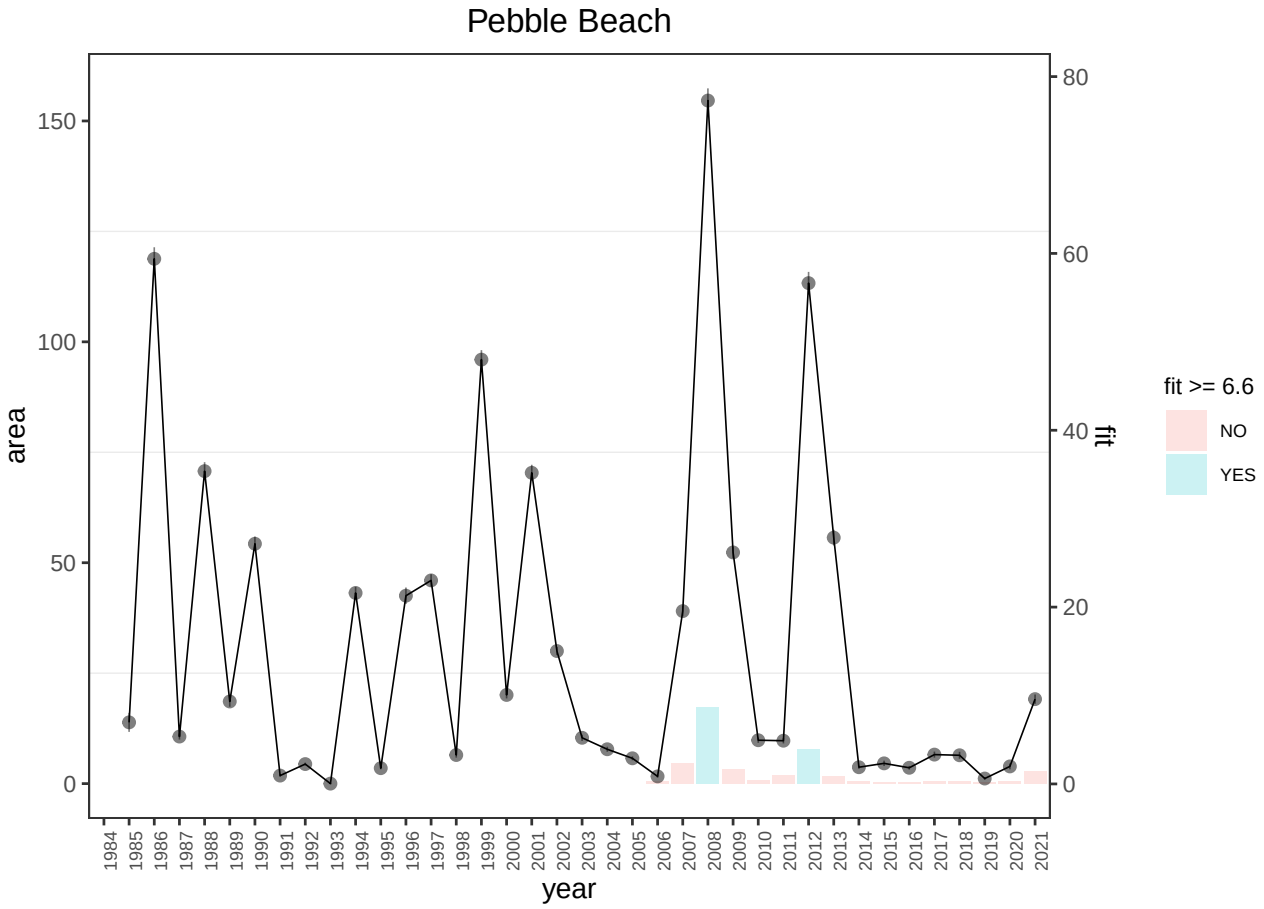

Glass Beach
area
fit
year
fit >= 6.6
NO
YES
0
25
50
75
100
0
20
40
60
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

