Analysis - Night setting around Tristan da Cunha

Cian Luck

09 Aug 2021

Table of Contents

Background

The critically endangered Tristan albatross (*Diomedea dabbenena*) is endemic to Gough Island, and has a decreasing population of fewer than 5000 individuals. The two major threats facing the Tristan albatross are predation by invasive mice and bycatch in longline fisheries.

To mitigate the risk of bycatch, ICCAT recommend that vessels adopt at least two of the following mitigation measures:

1. Night setting
2. Bird-scaring lines (a.k.a. Tori lines)
3. Line weighting

Night setting involves setting the lines in darkness, between nautical dusk and dawn, however it is only effective if sets are completed in darkness, and should be completed an hour before nautical dawn as best pratice. Previous analysis by GFW, in collaboration with BirdLife International showed that in 2017, 85% of sets were conducted with more than a two hour overlap with daylight, meaning only 15% of vessels set their lines fully at night. This was much lower than what was observed and self-reported.

The Research Team at Global Fishing Watch have developed an improved fishing detection algorithm for longline fishing vessels which can differentiate between setting and hauling of lines. In this analysis, we applied a newly developed dataset of longline fishing activity to determine when and where longlines have been set near Tristan da Cunha and assess if and to what extent night-setting is being used to mitigate Tristan albatross bycatch.

After applying the model, all longline setting events were grouped into categories as follows:

1. **Day Sets**: Sets that occurred entirely between nautical dawn and nautical dusk (when the center of the sun is 12º below the horizon)
2. **Night Sets**: Sets that occurred entirely between nautical dusk and nautical dawn.
3. **Sets Before Nautical Dusk**: Sets that began before nautical dusk and extended into the night
4. **Sets Over Nautical Dawn**: Sets that began before nautical dawn and extended into the day

We also defined several variants of sets (3) and (4). These were:

5. **Sets Over Dawn <2h**: Sets over nautical dawn, but extending less than 2 hours into day.
6. **Sets Over Dawn >2h**: Sets over nautical dawn, but extending more than 2 hours into day.
7. **Sets Over Dusk <2h**: Sets over nautical dusk, but starting less than 2 hours before dusk.
8. **Sets Over Dusk >2h**: Sets over nautical dawn, but starting more than 2 hours before dusk."

In this dataset, the column `cat` uses categories `[1,2,3,4]`, `cat2` uses categories `[1,2,5,6,7,8]`. The model is estimated to predict 90% of start and end times of setting events to within 2 hours of actual start and end times. For the purposes of this analysis, we simplified these categories to `night` and `day`, where `night` includes all sets carried out fully at night and sets that were carried out mostly at night but extended less than 2 hours into day (i.e. categories 1, 5, 7), and `day` includes all other sets (i.e. categories 1, 6, 8).

Setup

Load packages:

```
library(tidyverse)  # data manipulation and plotting

## Warning: package 'stringr' was built under R version 4.1.1

library(bigrquery)  # querying data through BigQuery
library(DBI)        # database interface
library(fishwatchr) # internal R package developed by Global Fishing Watch for common in-house analyses and functions
library(glue)       # used to format SQL queries in R
library(lubridate)  # format date time objects
library(here)       # useful package for specifying file locations
library(sf)         # simple features - used for spatial analysis
library(extrafont)  # load extra fonts for plotting
library(wesanderson) # additional colour palettes for plotting
```

Establish connection to Big Query project:

```
con <- DBI::dbConnect(drv = bigrquery::bigquery(),
                      project = "world-fishing-827",
                      use_legacy_sql = FALSE)
```

Read in shapefile of Tristan da Cunha:

```
# Shapefiles of Tristan da Cunha and Gough Island - sourced from OpenStreetMap
tdc_sf <- st_read("geodata/tdc_osm/tristan_da_cunha_archipelago_osm.shp")

st_crs(tdc_sf) <- 4326

# filter the shapefile to only include Gough Island - needed for spatial analysis
later
gough_sf <- tdc_sf %>% filter(FID == 138822)
```

This query pulled all setting events in the South Atlantic, between Jan. 1 2019 and June 30 2021:

```
query_1 <- readr::read_file(str_c("queries", "q_tdc_ns_sets_south_atlantic.sql",
sep="/"))
```

Run the query:

```
ll_ns <- fishwatchr::gfw_query(query = query_1,
                               run_query = TRUE,
                               con = con)$data
```

