SURE Project

September 14, 2022

Objective

Compare spatial predictions of kelp to "in situ" survey data. Compare each year and location for 150, 300, 600, 900 resolutions.

Extraction

Extract the predicted *log* kelps density of every year (2004 - 2021) for each site in the North Coast.

```
# set a directory
w.dir <- here()
d.dir <- here('data')
r1.dir <- here('spatial_data/sp_predictions_300m')
r2.dir <- here('spatial_data/sp_predictions_150m')
r4.dir <- here('spatial_data/sp_predictions_900m')
# r3.dir <- here('spatial_data/sp_predictions_150m_resolution')
# r4.dir <- here('spatial_data/sp_predictions_600m_resolution')

# read and transform the observed data to the log scale
df <- read.csv(paste(d.dir,
                     'RCCA_kelp_inverts_NC_depth-zones_wave_clim_temp_nit_subs_orbvel_npp.csv',
                     sep = '/')) %>%
  dplyr::select(site_name, year, transect, zone, latitude, longitude, den_NERLUE) %>%
  mutate_at(vars(year, transect, zone, site_name), list(as.factor)) %>%
  mutate(log_den_NERLUE = log(den_NERLUE))
head(df)
```

```
##   site_name year transect  zone latitude longitude den_NERLUE log_den_NERLUE
## 1    Caspar 2018        1 INNER 39.36173  -123.822          0           -Inf
## 2    Caspar 2018        2 INNER 39.36173  -123.822          0           -Inf
## 3    Caspar 2018        3 INNER 39.36173  -123.822          0           -Inf
## 4    Caspar 2018        4 OUTER 39.36173  -123.822          0           -Inf
## 5    Caspar 2018        5 OUTER 39.36173  -123.822          0           -Inf
## 6    Caspar 2018        6 OUTER 39.36173  -123.822          0           -Inf
```

Note that log(0) returns -Inf. How to deal with log(0)?

```
df$log_den_NERLUE <- replace(df$log_den_NERLUE, df$log_den_NERLUE == -Inf, 0)
```

Calculate the mean and standard error of kelps density of every year for each site by zone (INNER/OUTER).

```r
obs <- df %>%
  group_by(site_name, year, zone) %>%
  summarise_at(vars(log_den_NERLUE), list(mean = mean, se = std.error), na.rm = TRUE) %>%
  pivot_wider(names_from = zone, values_from = c(mean, se))
head(obs)
```

```
## # A tibble: 6 x 6
## # Groups:   site_name, year [6]
##   site_name year  mean_INNER mean_OUTER se_INNER se_OUTER
##   <fct>     <fct>      <dbl>      <dbl>    <dbl>    <dbl>
## 1 Caspar    2008       4.38        3.03   0.150     0.996
## 2 Caspar    2010       4.37        4.17   0.0664    0.586
## 3 Caspar    2014       0.799       0      0.799     0
## 4 Caspar    2015       0           0      0         0
## 5 Caspar    2016       0           0      0         0
## 6 Caspar    2017       0           0      0         0
```

Extract the predicted *log* kelps density of every year for each site at different resolutions.

```r
# kelp density predictions at 300m resolution

# read the .csv file
site <- read.csv(paste(d.dir, 'RCCA_North_Coast_sites.csv', sep = '/'))
# convert from .csv to .shp
site_shp <- st_as_sf(site, coords = c('longitude', 'latitude'), crs = 'EPSG:4326')

# declaring an empty data frame
pred <- data.frame(site_name = character(),
                   year = numeric(),
                   fit = numeric())

for (i in c(2006:2021)) {
  rast <- rast(paste0(r1.dir, paste0('/', i, '_Log_Nereo_NC.tif')))
  ext <- terra::extract(rast, vect(site_shp$geometry)) %>%
    mutate(site_name = site$site_name, year = as.factor(i), .before = fit) %>%
    dplyr::select(-ID)
  pred <- rbind(pred, ext)
}

head(pred)
```

```
##        site_name year       fit
## 1         Caspar 2006 0.1042327
## 2   Caspar North 2006 0.2256172
## 3     Dark Gulch 2006 0.2406471
## 4 Flat Iron Rock 2006 0.1637378
## 5      Fort Ross 2006       NaN
## 6    Frolic Cove 2006 0.1930680
```

```r
# write to cvs
merge_df <- left_join(pred,
                      site %>% dplyr::select(c(site_name, longitude, latitude)),
```

```
                      by = 'site_name')

# write.csv(merge_df, file.path(d.dir, 'NC_kelp_density_predictions_300m.csv'), row.names = FALSE)
```

```
# kelp density predictions at 150m resolution

# read the .csv file
site <- read.csv(paste(d.dir, 'RCCA_North_Coast_sites.csv', sep = '/'))
# convert from .csv to .shp
site_shp <- st_as_sf(site, coords = c('longitude', 'latitude'), crs = 'EPSG:4326')

# declaring an empty data frame
pred <- data.frame(site_name = character(),
                   year = numeric(),
                   fit = numeric())

for (i in c(2006:2021)) {
  rast <- rast(paste0(r2.dir, paste0('/', i, '_Log_Nereo_NC.tif')))
  ext <- terra::extract(rast, vect(site_shp$geometry)) %>%
    mutate(site_name = site$site_name, year = as.factor(i), .before = fit) %>%
    dplyr::select(-ID)
  pred <- rbind(pred, ext)
}

head(pred)
```

```
##        site_name year       fit
## 1         Caspar 2006 0.2277283
## 2   Caspar North 2006 0.2246296
## 3     Dark Gulch 2006 0.1929662
## 4 Flat Iron Rock 2006 0.1812405
## 5      Fort Ross 2006       NaN
## 6    Frolic Cove 2006 0.1510025
```

```
# write to cvs
merge_df <- left_join(pred,
                      site %>% dplyr::select(c(site_name, longitude, latitude)),
                      by = 'site_name')

# write.csv(merge_df, file.path(d.dir, 'NC_kelp_density_predictions_150m.csv'), row.names = FALSE)
```

```
# kelp density predictions at 900m resolution

# read the .csv file
site <- read.csv(paste(d.dir, 'RCCA_North_Coast_sites.csv', sep = '/'))
# convert from .csv to .shp
site_shp <- st_as_sf(site, coords = c('longitude', 'latitude'), crs = 'EPSG:4326')

# declaring an empty data frame
pred <- data.frame(site_name = character(),
                   year = numeric(),
                   fit = numeric())
```

```
for (i in c(2006:2021)) {
  rast <- rast(paste0(r4.dir, paste0('/', i, '_Log_Nereo_NC.tif')))
  ext <- terra::extract(rast, vect(site_shp$geometry)) %>%
    mutate(site_name = site$site_name, year = as.factor(i), .before = fit) %>%
    dplyr::select(-ID)
  pred <- rbind(pred, ext)
}

head(pred)
```

```
##        site_name year        fit
## 1         Caspar 2006 0.07571632
## 2   Caspar North 2006 0.07571632
## 3     Dark Gulch 2006 0.33232445
## 4 Flat Iron Rock 2006 0.18413925
## 5      Fort Ross 2006        NaN
## 6    Frolic Cove 2006 0.24223915
```

```
# write to cvs
merge_df <- left_join(pred,
                      site %>% dplyr::select(c(site_name, longitude, latitude)),
                      by = 'site_name')

write.csv(merge_df, file.path(d.dir, 'NC_kelp_density_predictions_900m.csv'), row.names = FALSE)
```

Comparison

```
# kelp density predictions at 300m resolution
pred_300m <- read.csv(paste(d.dir, 'NC_kelp_density_predictions_300m.csv', sep ='/')) %>%
  mutate_at(vars(year, site_name), list(as.factor))
head(pred_300m)
```

```
##        site_name year       fit longitude latitude
## 1         Caspar 2006 0.1042327 -123.8220 39.36173
## 2   Caspar North 2006 0.2256172 -123.8213 39.36443
## 3     Dark Gulch 2006 0.2406471 -123.7762 39.24030
## 4 Flat Iron Rock 2006 0.1637378 -124.1578 41.05942
## 5      Fort Ross 2006        NA -123.2450 38.51060
## 6    Frolic Cove 2006 0.1930680 -123.8239 39.35503
```

```
kelp_data_300m <- left_join(pred_300m, obs, by = c('site_name', 'year')) %>%
  group_by(site_name) %>%
  arrange(year, .by_group = TRUE) %>%
  relocate(fit, .after = last_col())
head(kelp_data_300m)
```

```
## # A tibble: 6 x 9
## # Groups:   site_name [1]
##   site_name year  longitude latitude mean_INNER mean_OUTER se_IN~1 se_OU~2   fit
```

```
##   <fct>     <fct>     <dbl>    <dbl>      <dbl>      <dbl>   <dbl>   <dbl> <dbl>
## 1 Caspar    2006      -124.     39.4      NA         NA    NA       NA     0.104
## 2 Caspar    2007      -124.     39.4      NA         NA    NA       NA     2.02
## 3 Caspar    2008      -124.     39.4       4.38       3.03  0.150    0.996 3.17
## 4 Caspar    2009      -124.     39.4      NA         NA    NA       NA     6.93
## 5 Caspar    2010      -124.     39.4       4.37       4.17  0.0664   0.586 0.522
## 6 Caspar    2011      -124.     39.4      NA         NA    NA       NA     2.00
## # ... with abbreviated variable names 1: se_INNER, 2: se_OUTER
```

```
# kelp density predictions at 150m resolution
pred_150m <- read.csv(paste(d.dir, 'NC_kelp_density_predictions_150m.csv', sep ='/')) %>%
  mutate_at(vars(year, site_name), list(as.factor))
head(pred_150m)
```

```
##        site_name year       fit longitude latitude
## 1         Caspar 2006 0.2277283 -123.8220 39.36173
## 2   Caspar North 2006 0.2246296 -123.8213 39.36443
## 3     Dark Gulch 2006 0.1929662 -123.7762 39.24030
## 4 Flat Iron Rock 2006 0.1812405 -124.1578 41.05942
## 5      Fort Ross 2006        NA -123.2450 38.51060
## 6    Frolic Cove 2006 0.1510025 -123.8239 39.35503
```

```
kelp_data_150m <- left_join(pred_150m, obs, by = c('site_name', 'year')) %>%
  group_by(site_name) %>%
  arrange(year, .by_group = TRUE) %>%
  relocate(fit, .after = last_col())
head(kelp_data_150m)
```

```
## # A tibble: 6 x 9
## # Groups:   site_name [1]
##   site_name year  longitude latitude mean_INNER mean_OU~1 se_IN~2 se_OU~3    fit
##   <fct>     <fct>     <dbl>    <dbl>      <dbl>     <dbl>   <dbl>   <dbl>  <dbl>
## 1 Caspar    2006      -124.     39.4      NA         NA    NA       NA      0.228
## 2 Caspar    2007      -124.     39.4      NA         NA    NA       NA      3.16
## 3 Caspar    2008      -124.     39.4       4.38       3.03  0.150    0.996  4.41
## 4 Caspar    2009      -124.     39.4      NA         NA    NA       NA     19.0
## 5 Caspar    2010      -124.     39.4       4.37       4.17  0.0664   0.586  0.871
## 6 Caspar    2011      -124.     39.4      NA         NA    NA       NA      3.06
## # ... with abbreviated variable names 1: mean_OUTER, 2: se_INNER, 3: se_OUTER
```

```
# kelp density predictions at 900m resolution
pred_900m <- read.csv(paste(d.dir, 'NC_kelp_density_predictions_900m.csv', sep ='/')) %>%
  mutate_at(vars(year, site_name), list(as.factor))
head(pred_900m)
```

```
##        site_name year        fit longitude latitude
## 1         Caspar 2006 0.07571632 -123.8220 39.36173
## 2   Caspar North 2006 0.07571632 -123.8213 39.36443
## 3     Dark Gulch 2006 0.33232445 -123.7762 39.24030
## 4 Flat Iron Rock 2006 0.18413925 -124.1578 41.05942
## 5      Fort Ross 2006         NA -123.2450 38.51060
## 6    Frolic Cove 2006 0.24223915 -123.8239 39.35503
```

```
kelp_data_900m <- left_join(pred_900m, obs, by = c('site_name', 'year')) %>%
  group_by(site_name) %>%
  arrange(year, .by_group = TRUE) %>%
  relocate(fit, .after = last_col())
head(kelp_data_900m)
```