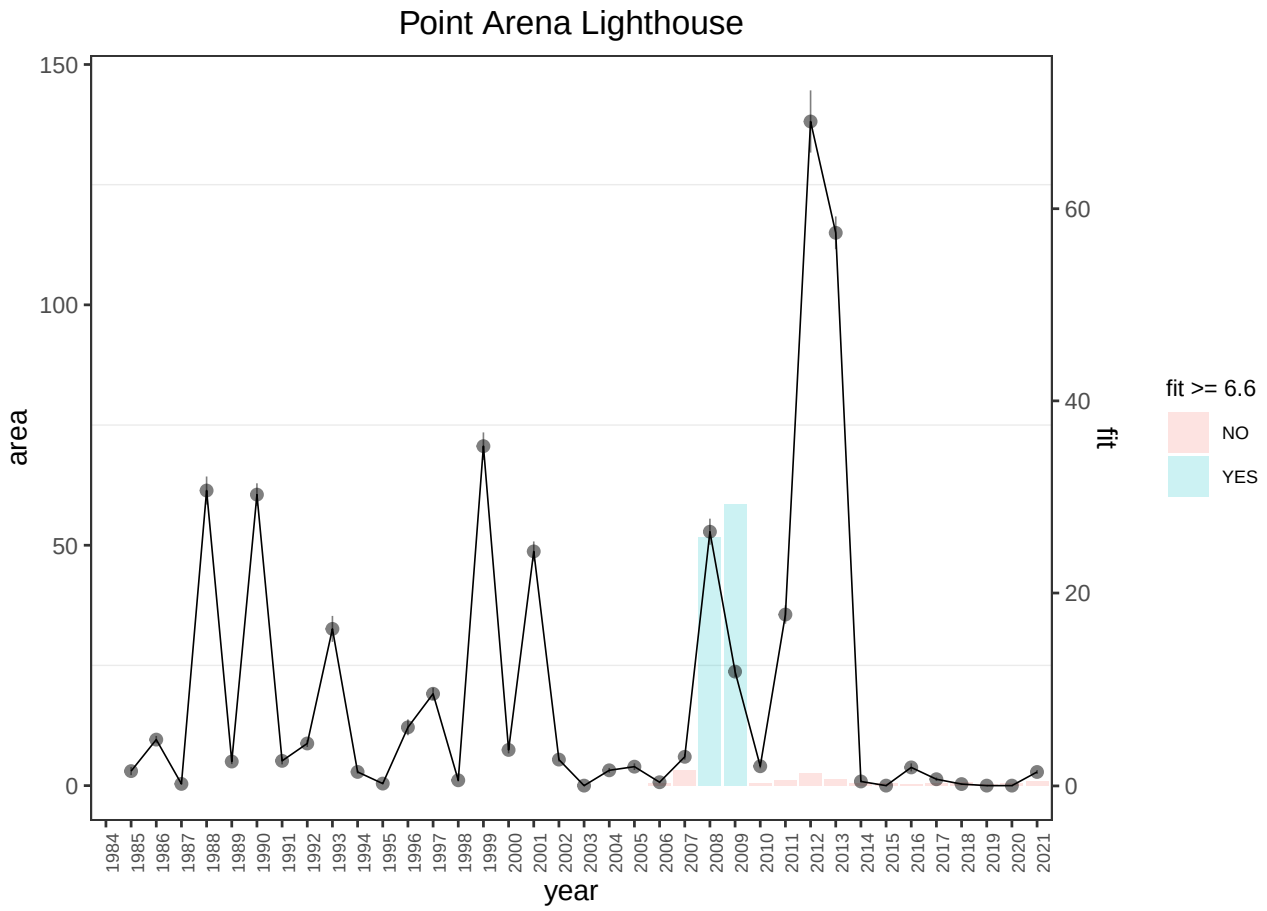

MacKerricher North
area
fit
year
fit >= 6.6
NO
YES
0
10
20
30
0
5
10
15
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

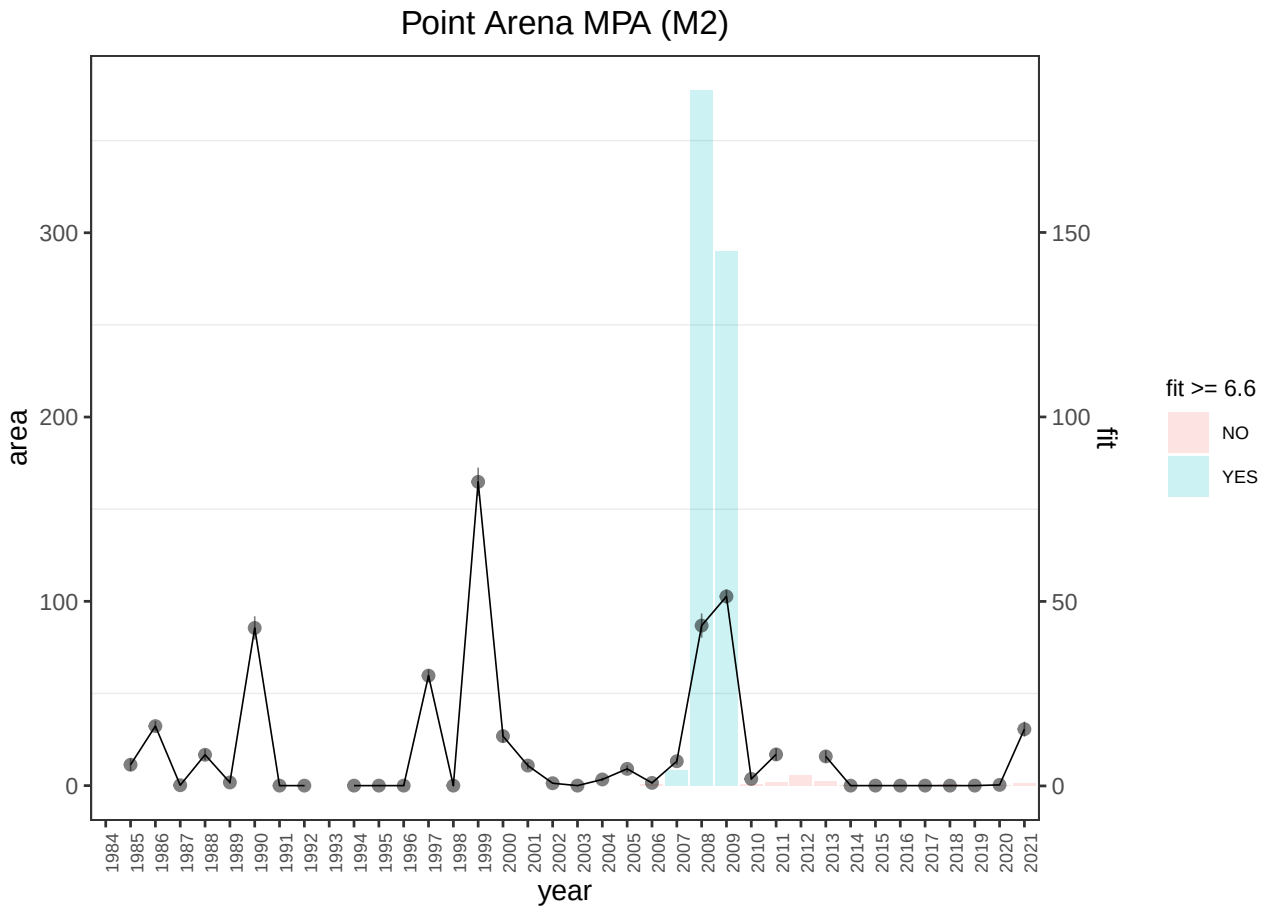

Mendocino Headlands
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
200
0
30
60
90
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