Alternatively, the queried data can be loaded locally here:

```
# ll_ns stands for longline night setting
ll_ns <- read_rds("data_production/data/night_setting_locations_south_atlantic.rds")
```

As described above, we created a third night-setting category (`cat3`) that grouped all sets carried out fully at night and all those which overlapped with day by less than 2 hours as `night`, and all other sets as day.

```
ll_ns <- ll_ns %>%
  mutate(cat3 = if_else(cat2 == 2 | cat2 == 5 | cat2 == 7, "night", "day"))
```

Night-setting in the South Atlantic

To create a map of night-setting across the South Atlantic, first we aggregated the setting events to a 0.1° x 0.1° grid, and at each location we calculated the ratio of longline sets that were carried out at night, where 1 means all sets were carried out at night and 0 means all sets were carried out during the day. Because each setting event had a start and end location, we used the latitude and longitude for each set. Each set had a start and end timestamp, and for this analysis we used the start timestamp as the date for each set.

```
ll_ns_gridded <- ll_ns %>%
  # select only required columns
  dplyr::select(start_lon, set_id, start_lat, end_lon, end_lat, start_time, cat2) %>%
```

```
  rowwise() %>%
  mutate(mean_lon = mean(c(start_lon, end_lon), na.rm = TRUE), # average longitude
and latitude between start and end locations
         mean_lat = mean(c(start_lat, end_lat), na.rm = TRUE),
         lon_bin = floor(mean_lon*10)/10, # set the grid resolution to one 10th of a
degree
         lat_bin = floor(mean_lat*10)/10,
         date = lubridate::date(start_time), # use the start time to assign a date to
each set
         cat3 = if_else(cat2 == 2 | cat2 == 5 | cat2 == 7, 1, 0) # recalculate cat3
as a numeric variable (1 = night, 0 = day)
         ) %>%
  group_by(lat_bin, lon_bin) %>%
  summarise(n_sets = n_distinct(set_id), # number of sets
            ratio_ns = mean(cat3, na.rm = TRUE), # ratio of sets at night:day
            .groups = "keep"
            )
```

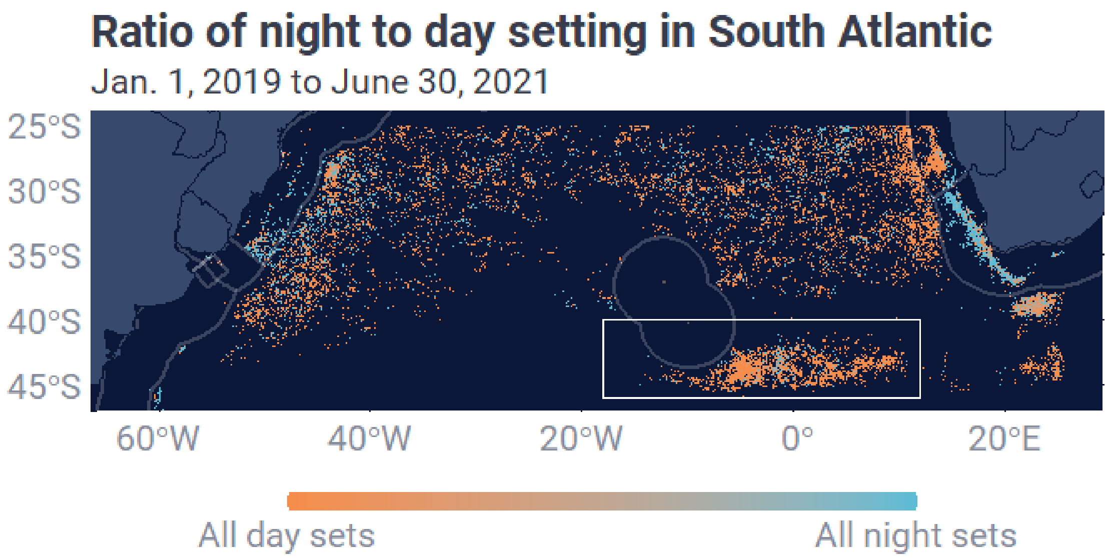
Map of ratio of night-setting to day-setting at each grid cell:

```
# Define a polygon of the area of interest to the southeast of Gough Island as a
dataframe
aoi_poly <- data.frame(
  id = 'aoi',
  x = c(-18, 12, 12, -18, -18),
  y = c(-40, -40, -46, -46, -40)
)

# set the bounding area for the map
bounding_1 <- fishwatchr::transform_box(xlim = c(-62, 25),
                                        ylim = c(-25, -46),
                                        output_crs = "+proj=longlat +ellps=WGS84
+datum=WGS84 +no_defs")

# make map with ggplot2
ll_ns_gridded %>%
  # only include data south of 25°S, where night setting is required
  filter(lat_bin < -25) %>%
ggplot() +
  geom_raster(aes(x = lon_bin,
                  y = lat_bin,
                  fill = ratio_ns)) +
  fishwatchr::geom_gfw_eez(lwd = 1) +
  fishwatchr::geom_gfw_land() +
  geom_sf(data = tdc_sf, fill = gfw_palette("map_country_dark")[1]) +
  geom_path(data = aoi_poly, aes(x = x, y = y, group = id), colour = "white") +
  labs(title = "Ratio of night to day setting in South Atlantic",
       subtitle = "Jan. 1, 2019 to June 30, 2021",
       x = "lon",
       y = "lat") +
  scale_fill_gradientn(colours = c(gfw_palette("orange")[1],
wes_palette("Darjeeling1")[5]),
                       limits = c(0,1),
                       oob = scales::squish,
                       na.value = NA,
```