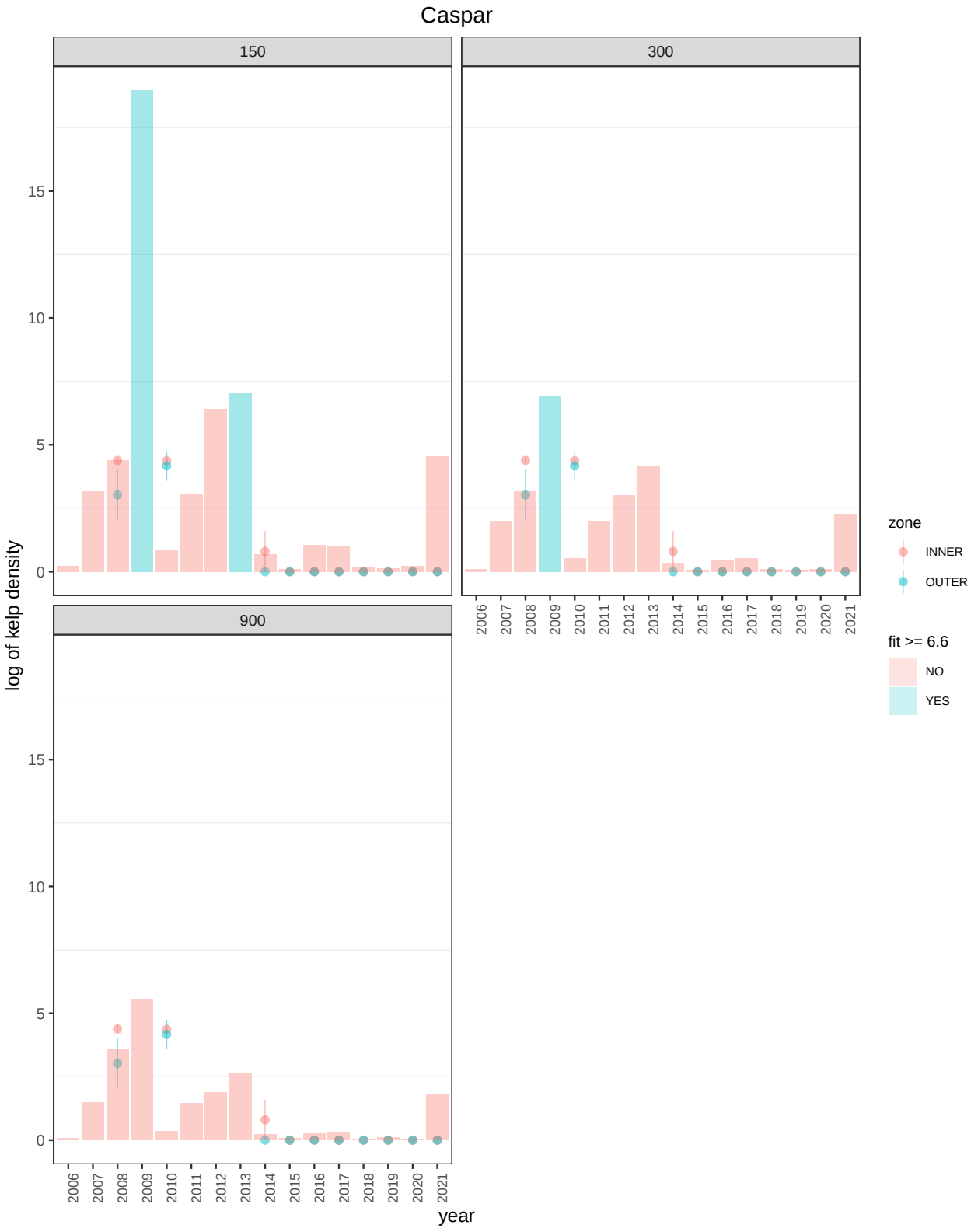
```
## # A tibble: 6 x 9
## # Groups:   site_name [1]
##   site_name year  longitude latitude mean_INNER mean_OU~1 se_IN~2 se_OU~3    fit
##   <fct>     <fct>     <dbl>    <dbl>      <dbl>     <dbl>   <dbl>   <dbl>  <dbl>
## 1 Caspar    2006      -124.     39.4      NA        NA     NA       NA     0.0757
## 2 Caspar    2007      -124.     39.4      NA        NA     NA       NA     1.50
## 3 Caspar    2008      -124.     39.4       4.38      3.03  0.150    0.996 3.57
## 4 Caspar    2009      -124.     39.4      NA        NA     NA       NA     5.56
## 5 Caspar    2010      -124.     39.4       4.37      4.17  0.0664   0.586 0.339
## 6 Caspar    2011      -124.     39.4      NA        NA     NA       NA     1.44
## # ... with abbreviated variable names 1: mean_OUTER, 2: se_INNER, 3: se_OUTER
```

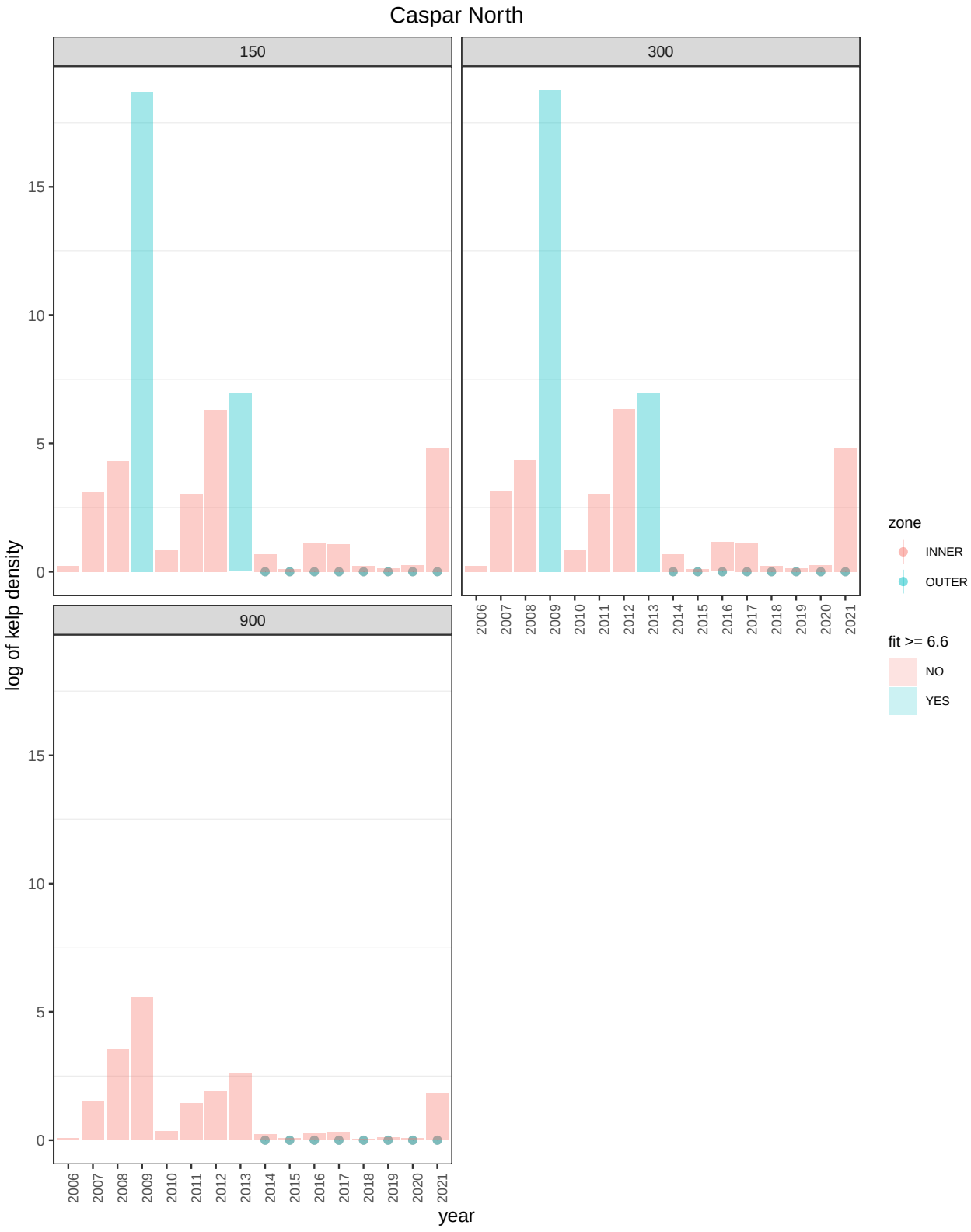
Plotting

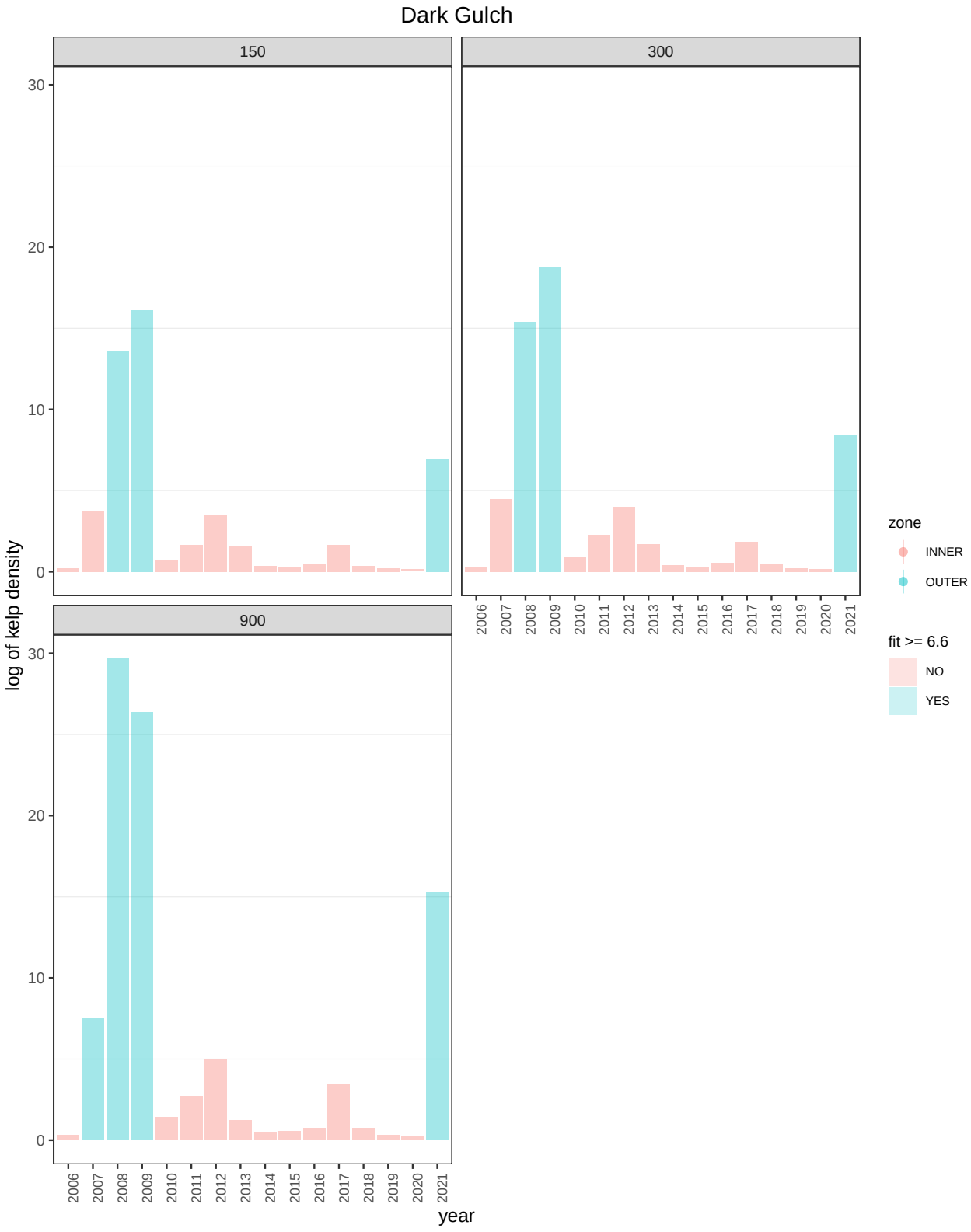
Plot log of kelps density vs year for each site at different resolutions.

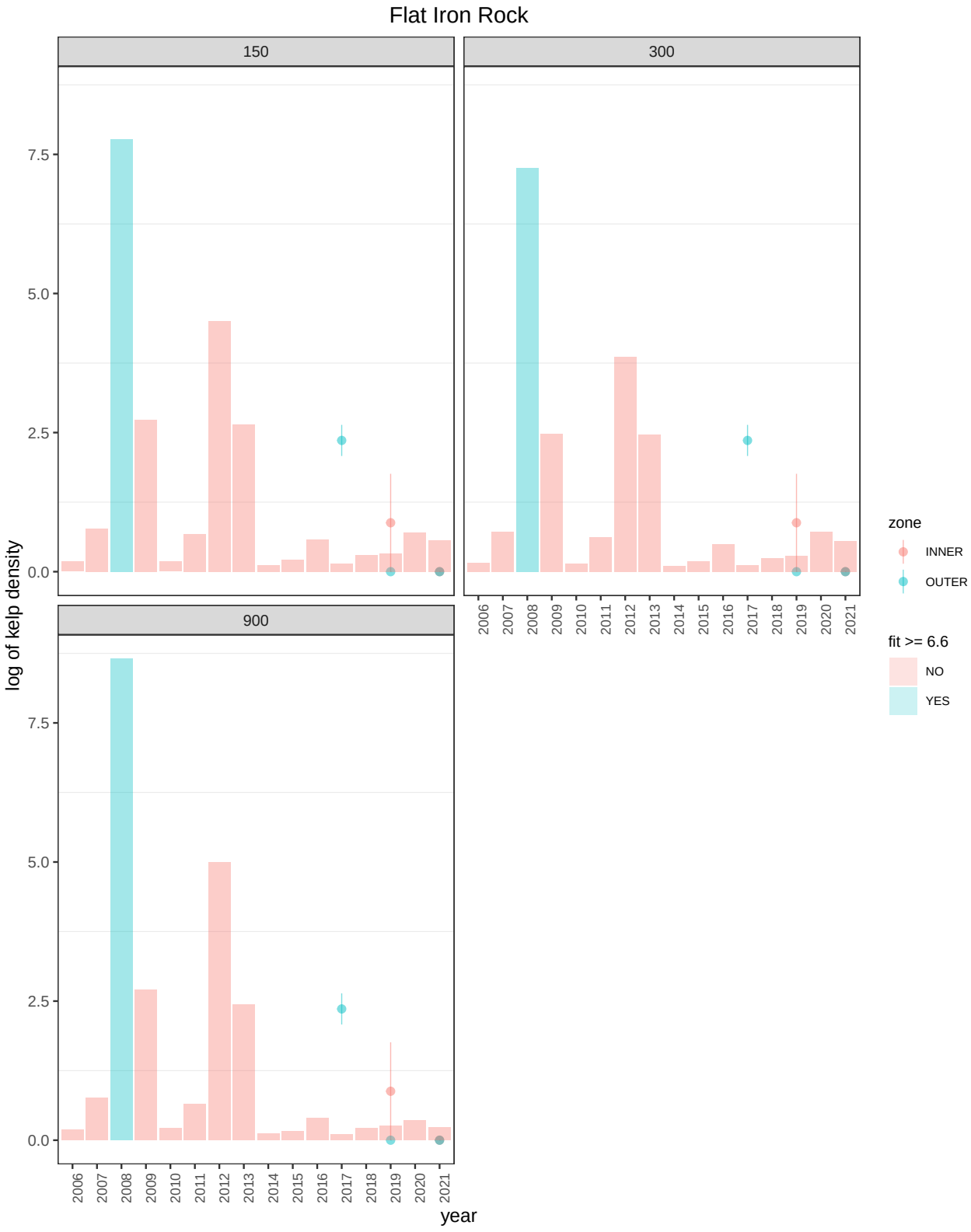
```
sites <- unique(kelp_data_300m$site_name)

kelp_longer_300m <- kelp_data_300m %>%
  dplyr::select(-c(longitude, latitude)) %>%
  pivot_longer(
    -c('site_name', 'year', 'fit'),
    names_to = c('.value', 'zone'),
    names_sep = '_'
    ) %>%
  mutate(resolution = as.factor(300))

kelp_longer_150m <- kelp_data_150m %>%
  dplyr::select(-c(longitude, latitude)) %>%
  pivot_longer(
    -c('site_name', 'year', 'fit'),
    names_to = c('.value', 'zone'),
    names_sep = '_'
    ) %>%
  mutate(resolution = as.factor(150))

kelp_longer_900m <- kelp_data_900m %>%
  dplyr::select(-c(longitude, latitude)) %>%
  pivot_longer(
    -c('site_name', 'year', 'fit'),
    names_to = c('.value', 'zone'),
    names_sep = '_'
    ) %>%
  mutate(resolution = as.factor(900))

# kelp_longer_600m <- kelp_data_600m %>%
```