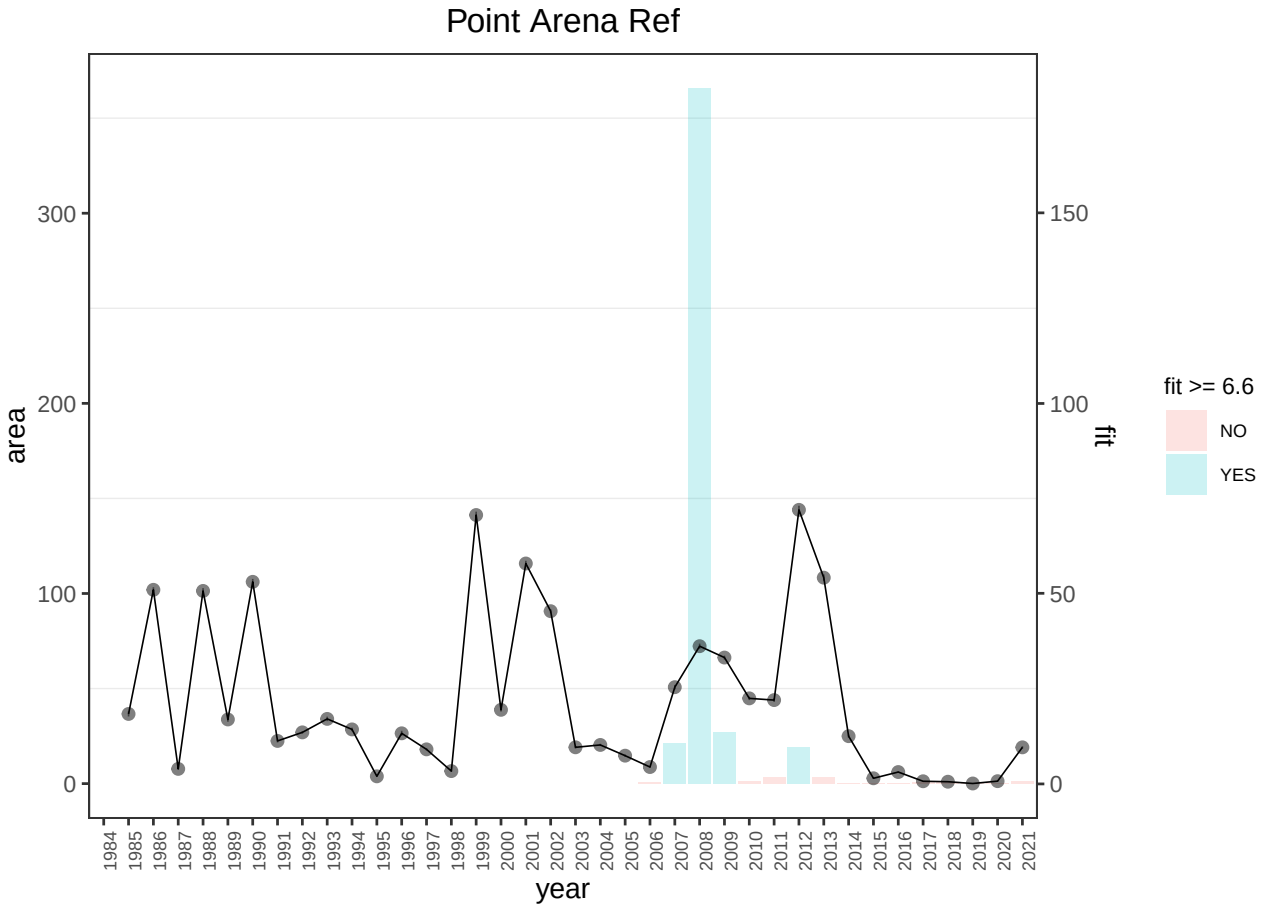

Monument
area
fit
year
fit >= 6.6
NO
YES
300
200
100
0
150
100
50
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

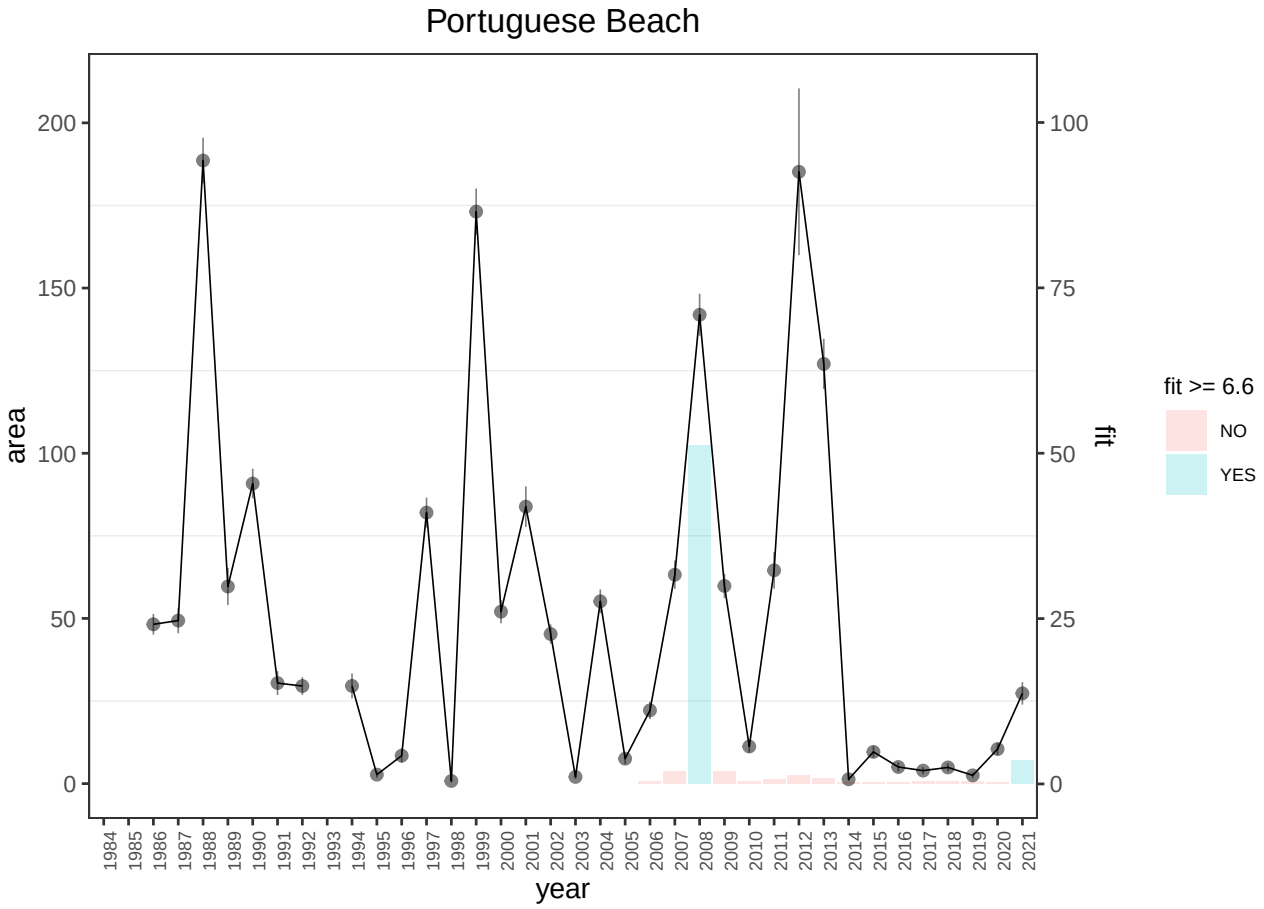

Ocean Cove
area
fit
year
fit >= 6.6
NO
300
200
100
0
150
100
50
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