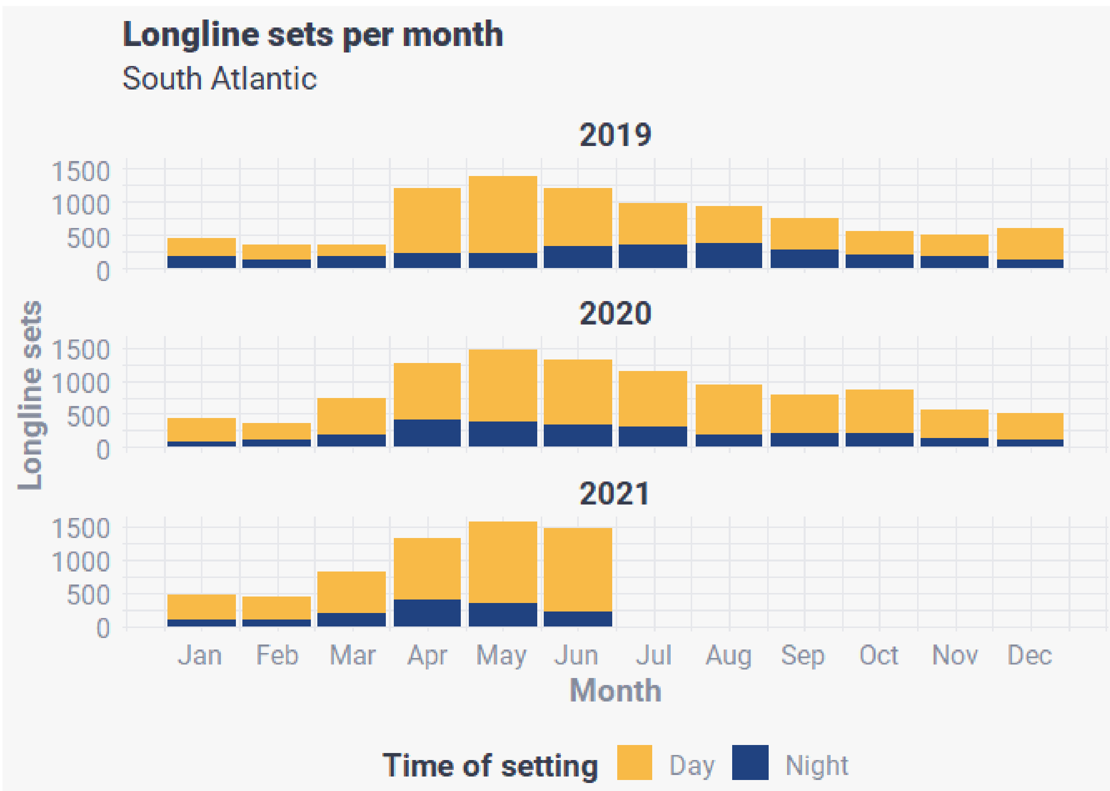
```
                       name = NULL,
                       breaks = c(0,1),
                       labels = c("All day sets", "All night sets")) +
  theme_gfw_map_cian() +
  theme(plot.title = element_text(size = 20),
        plot.subtitle = element_text(size = 16),
        axis.title = element_blank(),
        axis.text = element_text(size = 16),
        legend.text = element_text(size = 16)) +
  coord_sf(xlim = c(bounding_1$box_out[['xmin']], bounding_1$box_out[['xmax']]),
           ylim = c(bounding_1$box_out[['ymin']], bounding_1$box_out[['ymax']]),
           crs = bounding_1$out_crs)
```