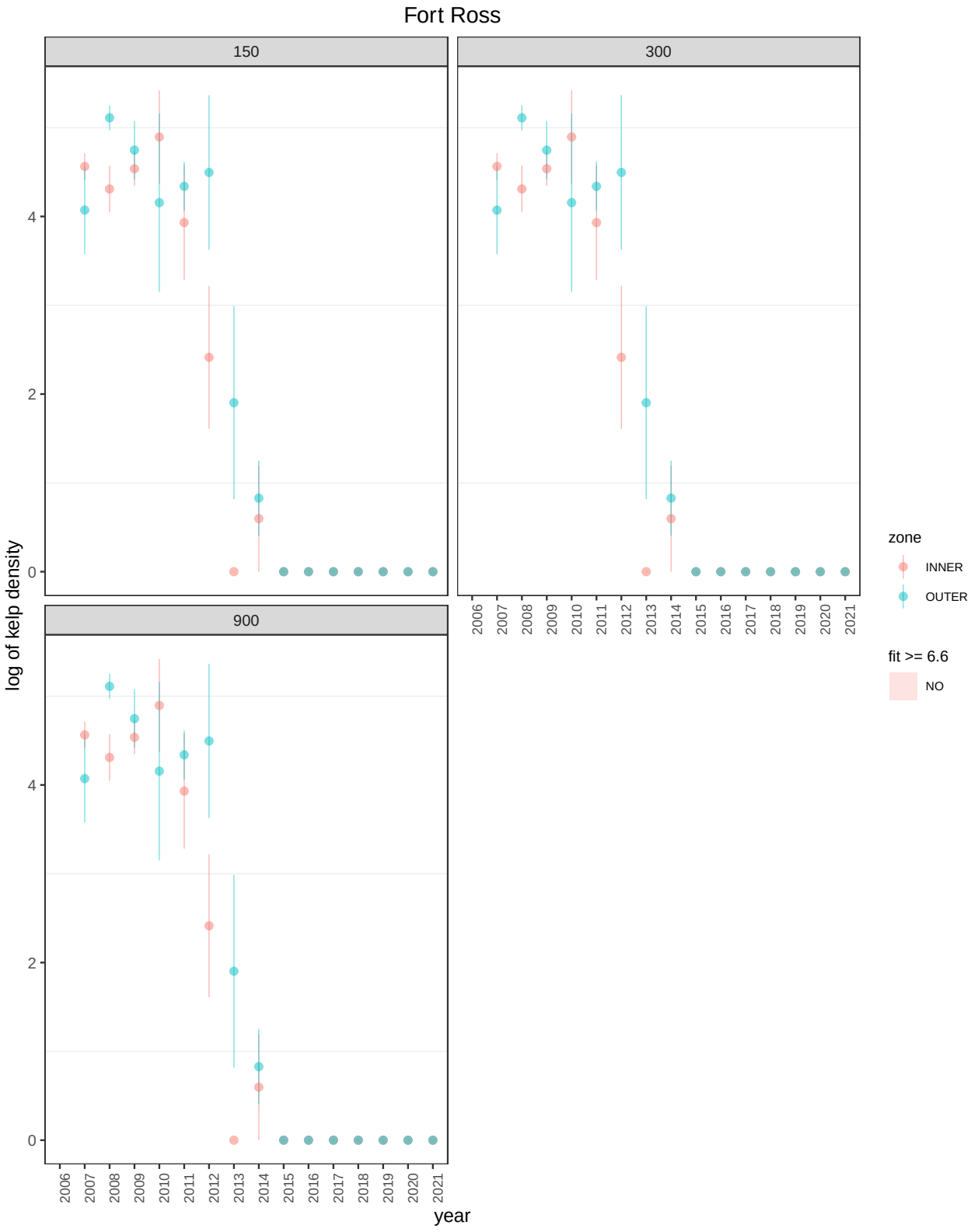
```
#   dplyr::select(-c(longitude, latitude)) %>%
#   pivot_longer(
#     -c('site_name', 'year', 'fit'),
#     names_to = c('.value', 'zone'),
#     names_sep = '_'
#     ) %>%
#   mutate(resolution = as.factor(600))

# kelp_longer <- rbind(kelp_longer_150m, kelp_longer_300m,
#                      kelp_longer_600m, kelp_longer_900m)

kelp_longer <- rbind(kelp_longer_150m, kelp_longer_300m,
                     kelp_longer_900m)

for (i in sites) {
  plot <- kelp_longer %>%
    filter(site_name == i) %>%
    ggplot() +
    geom_pointrange(aes(
      x = year, y = mean, group = zone, color = zone,
      ymin = mean - se, ymax = mean + se
      ), alpha = 0.5, size = 0.3) +
    geom_bar(aes(x = year, y = fit,
                 fill = ifelse(!is.na(fit) & fit >= 6.6, 'YES', 'NO')),
             stat = 'identity', position = 'dodge', alpha = 0.2) +
    facet_wrap(. ~ resolution, nrow = 2) +
    theme_bw() +
    theme(axis.text.x = element_text(angle = 90, size = 8),
          plot.title = element_text(hjust = 0.5),
          panel.grid.major = element_blank(),
          legend.title = element_text(size = 9),
          legend.text = element_text(size = 7)) +
    labs(y = 'log of kelp density', title = i, fill = 'fit >= 6.6')
  print(plot)
}
```

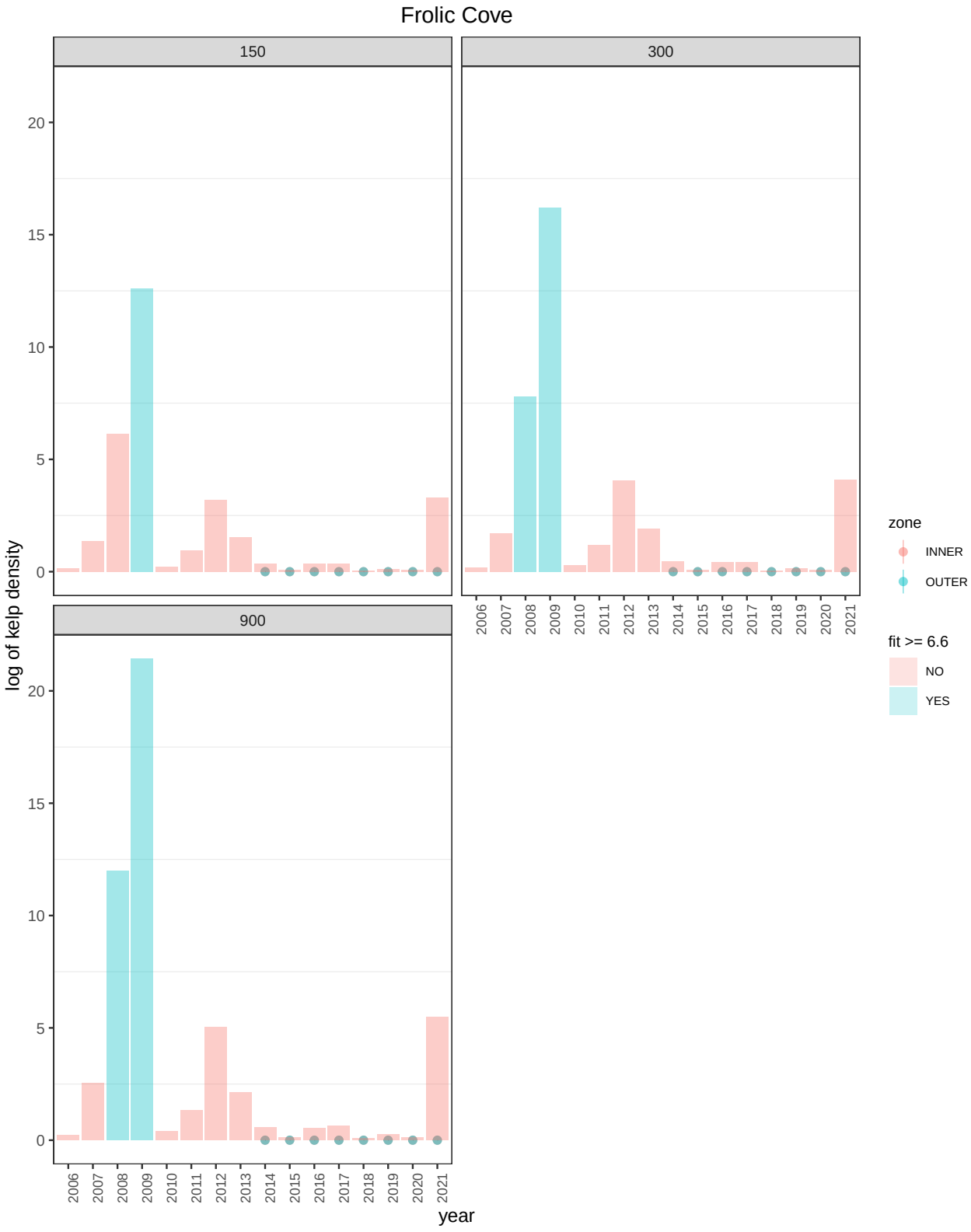

Caspar
150
300
900
log of kelp density
0
5
10
15
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

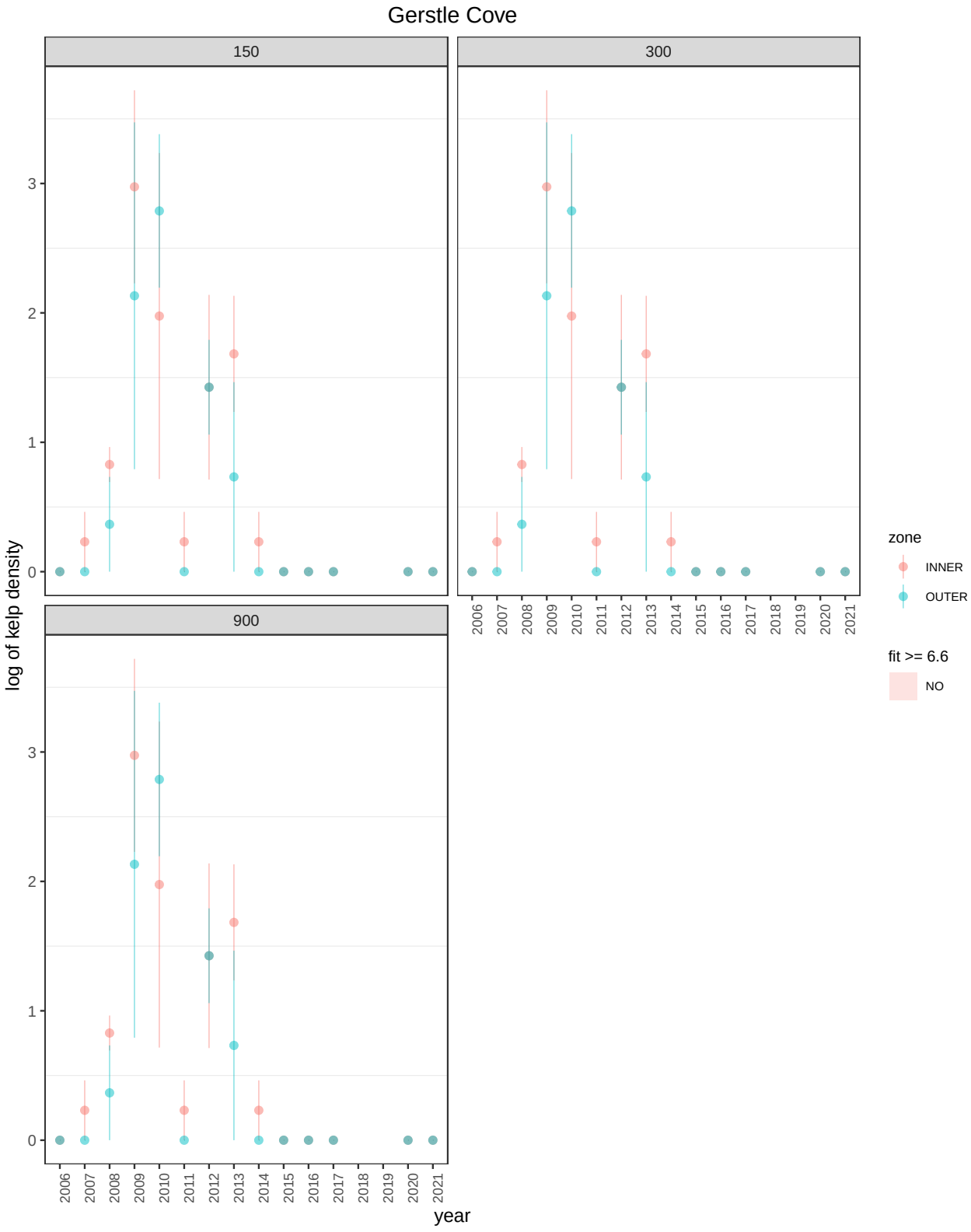

Caspar North
150
300
900
log of kelp density
0
5
10
15
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