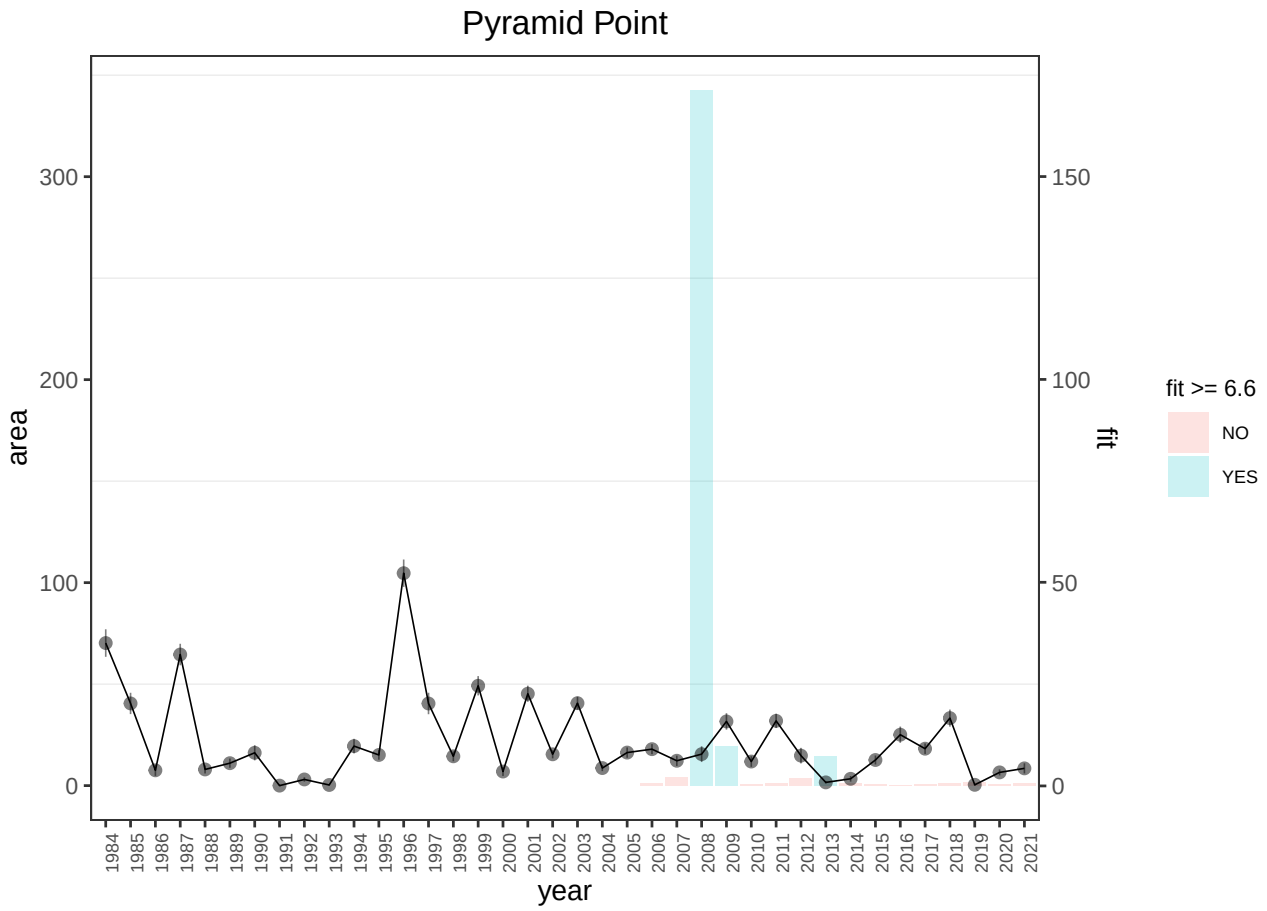

Pebble Beach
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
0
20
40
60
80
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