Next, we plotted the number of longlines set per month throughout the South Atlantic (south of 25°S), coloured by the time of setting (night or day):

```
ll_ns %>%
  # extract the month and year from the setting start_time
  mutate(month = lubridate::month(start_time),
         year = lubridate::year(start_time)) %>%
  # filter all sets that started OR ended south of 25°S
  filter(start_lat <= -25 | end_lat <= -25) %>%
  ggplot() +
    geom_bar(aes(x = month, fill = cat3)) +
    facet_wrap(~year, ncol = 1) +
    scale_fill_discrete(type = rev(gfw_palette("chart")[c(1,7)]),
                        name = "Time of setting",
                          labels = c("Day", "Night")) +
```

```
    scale_x_continuous(breaks = c(1,2,3,4,5,6,7,8,9,10,11,12),
                       labels = c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul",
"Aug", "Sep", "Oct", "Nov", "Dec")) +
    labs(x = "Month",
         y = "Longline sets",
         title = "Longline sets per month",
         subtitle = "South Atlantic") +
      theme_gfw_cian() +
    theme(plot.title = element_text(size = 16),
          plot.subtitle = element_text(size = 14),
          axis.text = element_text(size = 12),
          axis.title = element_text(size = 14),
          legend.title = element_text(size = 14),
          legend.text = element_text(size = 12),
          strip.text = element_text(size = 14),
          legend.position = "bottom")
```