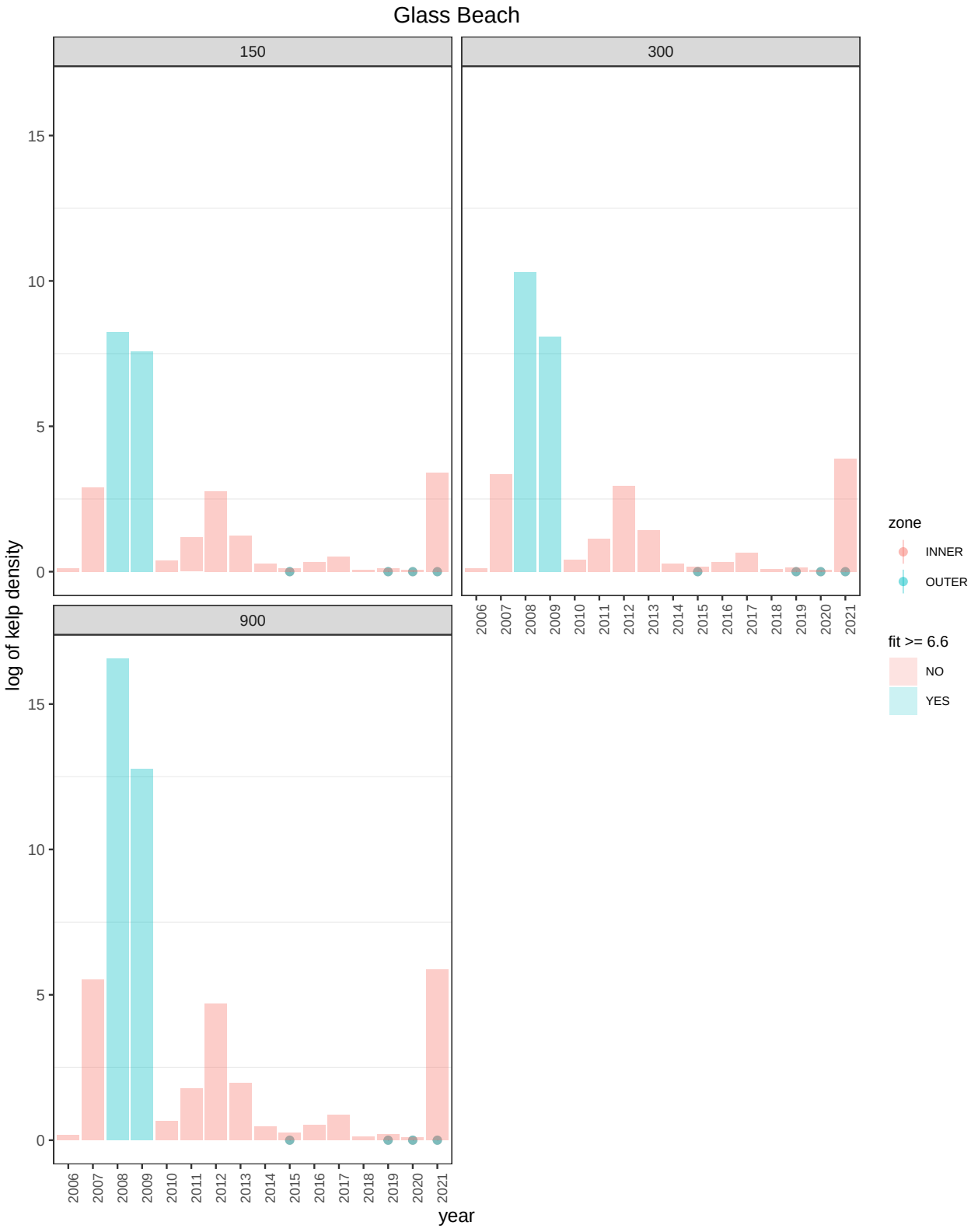

Dark Gulch
150
300
900
log of kelp density
30
20
10
0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

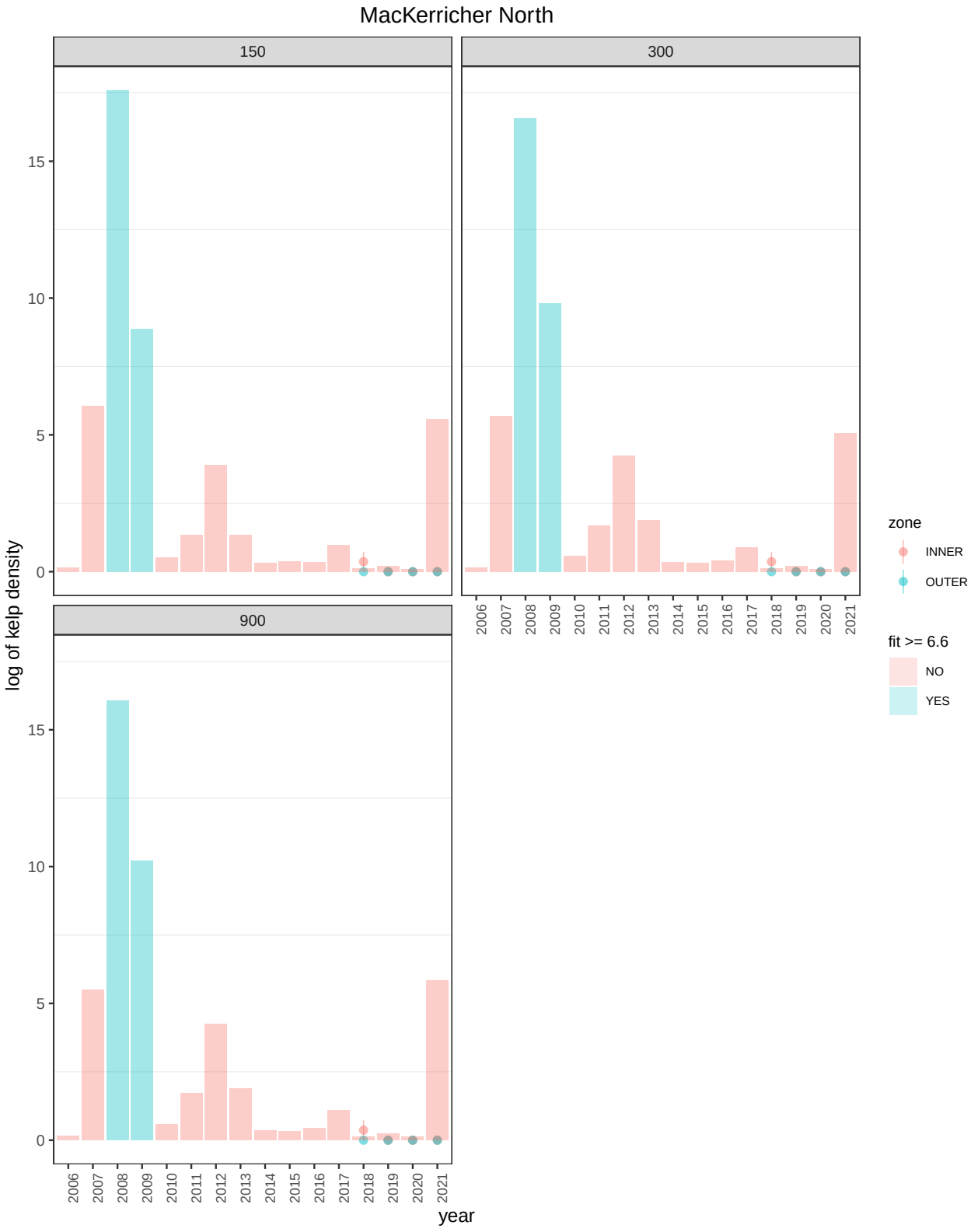

Flat Iron Rock
150
300
900
log of kelp density
7.5
5.0
2.5
0.0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

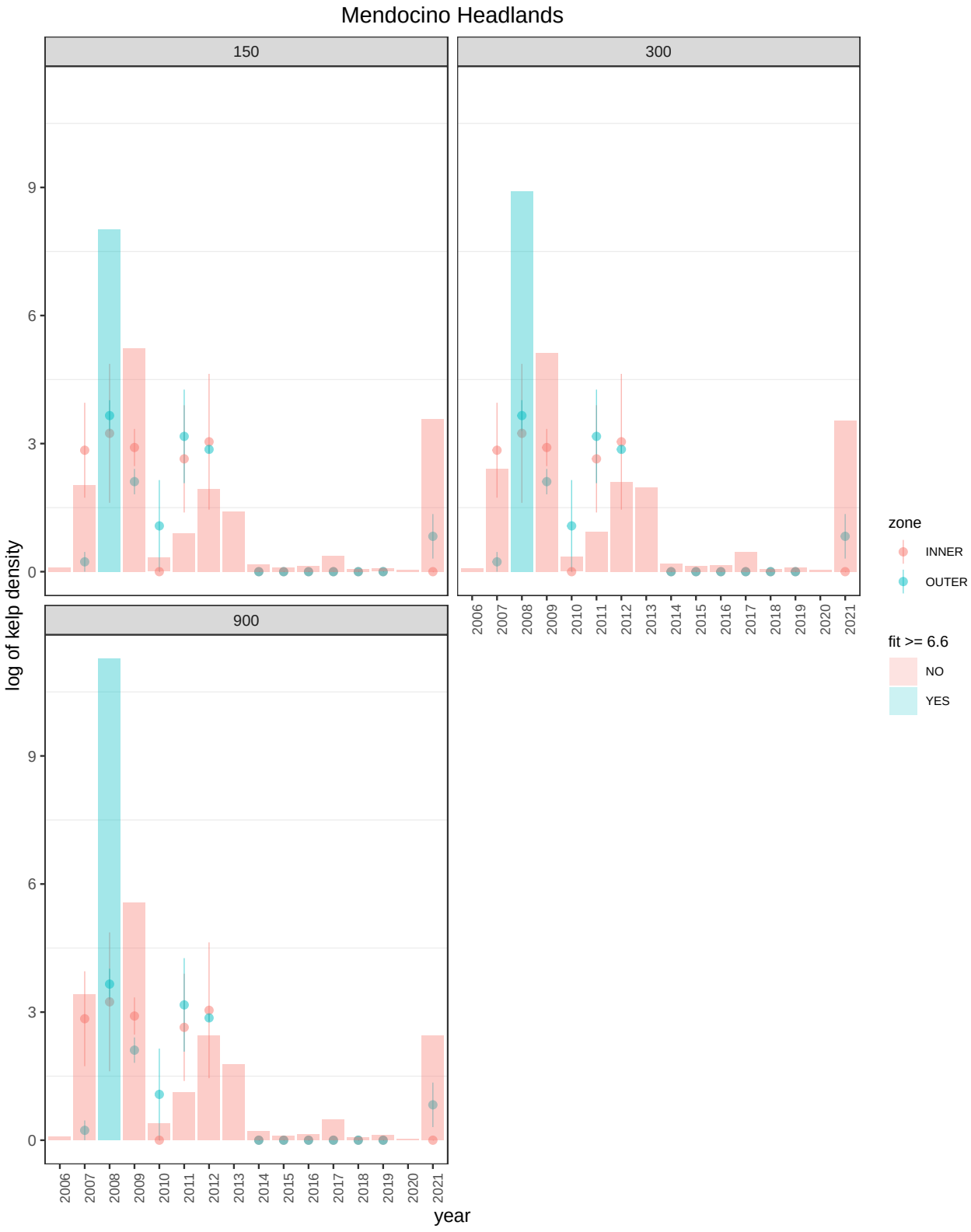

Fort Ross
150
300
900
log of kelp density
0
2
4
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO

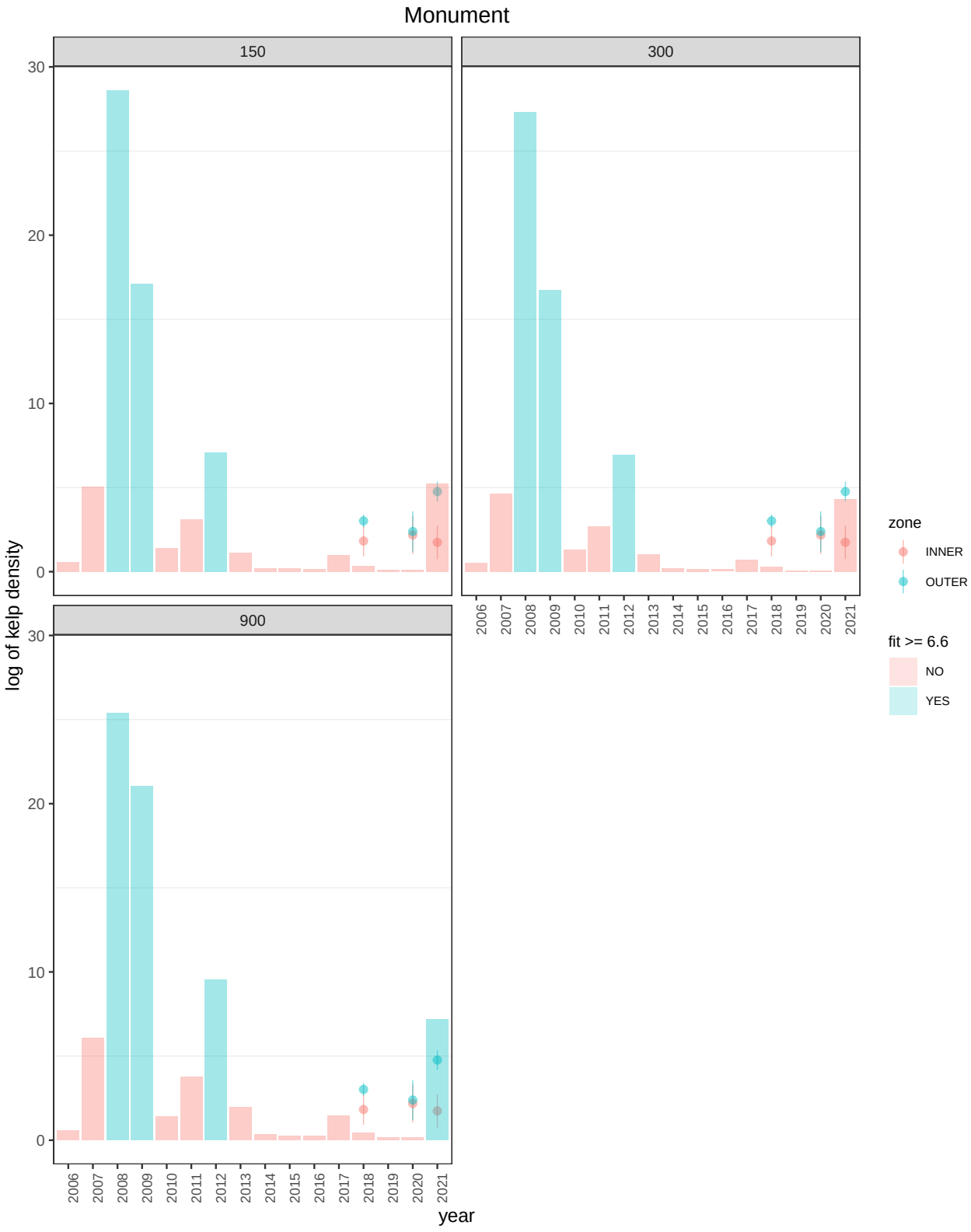

Frolic Cove
150
300
900
log of kelp density
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