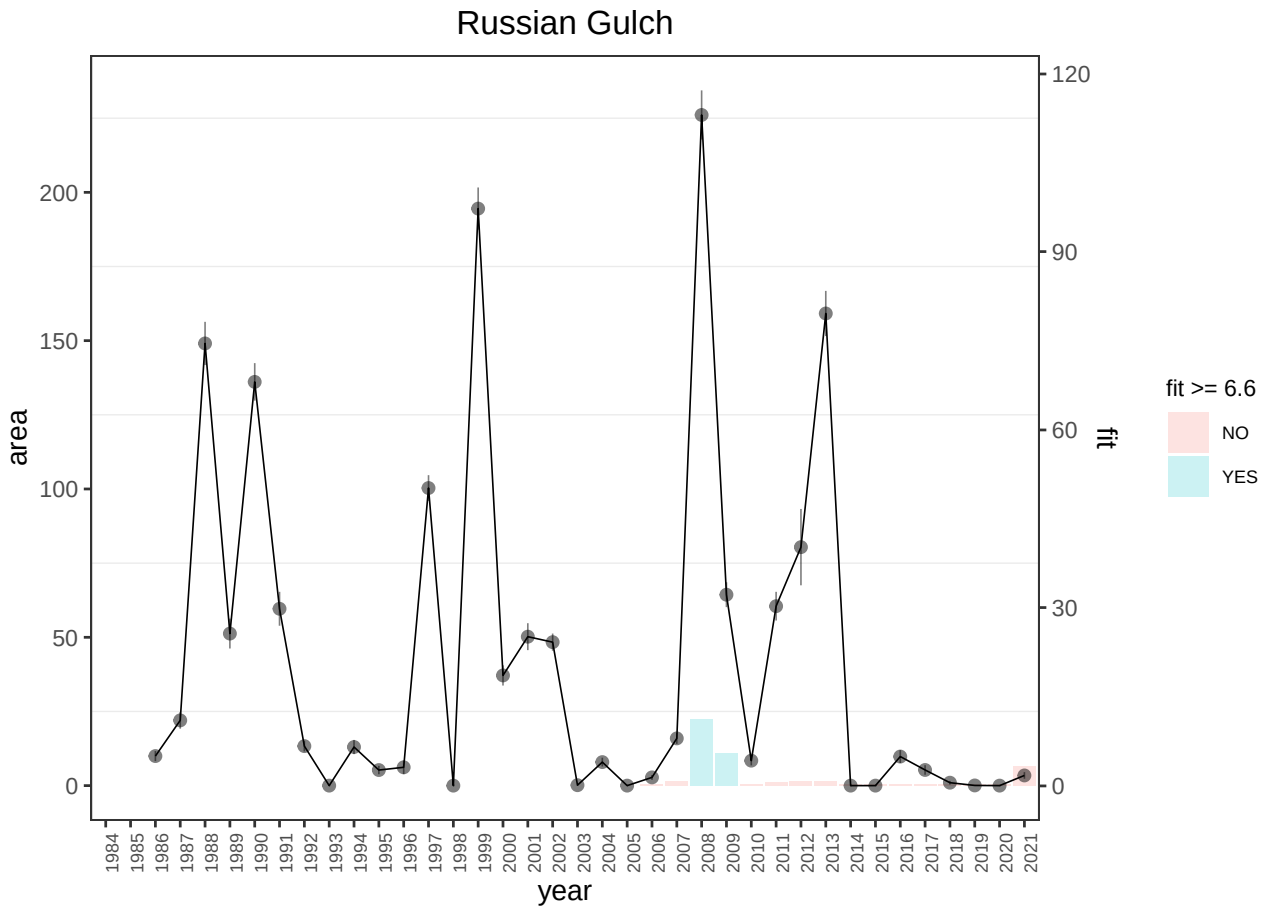

Point Arena Lighthouse
area
fit
year
fit >= 6.6
NO
YES
150
100
50
0
60
40
20
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

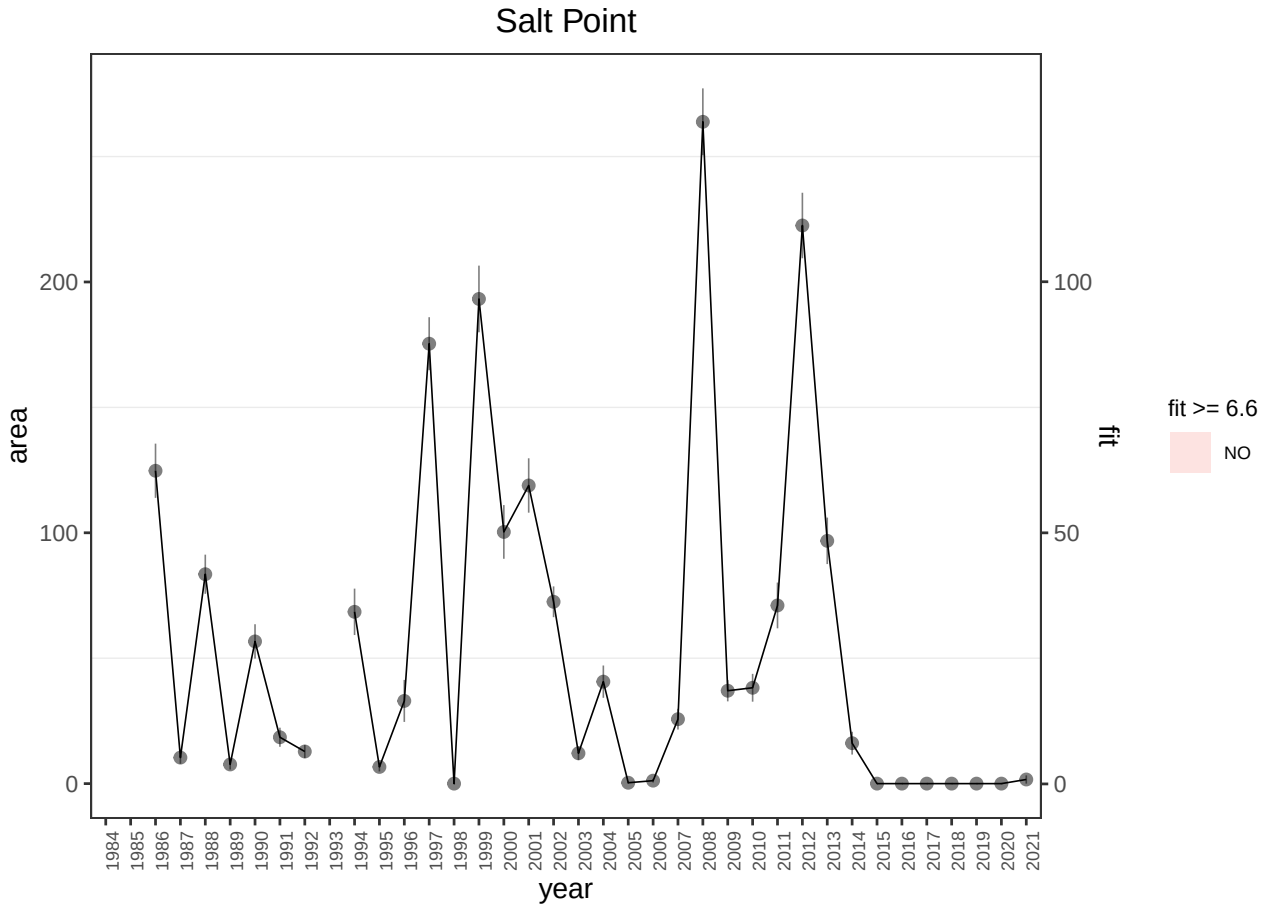

Point Arena MPA (M2)
area
fit
year
fit >= 6.6
NO
YES
0
100
200
300
50
150
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