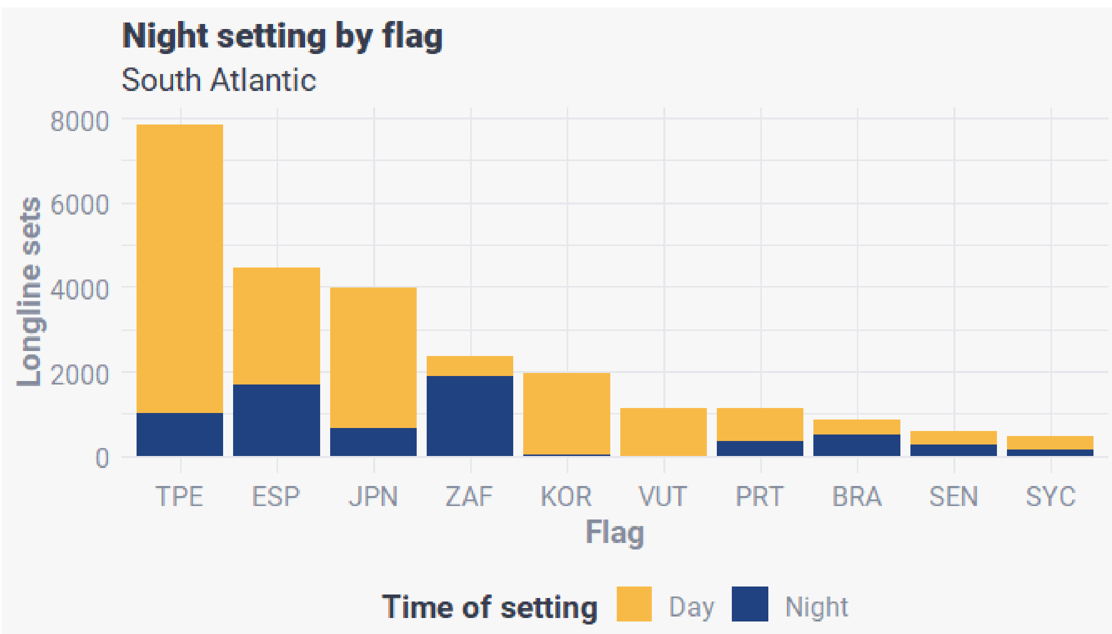


We also plotted the number of longline sets deployed in the South Atlantic (south of 25°S) by vessels belonging to the 10 most active flags:

```
# create a dataframe of the 10 flags that deployed the most sets in the region
# order these flags by the number of longline sets
# used to reorder the flags in the next plot
top_10_by_flag <- ll_ns %>%
  filter(!is.na(best_flag),
         start_lat <= -25) %>%
  group_by(best_flag) %>%
  summarise(n = n()) %>%
```

```
  top_n(10, wt = n) %>%
  arrange(desc(n))

# plot longline sets per flag, coloured by time of setting (night or day)
ll_ns %>%
  # include only flags from top 10 and sets south of 25°S
  filter(best_flag %in% top_10_by_flag$best_flag,
         (start_lat <= -25 | end_lat <= -25)) %>%
  # reorder best_flag according to the order in top_10_by_flag
  mutate(best_flag = best_flag %>%
           factor(levels = top_10_by_flag[["best_flag"]]) %>%
           recode_factor("TWN" = "TPE")) %>%
    ggplot() +
    geom_bar(aes(x = best_flag, fill = as.factor(cat3))) +
    scale_fill_discrete(type = rev(gfw_palette("chart")[c(1,7)]),
                        name = "Time of setting",
                        labels = c("Day", "Night")) +
    labs(x = "Flag",
         y = "Longline sets",
         title = "Night setting by flag",
         subtitle = "South Atlantic") +
    theme_gfw_cian() +
    theme(plot.title = element_text(size = 16),
          plot.subtitle = element_text(size = 14),
          axis.text = element_text(size = 12),
          axis.title = element_text(size = 14),
          legend.title = element_text(size = 14),
          legend.text = element_text(size = 12),
          legend.position = "bottom")
```



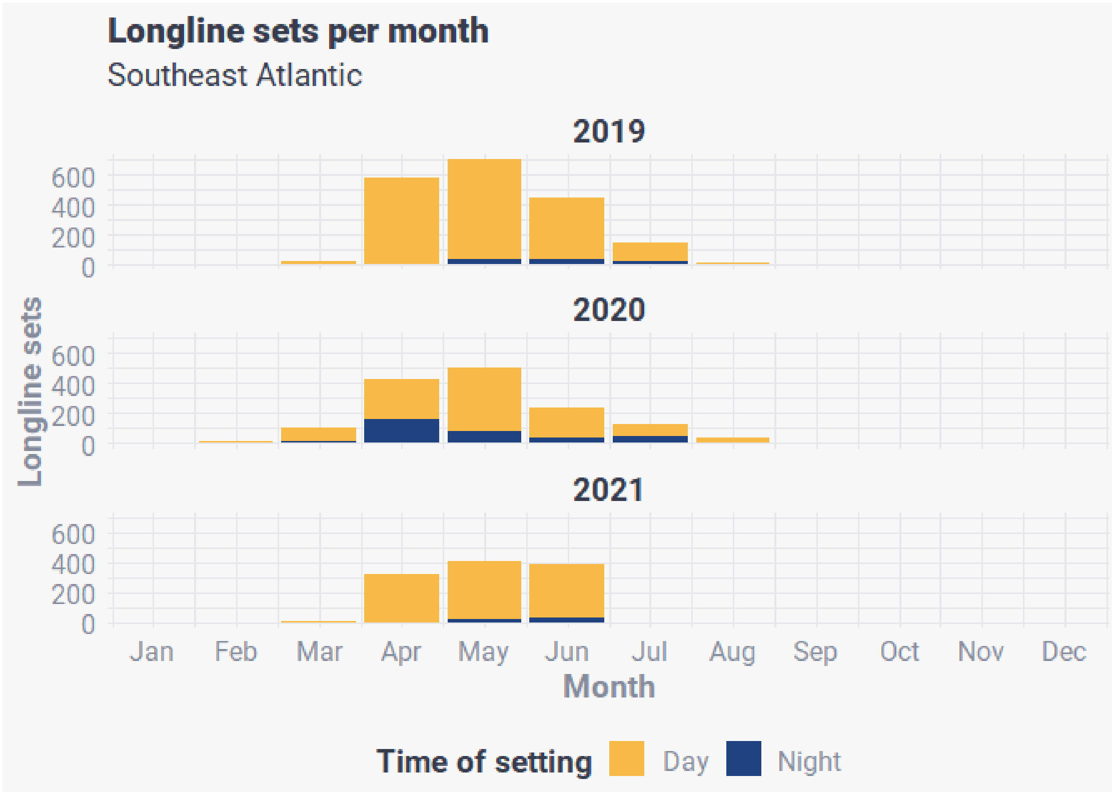
Night-setting in the high seas near Gough Island