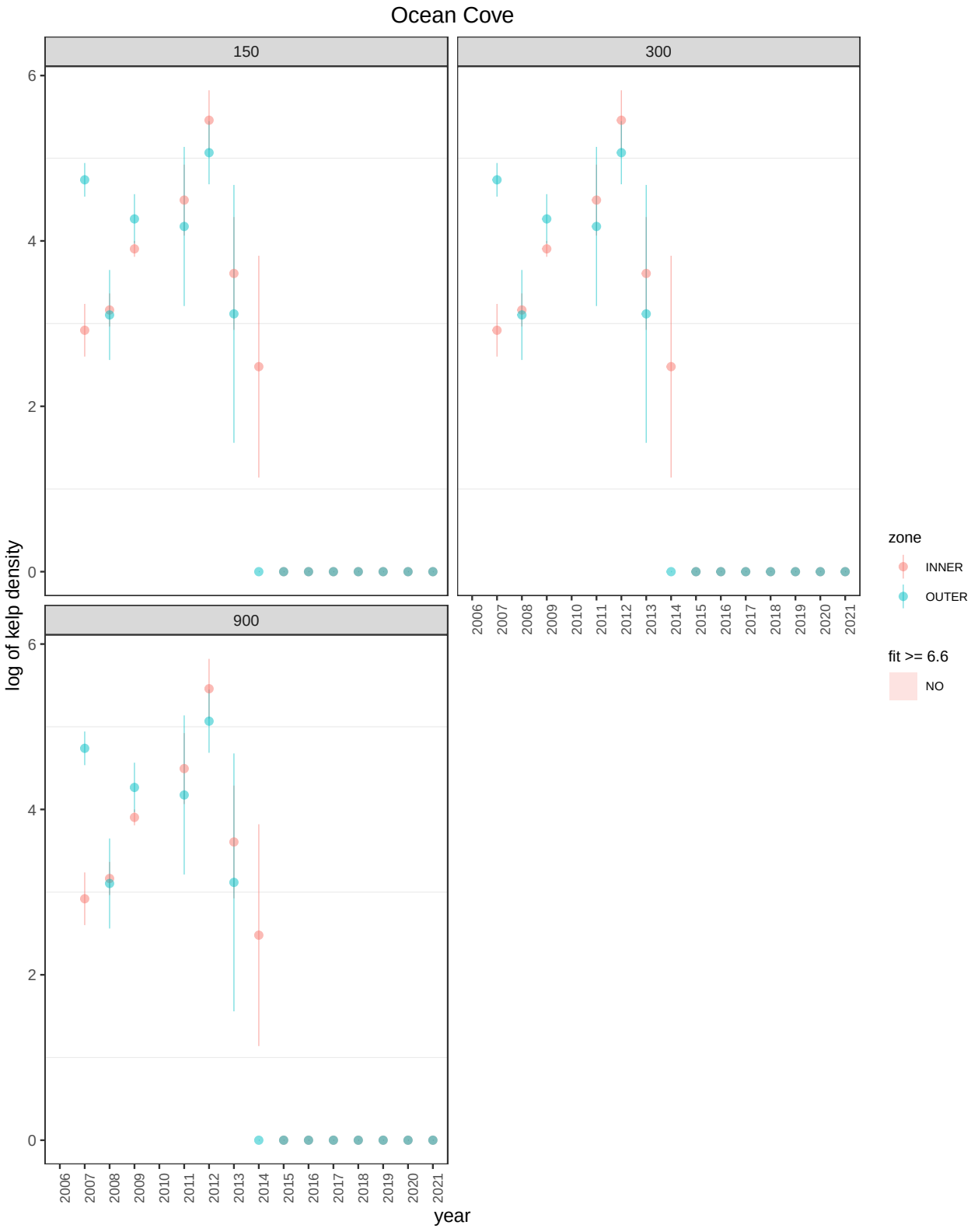

Gerstle Cove
150
300
900
log of kelp density
0
1
2
3
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO

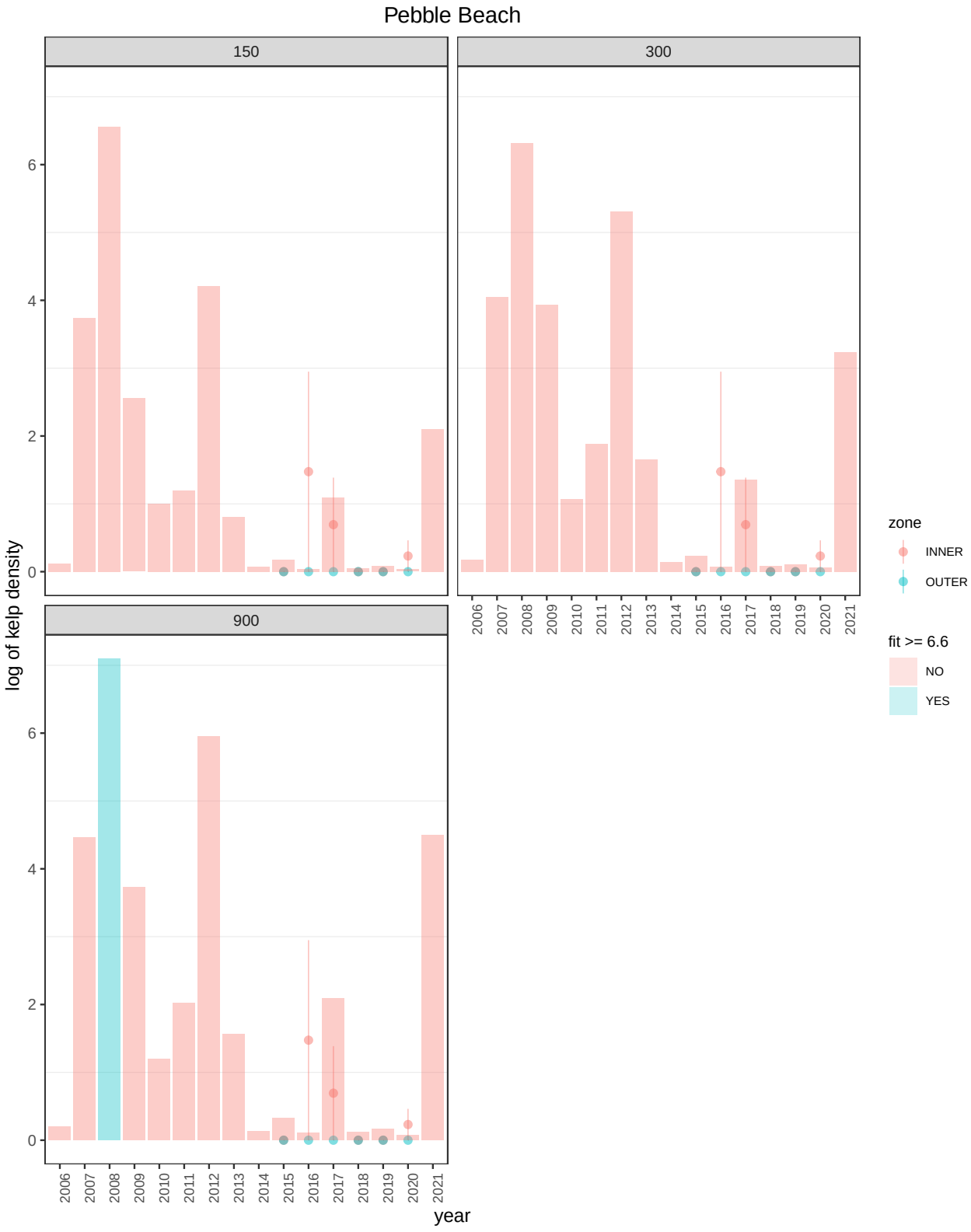

Glass Beach
150
300
900
log of kelp density
0
5
10
15
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

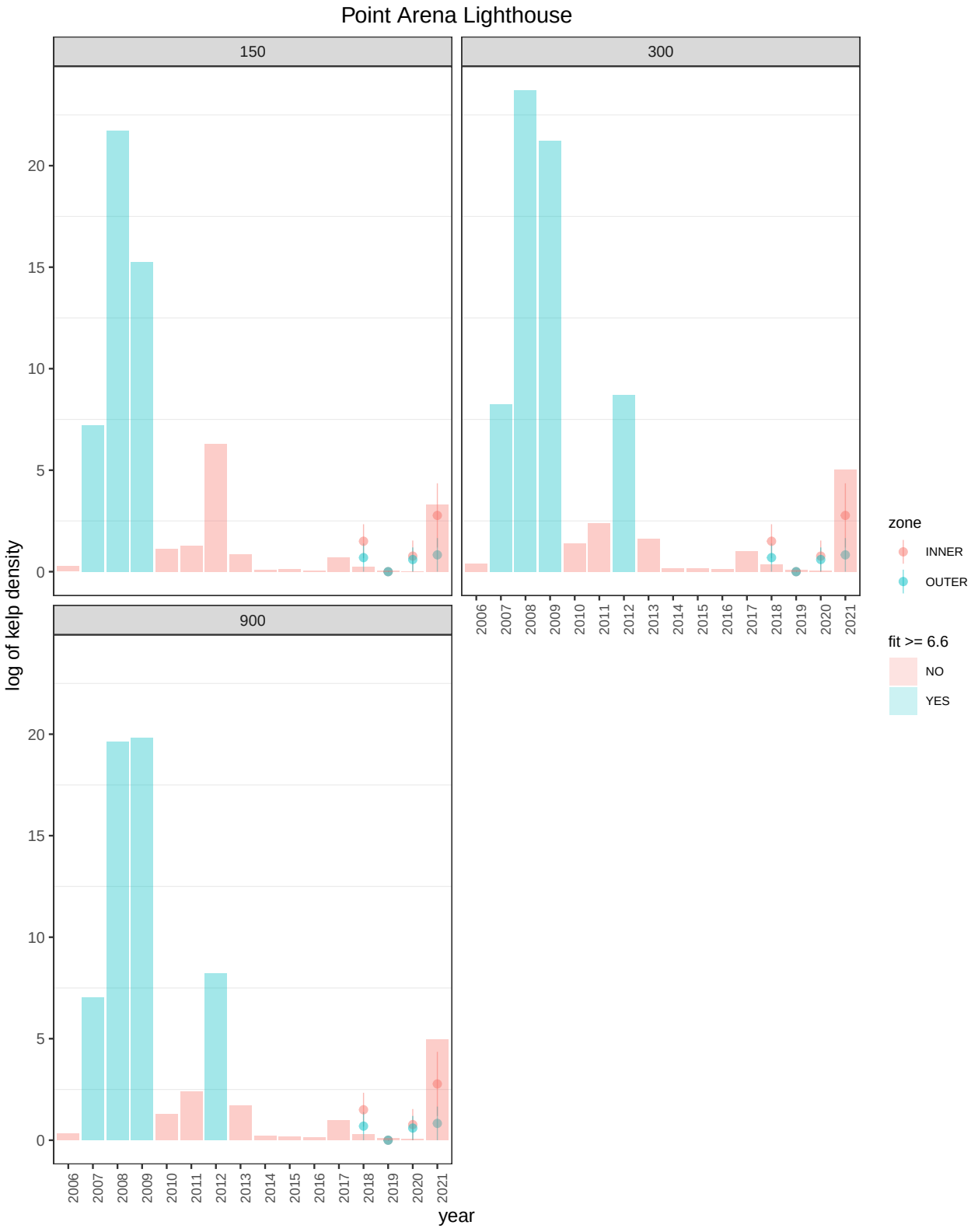

MacKerricher North
150
300
900
log of kelp density
0
5
10
15
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