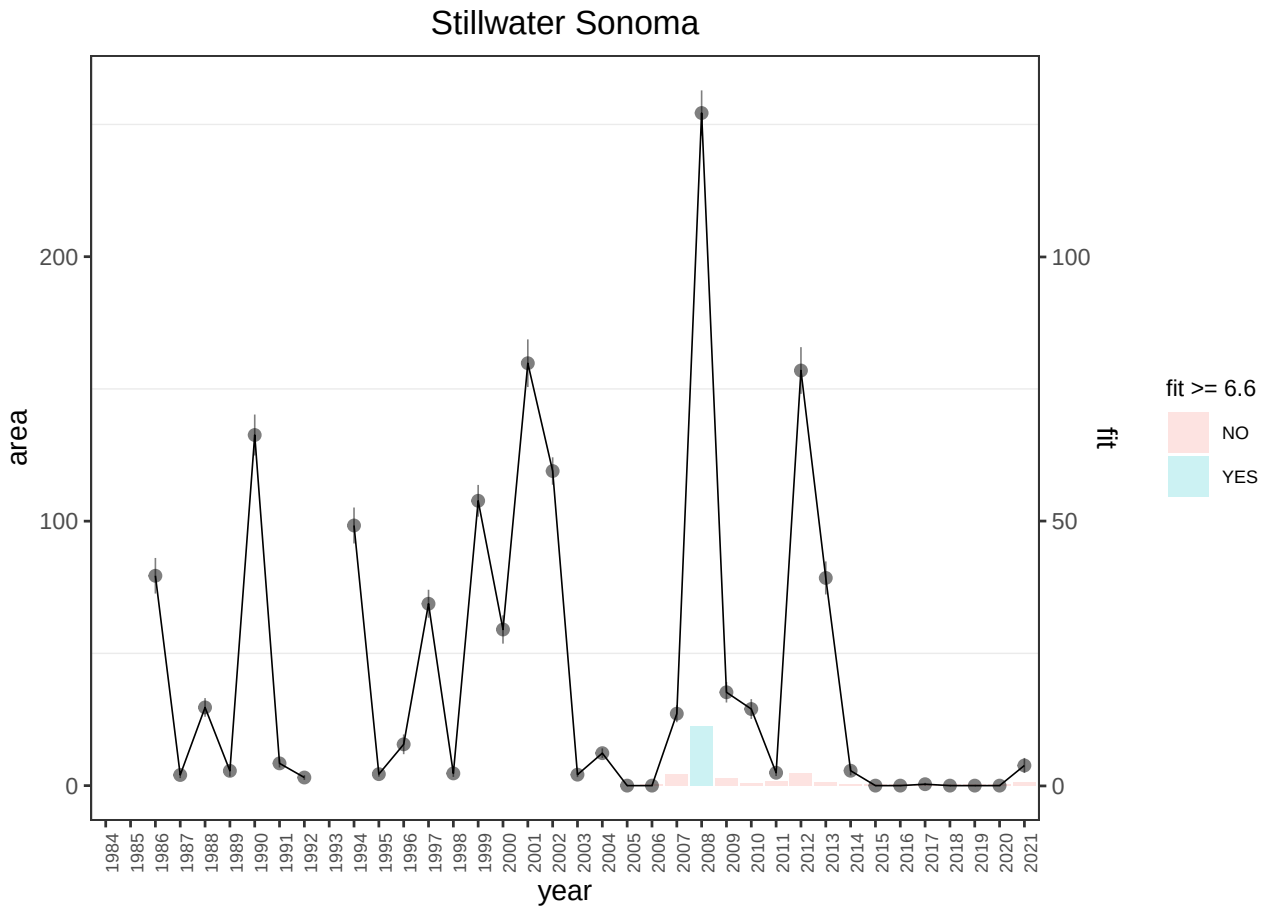

Point Arena Ref
area
fit
year
fit >= 6.6
NO
YES
0
100
200
300
50
150
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

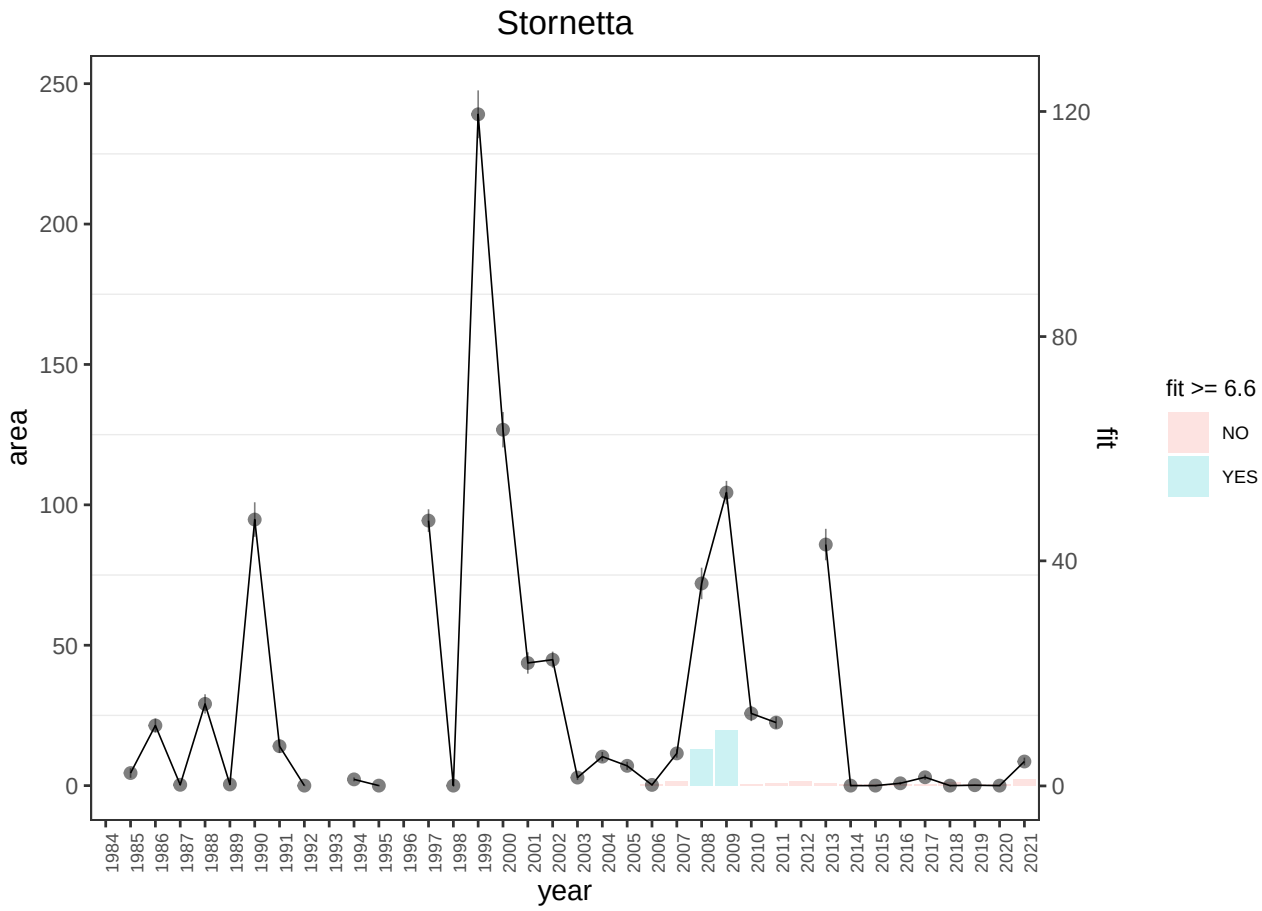

Portuguese Beach
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
200
0
25
50
75
100
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