We subset `ll_ns` to include only sets that started or finished within the area of interest to the southeast of Gough Island.

```
ll_ns_tdc <- ll_ns %>%
  filter((start_lon %>% between(-18, 12) & start_lat %>% between(-46, -40)) |
           (end_lon %>% between(-18, 12) & end_lat %>% between(-46, -40)))
```

We plotted the number of longlines set per month throughout area of interest, coloured by the time of setting (`night` or `day`):

```
ll_ns_tdc %>%
  # extract the month and year from set start_time
  mutate(month = lubridate::month(start_time),
         year = lubridate::year(start_time)) %>%
  ggplot() +
    geom_bar(aes(x = month, fill = cat3)) +
    # create separate rows for each year
    facet_wrap(~year, ncol = 1) +
    scale_fill_discrete(type = rev(gfw_palette("chart")[c(1,7)]),
                        name = "Time of setting",
                        labels = c("Day", "Night")) +
    scale_x_continuous(breaks = c(1,2,3,4,5,6,7,8,9,10,11,12),
                       labels = c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul",
"Aug", "Sep", "Oct", "Nov", "Dec"),
                       limits = c(1,12)) +
    labs(x = "Month",
         y = "Longline sets",
         title = "Longline sets per month",
         subtitle = "Southeast Atlantic") +
theme_gfw_cian() +
    theme(plot.title = element_text(size = 16),
          plot.subtitle = element_text(size = 14),
          axis.text = element_text(size = 12),
          axis.title = element_text(size = 14),
          legend.title = element_text(size = 14),
          legend.text = element_text(size = 12),
          strip.text = element_text(size = 14),
          legend.position = "bottom")
```