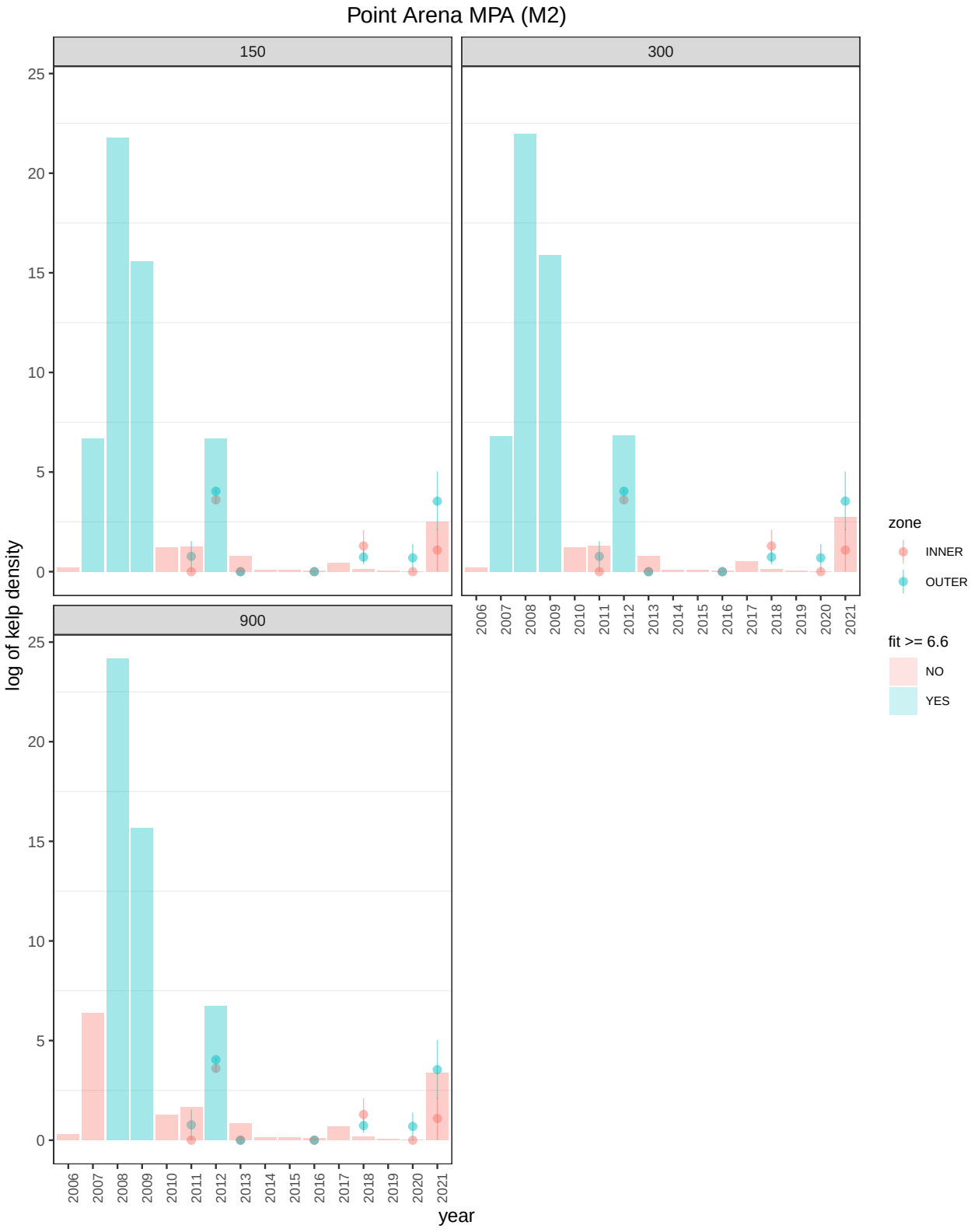

Mendocino Headlands
150
300
900
log of kelp density
0
3
6
9
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

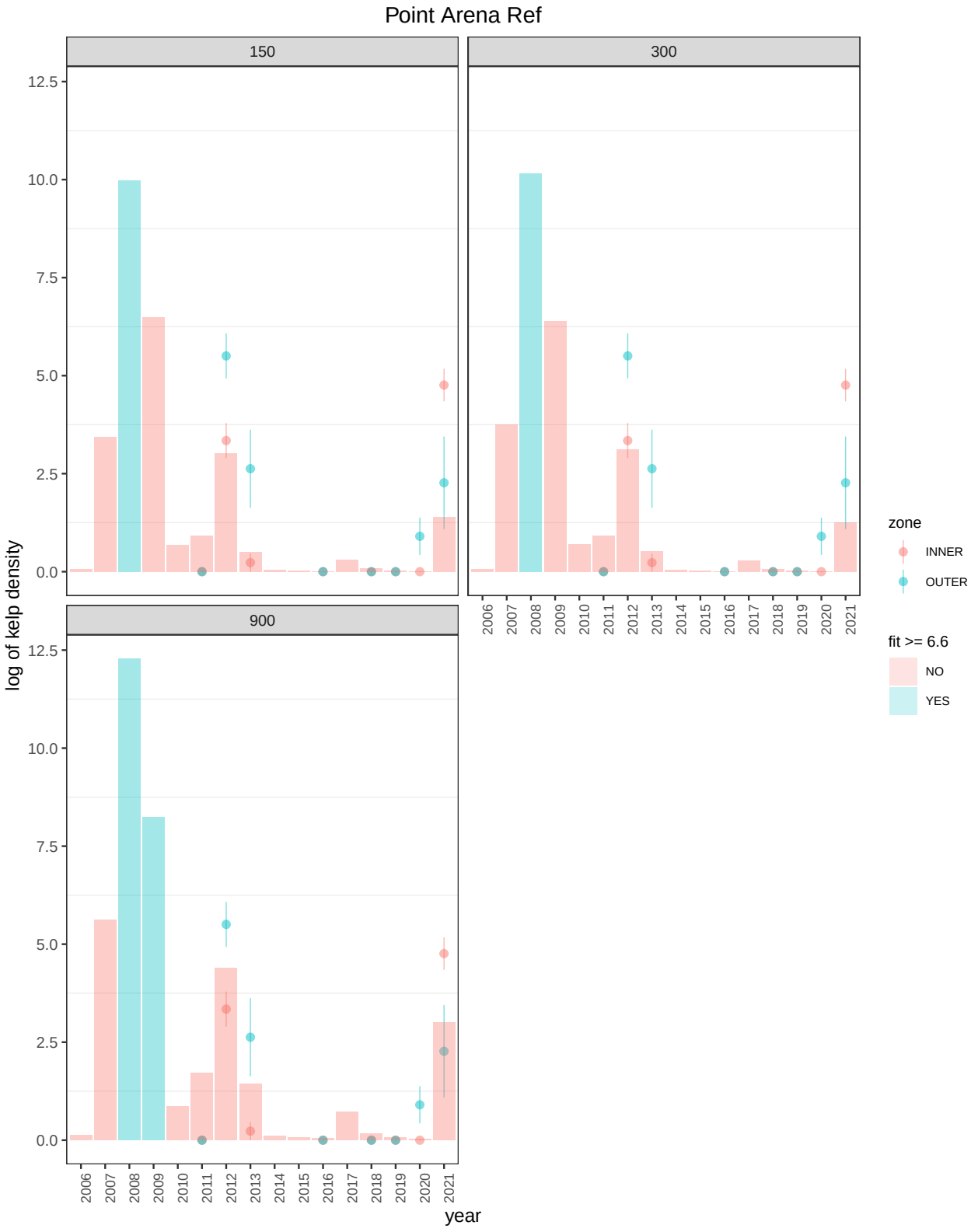

Monument
150
300
900
log of kelp density
30
20
10
0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

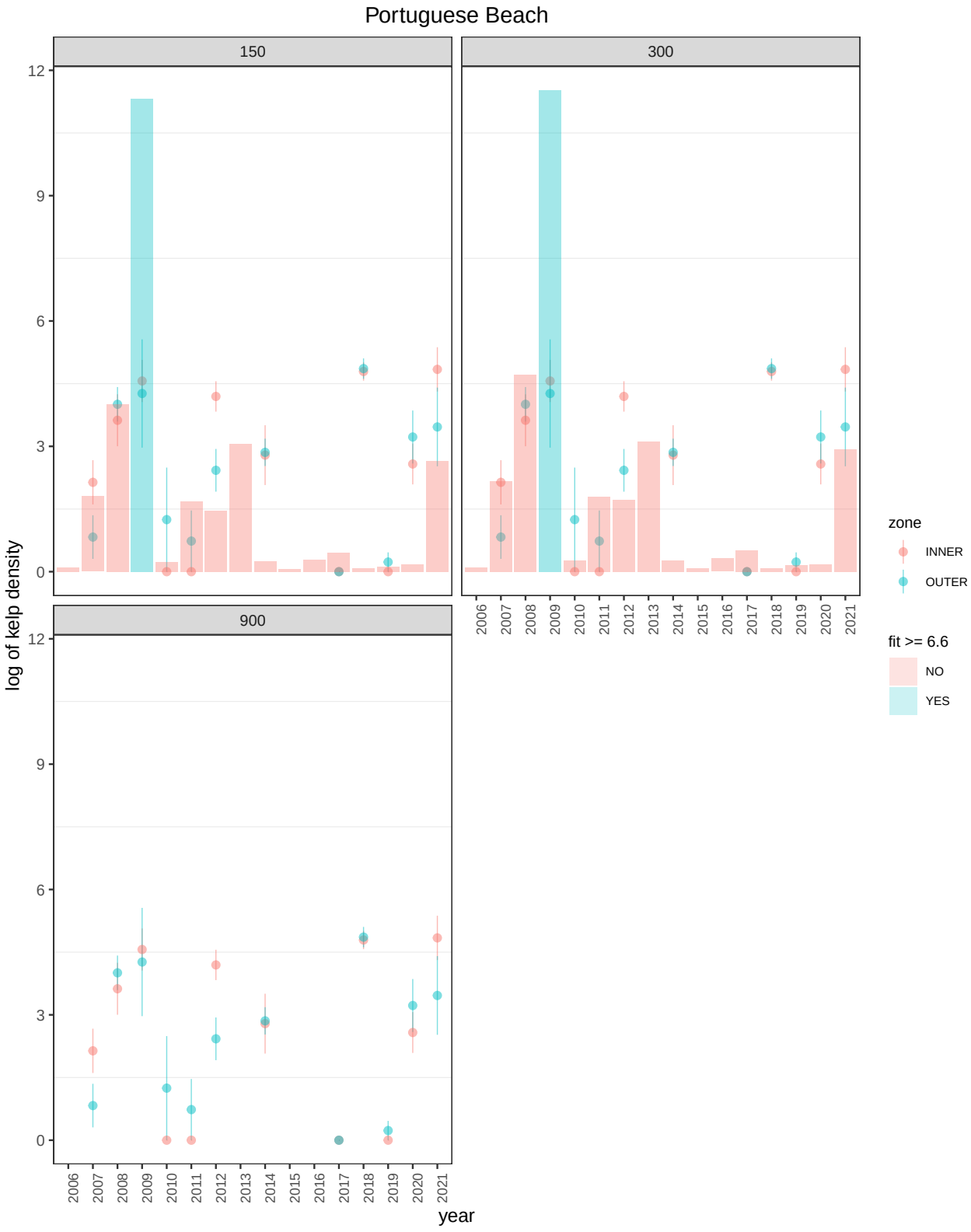

Ocean Cove
150
300
900
log of kelp density
year
zone
INNER
OUTER
fit >= 6.6
NO

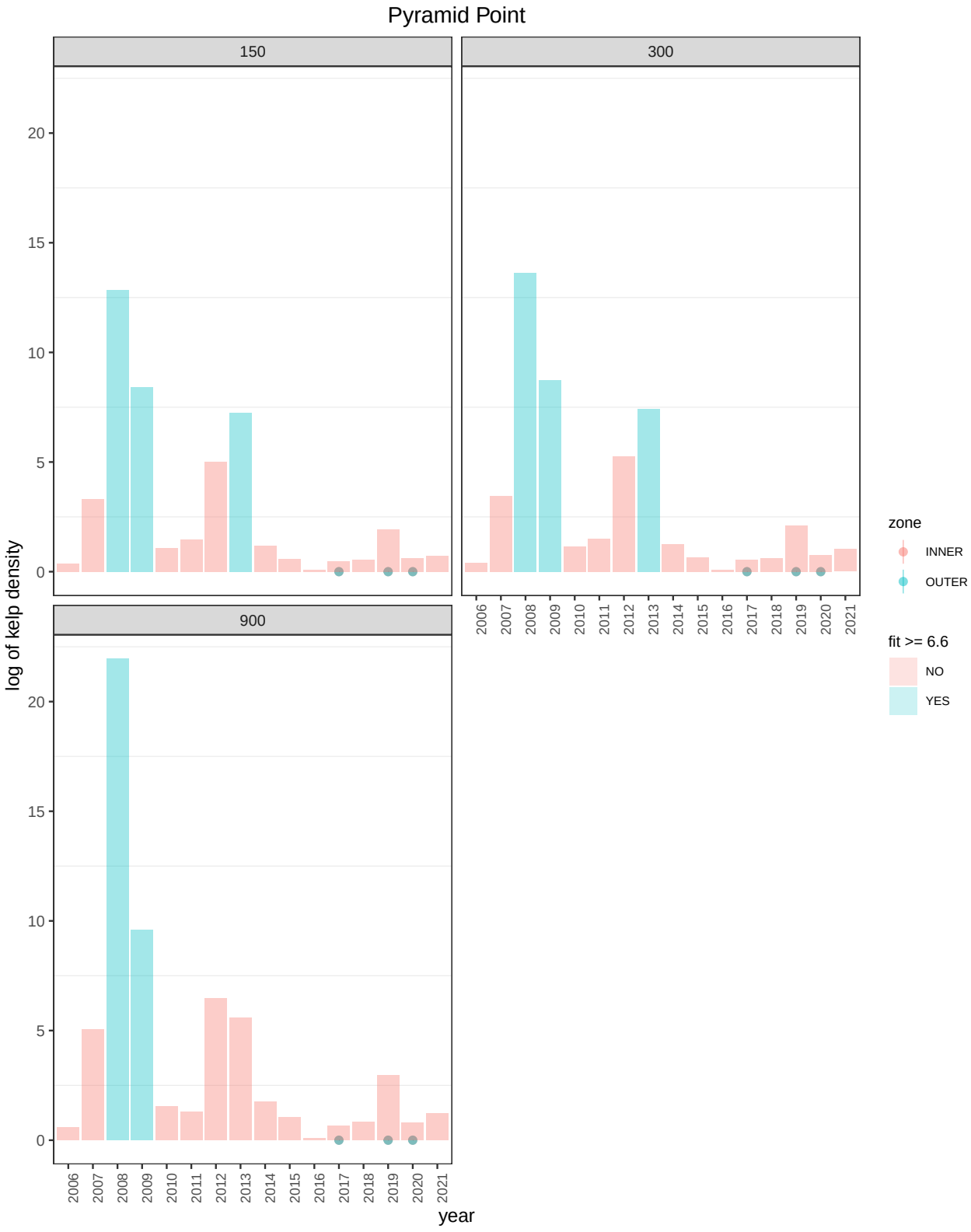

Pebble Beach
150
300
900
log of kelp density
0
2
4
6
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