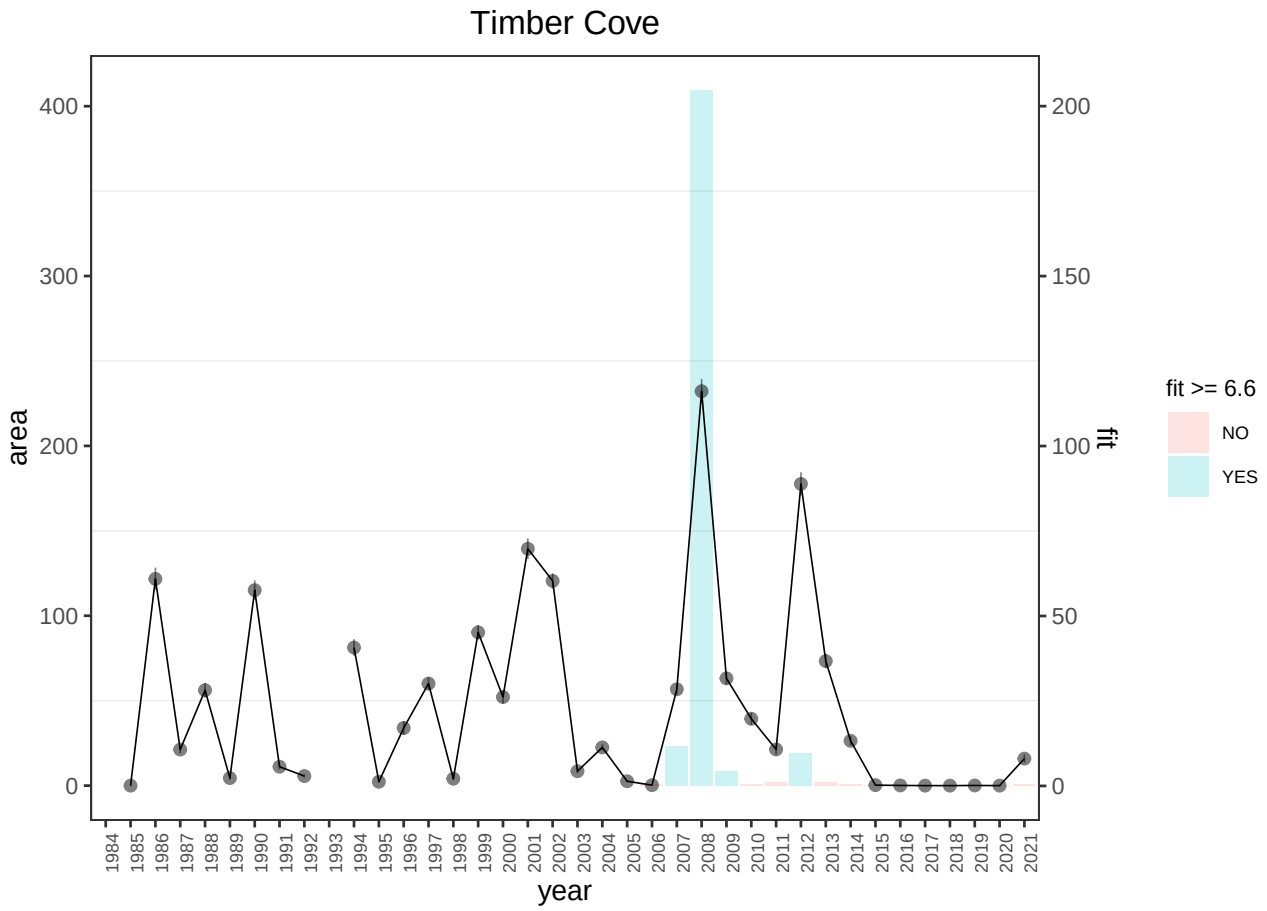

Pyramid Point
area
fit
year
fit >= 6.6
NO
YES
0
100
200
300
50
150
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