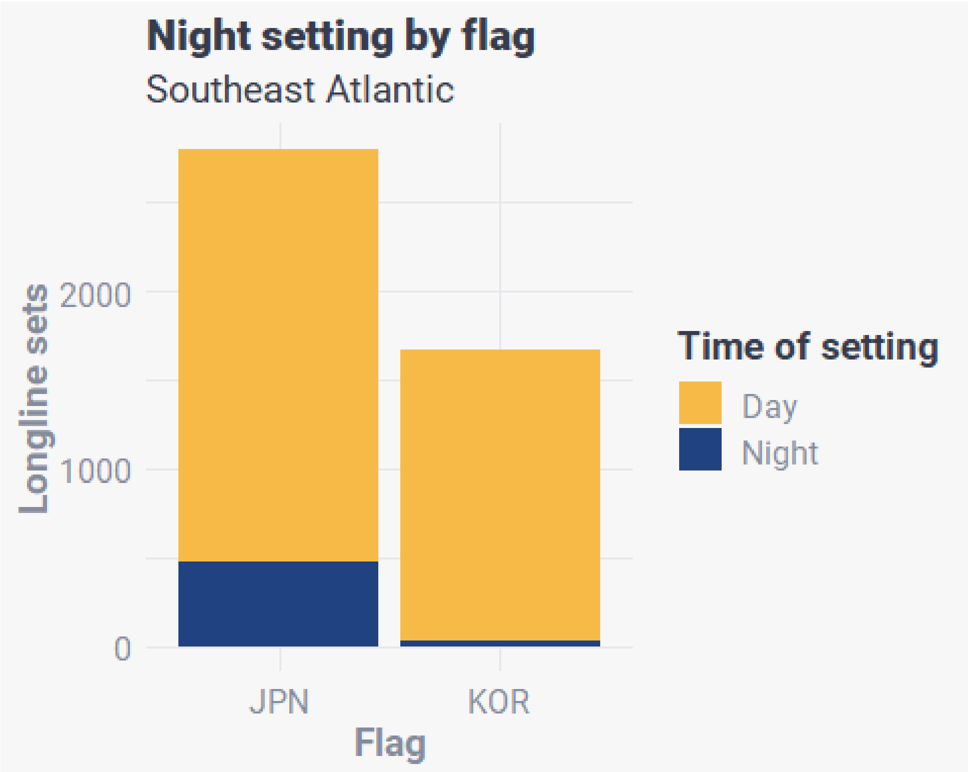


Plotted the number of longline sets deployed in the area of interest by Japanese and Korean vessels, which were responsible for 99% of longline sets in the area:

```
# create a dataframe of the 10 flags that deployed the most sets in the region
# order these flags by the number of longline sets
# used to reorder the flags in the next plot
top_aoi_by_flag <- ll_ns_tdc %>%
  filter(!is.na(best_flag)) %>%
  group_by(best_flag) %>%
  summarise(n = n(),
            prop = n/nrow(ll_ns_tdc),
            n_vessels = n_distinct(ssvid)) %>%
  arrange(desc(n))

# plot longline sets per flag, coloured by time of setting (night or day)
ll_ns_tdc %>%
  # extract month and year from start_time
  # reorder best_flag according to the order they appear in top_aoi_by_flag
  mutate(month = lubridate::month(start_time),
         year = lubridate::year(start_time),
         best_flag = best_flag %>% factor(levels = top_aoi_by_flag[["best_flag"]]))
%>%
  # filter to include only japanese and korean vessels
  filter(best_flag %in% c("JPN", "KOR")) %>%
    ggplot() +
    geom_bar(aes(x = best_flag, fill = as.factor(cat3))) +
    scale_fill_discrete(type = rev(gfw_palette("chart")[c(1,7)]),
                        name = "Time of setting",
```

```
                        labels = c("Day", "Night")) +
    labs(x = "Flag",
         y = "Longline sets",
         title = "Night setting by flag",
         subtitle = "Southeast Atlantic") +
      theme_gfw_cian() +
    theme(plot.title = element_text(size = 16),
          plot.subtitle = element_text(size = 14),
          axis.text = element_text(size = 12),
          axis.title = element_text(size = 14),
          legend.title = element_text(size = 14),
          legend.text = element_text(size = 12),
          legend.position = "right")
```



We mapped the locations of longlines set by Japanese and Korean flagged vessels relative to other flags.

First, we re-gridded the data and identified which flag set the most longlines in each grid cell. Then we mapped these locations coloured by the most active flag at each location, with a colour scale that highlights Japanese and Korean-flagged vessels.

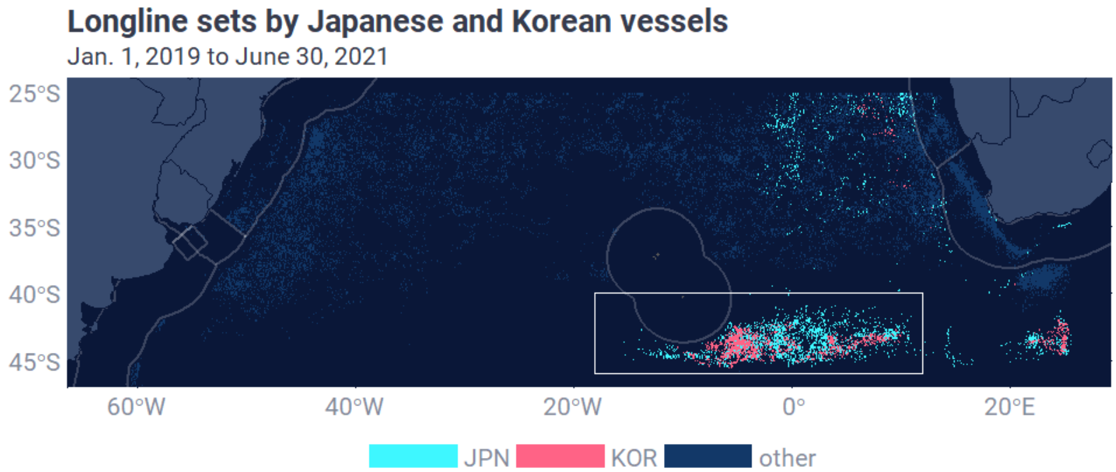
```
# regrid the data to highlight the most active flag at each location
ll_ns_gridded_2 <- ll_ns %>%
  dplyr::select(best_flag, start_lon, set_id, start_lat, end_lon, end_lat,
start_time) %>%
  rowwise() %>%
  mutate(mean_lon = mean(c(start_lon, end_lon), na.rm = TRUE), # average longitude
and latitude between start and end locations
         mean_lat = mean(c(start_lat, end_lat), na.rm = TRUE),
```

```
         lon_bin = floor(mean_lon*10)/10, # set the grid resolution to one 10th of a
degree
         lat_bin = floor(mean_lat*10)/10,
         date = lubridate::date(start_time) # use the start time to assign a date to
each set
         ) %>%
  group_by(lat_bin, lon_bin) %>%
  # filter to keep only the most prominent flag at each location
  filter(best_flag == max(best_flag)) %>%
  # create a new variable called flag_2 that groups all non-Japanese and non-Korean
vessels as "other"
  mutate(flag_2 = if_else(best_flag %in% c("JPN", "KOR"), best_flag, "other"))
```