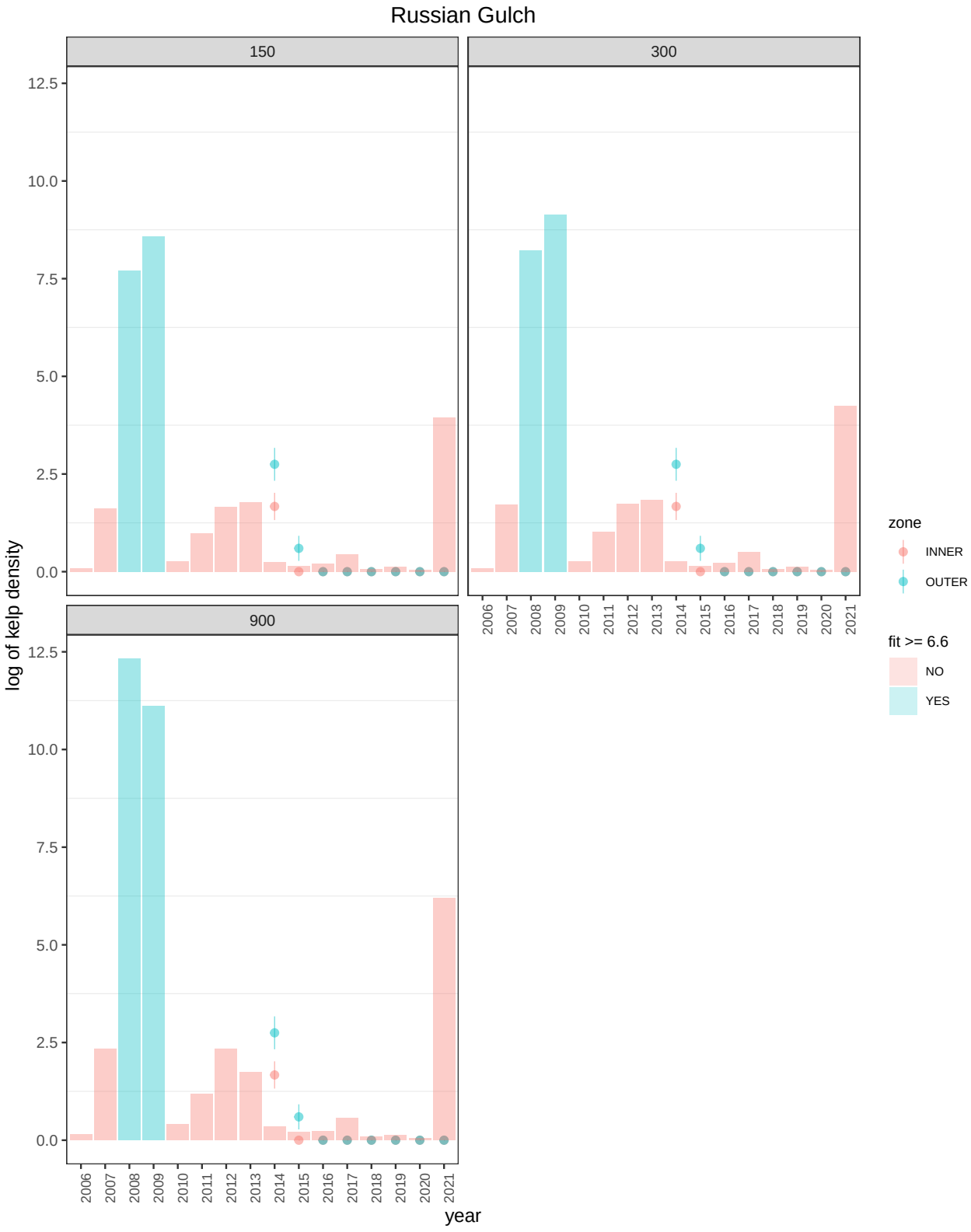

Point Arena Lighthouse
150
300
900
log of kelp density
0
5
10
15
20
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

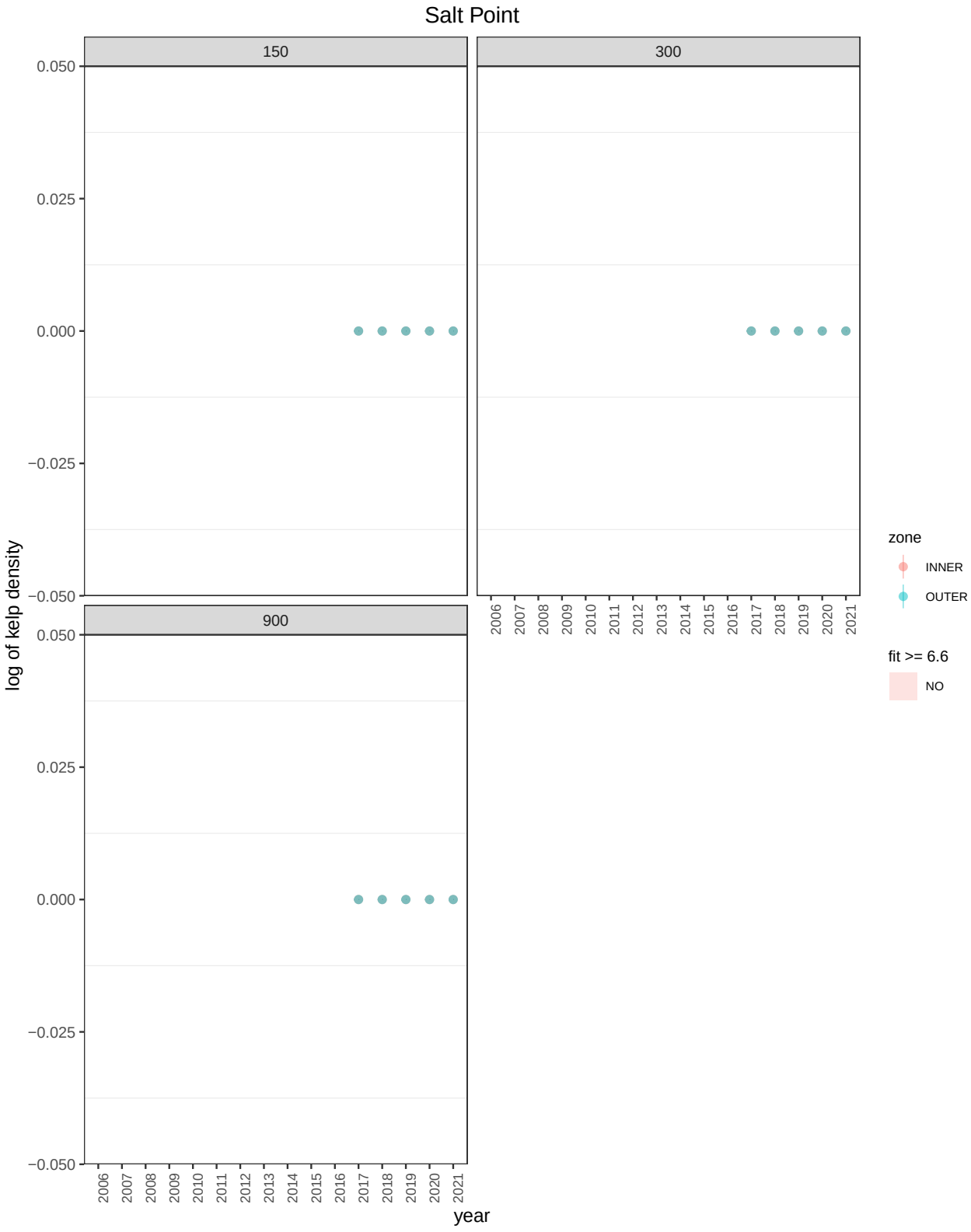

Point Arena MPA (M2)
150
300
900
log of kelp density
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

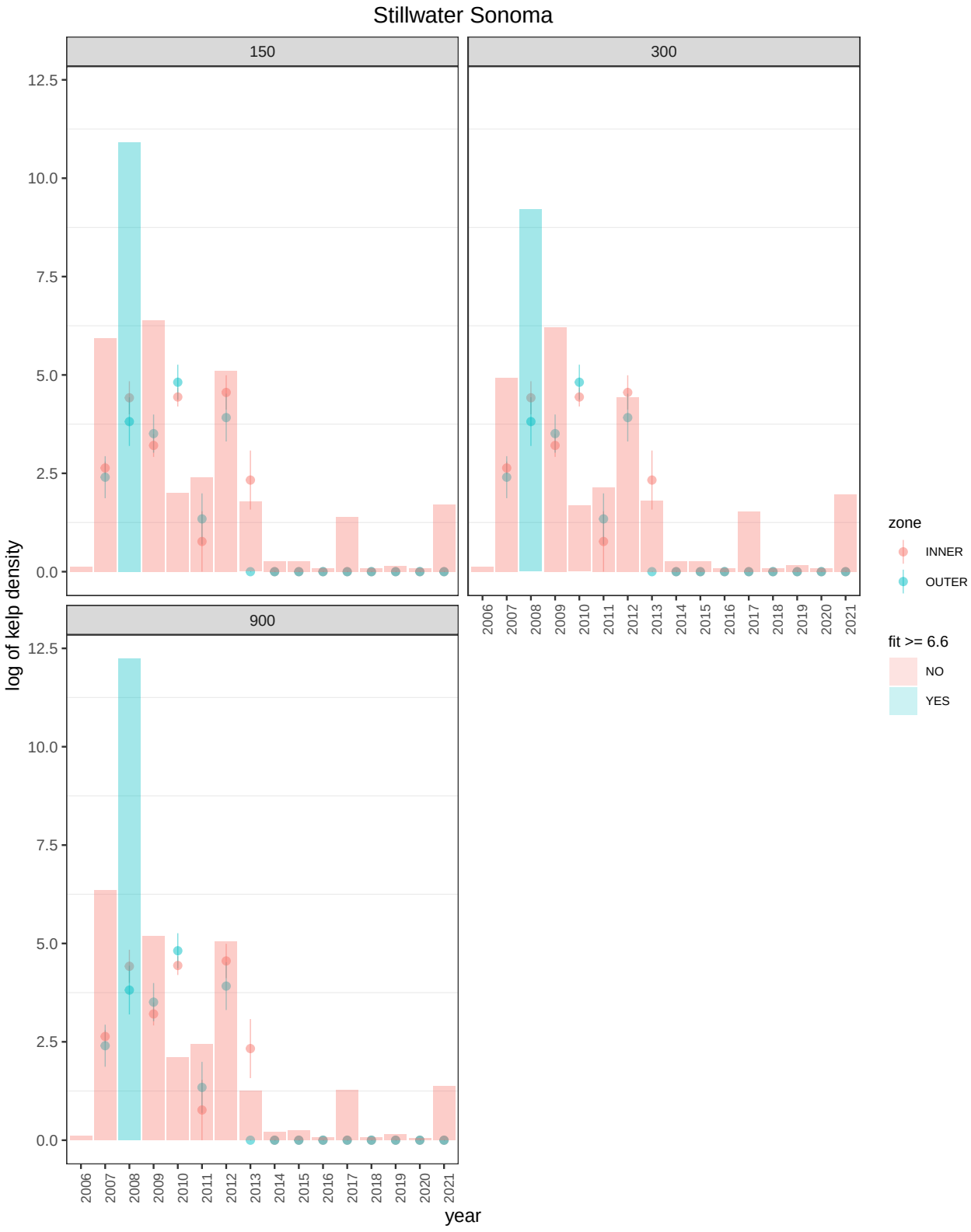

Point Arena Ref
150
300
900
log of kelp density
12.5
10.0
7.5
5.0
2.5
0.0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

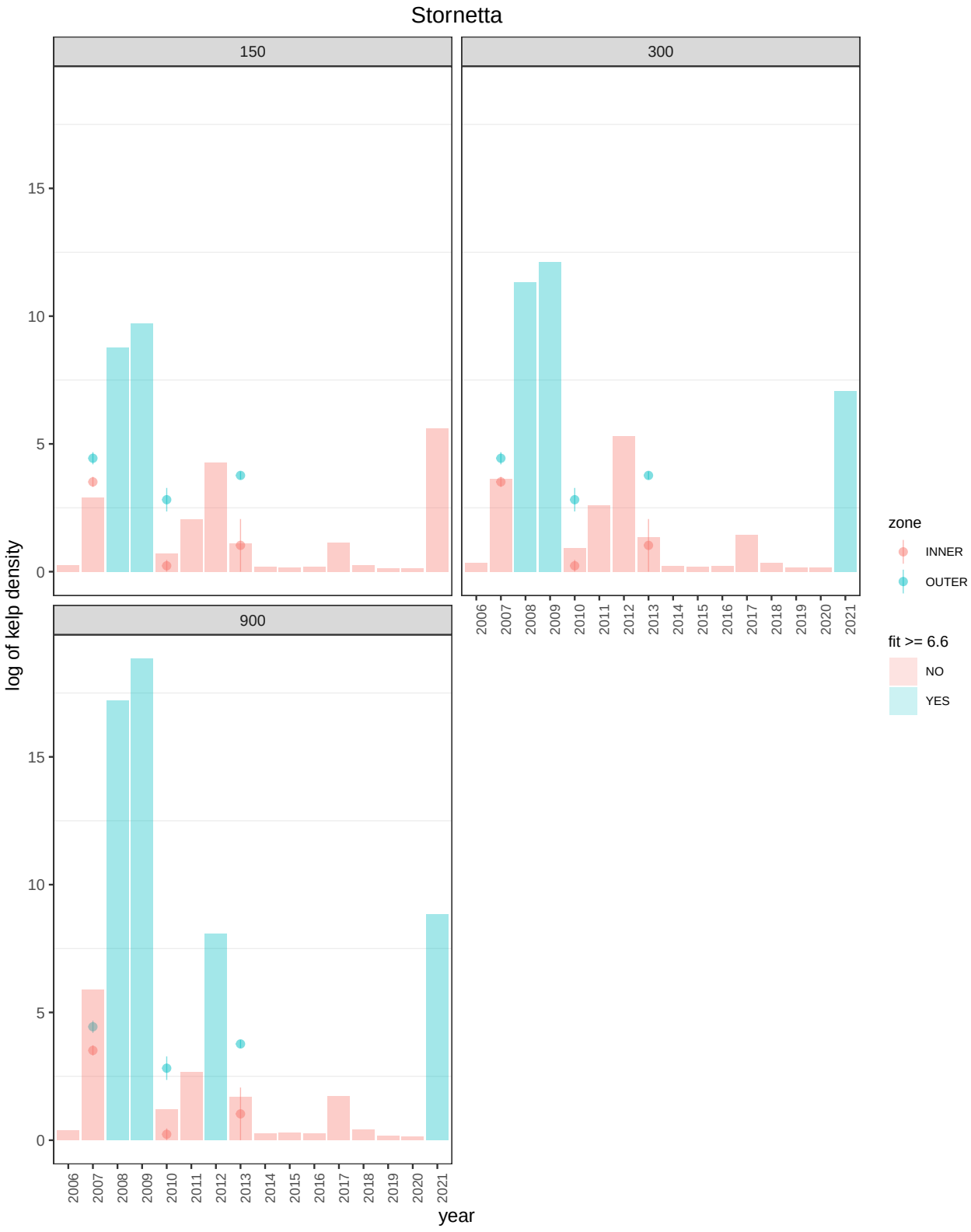

Portuguese Beach
150
300
900
log of kelp density
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