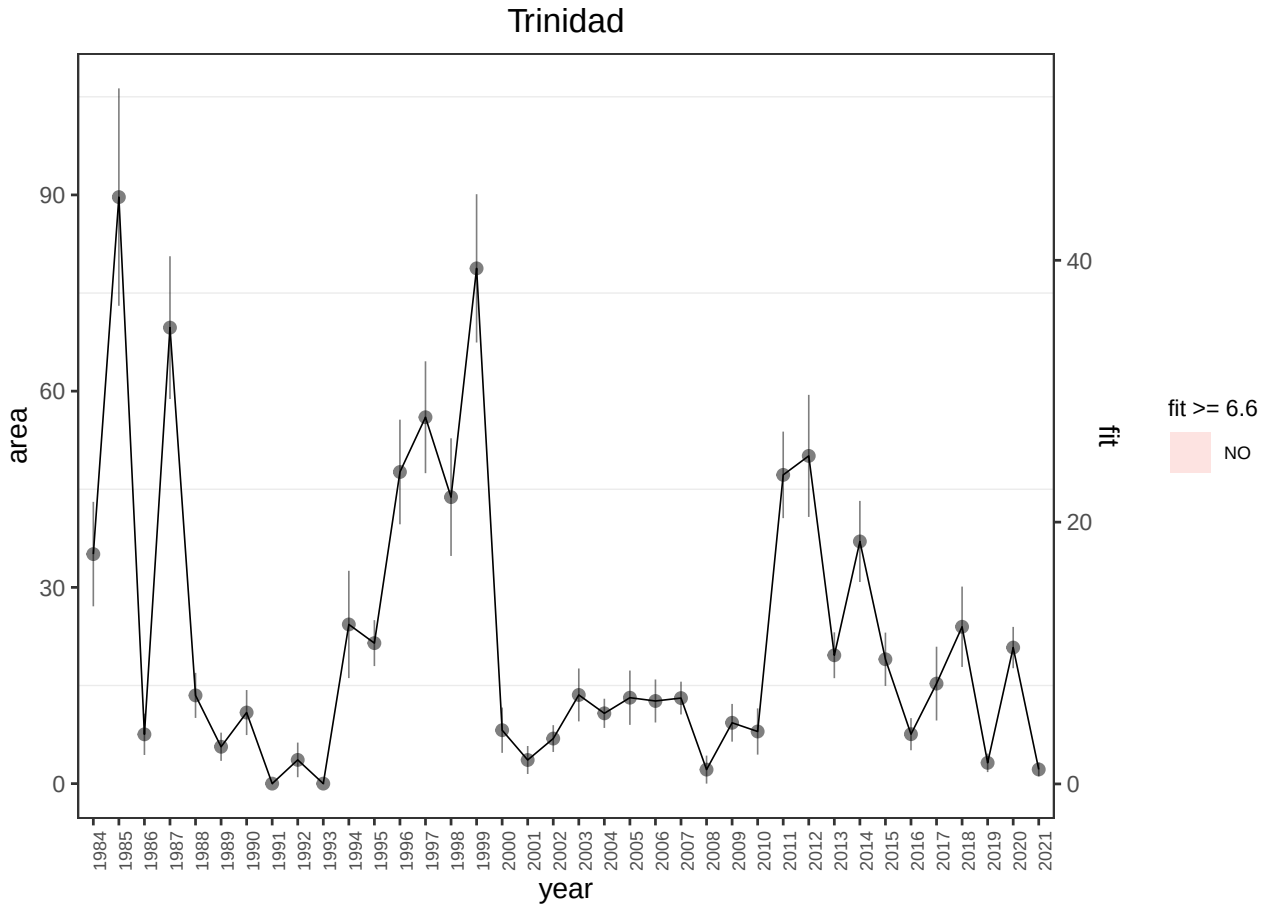

Russian Gulch
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
200
0
30
60
90
120
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021

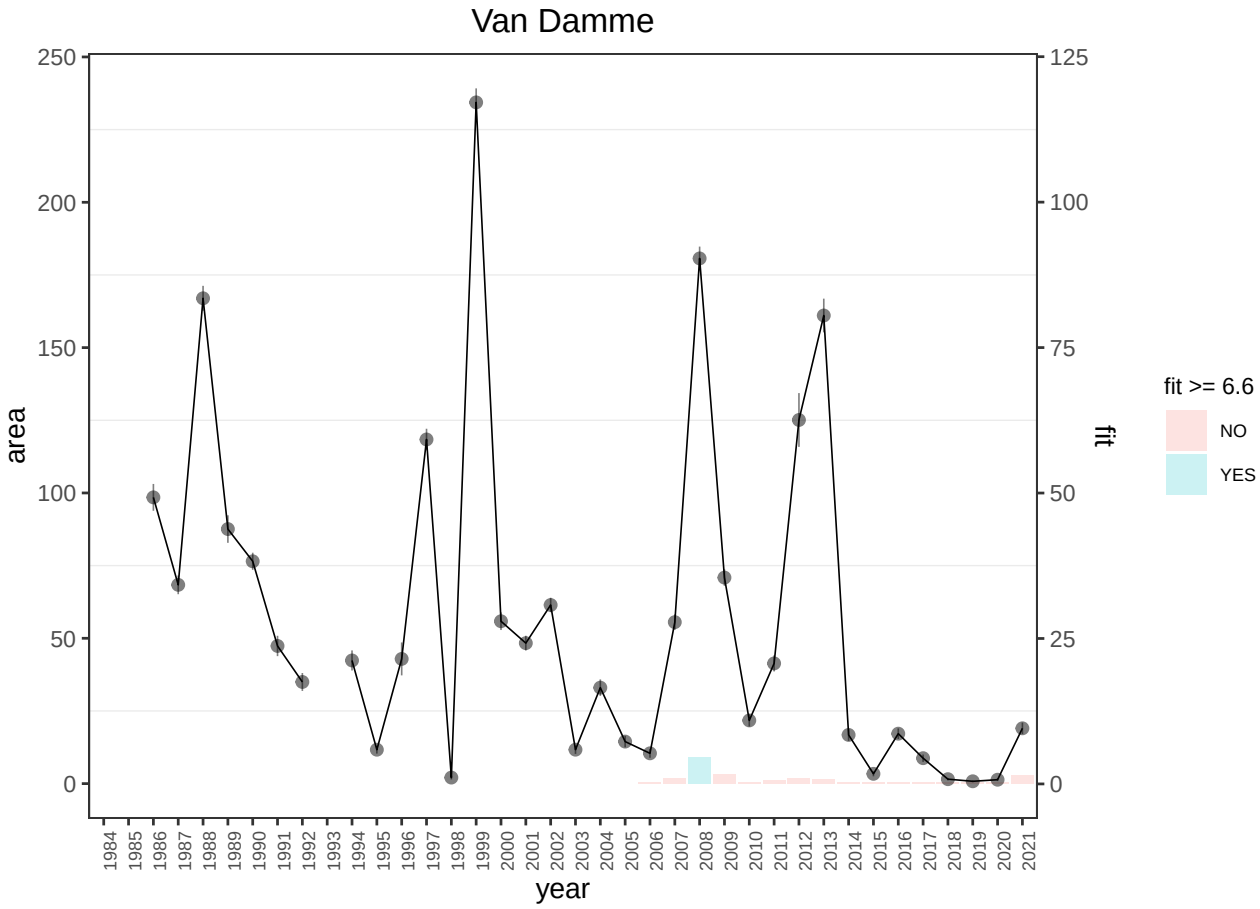

Salt Point
area
fit
year
200
100
0
100
50
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
fit >= 6.6
NO


Stillwater Sonoma
area
fit
year
fit >= 6.6
NO
YES
200
100
0
100
50
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021


Stornetta
area
fit
year
fit >= 6.6
NO
YES
0
50
100
150
200
250
0
40
80
120
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021


Timber Cove
area
fit
year
fit >= 6.6
NO
YES
0
100
200
300
400
50
150
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021


Trinidad
area
fit
year
fit >= 6.6
NO
0
30
60
90
0
20
40
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021


Van Damme
area
fit
year
fit >= 6.6
NO
YES
250
200
150
100
50
0
125
100
75
50
25
0
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021