Map of setting in South Atlantic coloured by Japanese and Korean vessel:

```
# set the bounding area for the map
bounding_1 <- fishwatchr::transform_box(xlim = c(-62, 25),
                                        ylim = c(-25, -46),
                                        output_crs = "+proj=longlat +ellps=WGS84
+datum=WGS84 +no_defs")

# make map with ggplot2
ll_ns_gridded_2 %>%
  filter(lat_bin < -25) %>%
ggplot() +
  geom_raster(aes(x = lon_bin,
                  y = lat_bin,
                  fill = flag_2)) +
  fishwatchr::geom_gfw_eez(lwd = 1) +
  fishwatchr::geom_gfw_land() +
  geom_sf(data = tdc_sf, fill = gfw_palette("map_country_dark")[1]) +
  labs(title = "Longline sets by Japanese and Korean vessels",
       subtitle = "Jan. 1, 2019 to June 30, 2021",
       x = "lon",
       y = "lat") +
  scale_fill_manual(values = gfw_palette("diverging")[c(1,7,4)],
                    labels = c("JPN", "KOR", "other"),
                    na.value = NA,
                    name = NULL) +
  geom_path(data = aoi_poly, aes(x = x, y = y, group = id), colour = "white") +
  theme_gfw_map_cian() +
  theme(plot.title = element_text(size = 20),
        plot.subtitle = element_text(size = 16),
        axis.title = element_blank(),
        axis.text = element_text(size = 16),
        legend.text = element_text(size = 16)) +
  coord_sf(xlim = c(bounding_1$box_out[['xmin']], bounding_1$box_out[['xmax']]),
           ylim = c(bounding_1$box_out[['ymin']], bounding_1$box_out[['ymax']]),
           crs = bounding_1$out_crs)
```



Setting over dawn

To understand the potential for increased bycatch risk for Tristan albatross during the dawn period, we compared the timing of longline setting activity with the mean time of nautical dawn in the area. First, we calculated the mean time of nautical dawn at the time and location where each longline was set, within the area of interest to the southeast of Gough Island. We then formatted this as a datetime object, with a dummy date (January 1st 2020), for the purpose of plotting:

```
dawn_times <- ll_ns_tdc %>%
  summarise(start_ndawn = mean(start_ndawn)) %>%
  mutate(start_ndawn = ymd_hms("2020-01-01 00:00:00") + (start_ndawn*60*60))
```

Plotted the density of the time of day when longline setting began and finished, including the mean time of nautical dawn for reference. For this we had to reformat the data into a structure with one column called `Set` which specified `start` or end, and another column `time_of_day`, which included the time of day that each setting event started or ended. `time_of_day` was formatted as a datetime object, with the same dummy date (January 1st 2020), for the purposes of plotting.

```
ll_ns_tdc %>%
  mutate(start = ymd_hms("2020-01-01 00:00:00") + local_time(start_time, units =
"hours"),
          end = ymd_hms("2020-01-01 00:00:00") + local_time(end_time, units =
"hours")) %>%
  dplyr::select(start, end) %>%
  pivot_longer(cols = start:end, names_to = "Set", values_to = "time_of_day") %>%
  mutate(Set = Set %>% factor(levels = c("start", "end"))) %>%
  ggplot() +
  # use stat_density instead of geom_density so the legend symbols are lines rather
than boxes
    stat_density(aes(x = time_of_day, colour = Set, linetype = Set),
                 geom = "line", position = "identity", alpha = 1, size = 1) +
```

```r
    geom_vline(data = dawn_times, aes(xintercept = start_ndawn), colour =
gfw_palette("chart")[5], size = 1) +
    scale_x_datetime(date_labels = "%H:%M") +
    scale_colour_manual(name = "Setting",
                        values = c(gfw_palette("chart")[1], "grey50"),
                        labels = c("Start", "End")) +
    scale_linetype_manual(name = "Setting",
                          values = c(1,2),
                          labels = c("Start", "End")) +
    labs(x = "Time of day",
         y = "Proportion of sets",
         title = "Time of setting and hauling of longlines",
         subtitle = "Southeast Atlantic") +
    theme_gfw_cian() +
    theme() +
        theme(plot.title = element_text(size = 16),
          plot.subtitle = element_text(size = 14),
          axis.text = element_text(size = 12),
          axis.title = element_text(size = 14),
          legend.title = element_text(size = 14),
          legend.text = element_text(size = 12))
```