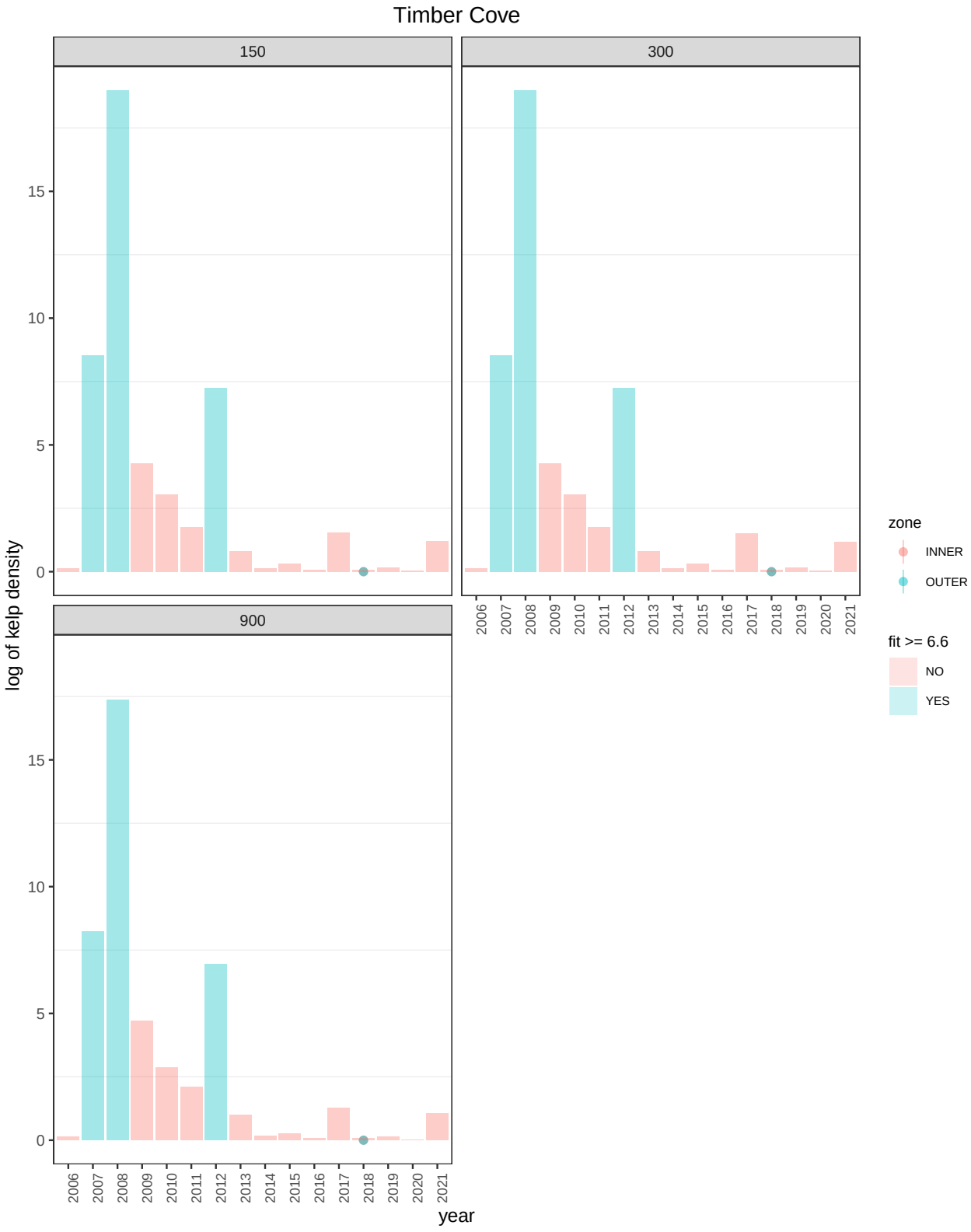

Pyramid Point
150
300
900
log of kelp density
0
5
10
15
20
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

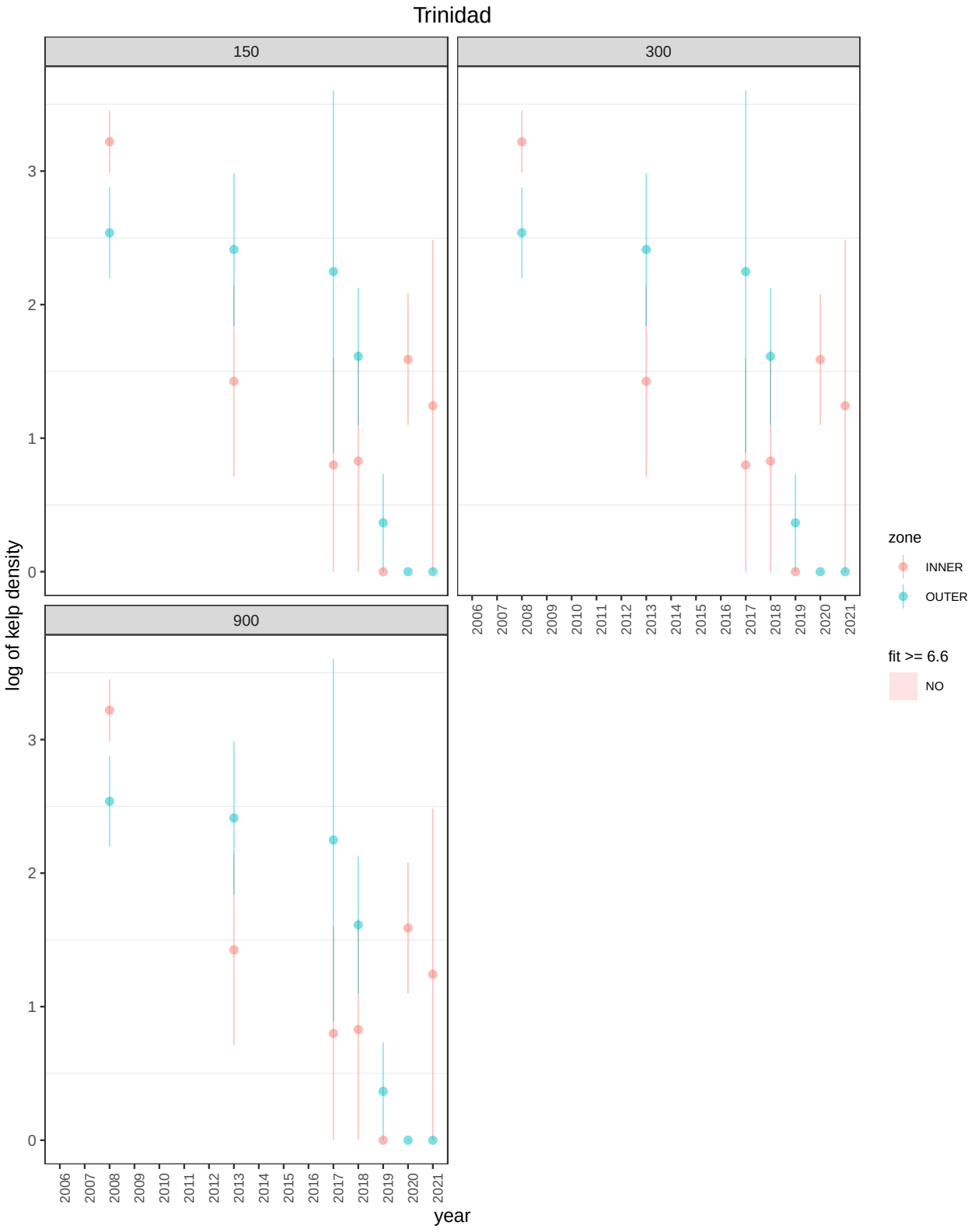

Russian Gulch
150
300
900
log of kelp density
12.5
10.0
7.5
5.0
2.5
0.0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES

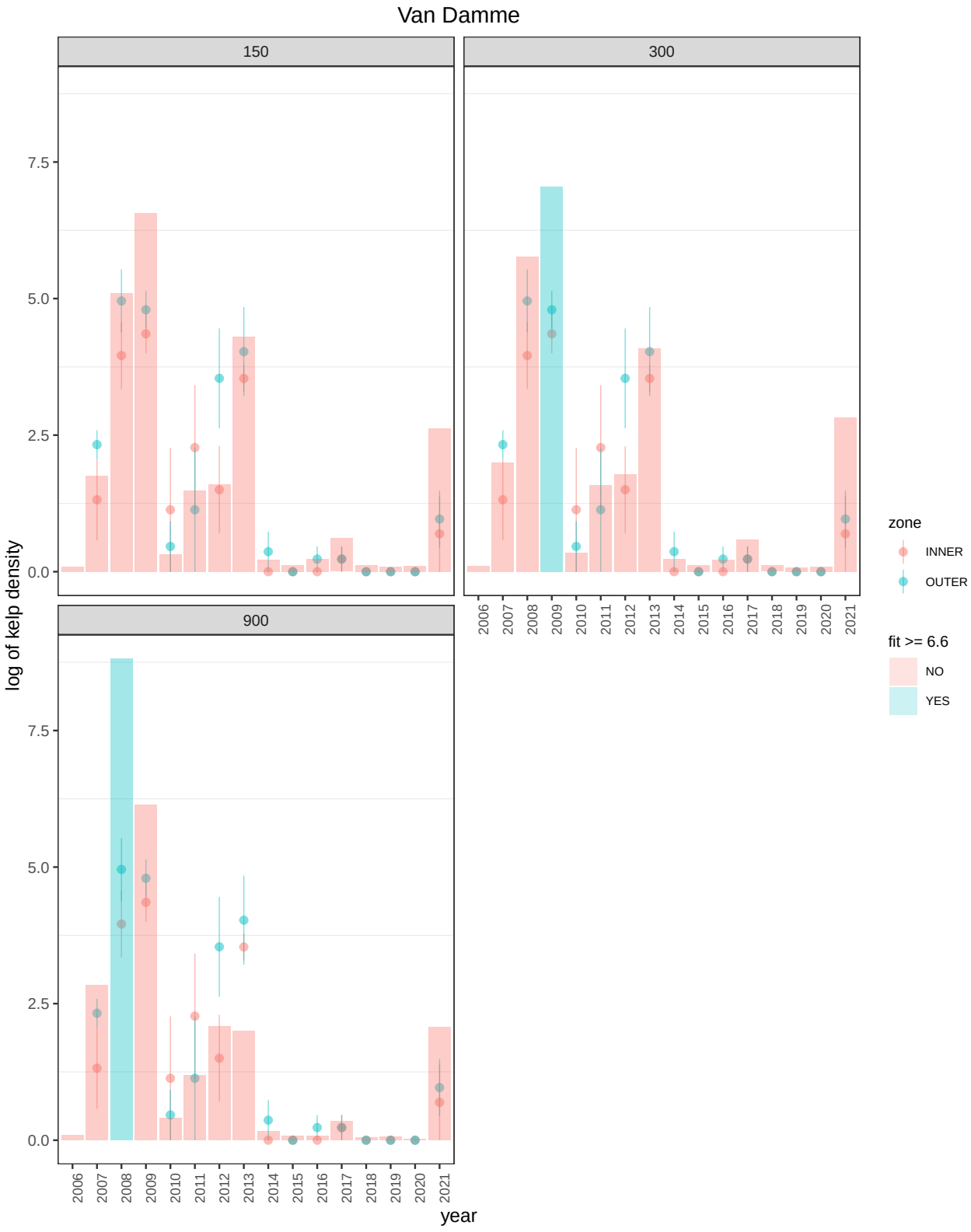

Salt Point
150
300
900
0.050
0.025
0.000
−0.025
−0.050
log of kelp density
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO


Stillwater Sonoma
150
300
900
log of kelp density
12.5
10.0
7.5
5.0
2.5
0.0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES


Stornetta
150
300
900
log of kelp density
0
5
10
15
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES


Timber Cove
150
300
900
log of kelp density
0
5
10
15
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES


Trinidad
150
300
900
log of kelp density
year
zone
INNER
OUTER
fit >= 6.6
NO


Van Damme
150
300
900
log of kelp density
7.5
5.0
2.5
0.0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
year
zone
INNER
OUTER
fit >= 6.6
NO
YES