BIO710 - Independent Project Introduction

Shifting signals: Exploring geographic variation in echolocation clicks of Risso's dolphins in the California Current

Sarah Shiers

2025-03-09

Inroduction

Risso's dolphins (*Grampus griseus*) are deep-diving cetaceans that rely on echolocation for foraging and navigation, making them particularly sensitive to changes in their acoustic environment (Arranz et all., 2016). While these dolphins are broadly distributed, their population structure remains poorly understood, particularly in the California Current Ecosystem. Current stock assessments rely on large-scale geographic distinctions which may overlook important regional variations (Soldevilla et al., 2017). Understanding finer-scale population structure is crucial for conservation efforts, especially as offshore wind energy development expands along the U.S. West Coast, introducing potential acoustic disturbances that could impact this species. Passive acoustic monitoring (PAM) has proven to be an effective tool for studying cetacean populations over long time scales, offering insights into species distribution and behavior through the analysis of their echolocation signals (Soldevilla et al., 2008; Van Parijs et al., 2009).

This study examines spatial, temporal, and geographic variation in the spectral characteristics of Risso's dolphin echolocation clicks using 273 detections of Risso's dolphins from the California Current Ecosystem collected over six years across multiple passive acoustic surveys (Rankin et al., 2024). The analysis follows the methodology established by Soldevilla et al. (2017) while being tailored to this unique dataset. Click analysis will be conducted using custom scripts in Matlab and R, with all code made publicly available on GitHub to ensure transparency and reproducibility. The primary focus will be on spectral features, specifically peak frequencies and notch depths, which characterize species-specific banding patterns (Soldevilla et al., 2008). Mean spectral peak and notch frequencies will be calculated for each encounter, using the R package "PAMpal" to assist in feature extraction (Sakai, 2023).

As in Soldevilla et al.'s (2017) research, to assess geographic variation, statistical tests will be employed to determine whether click characteristics differ between regions. K-means clustering will be applied to spectral peak and notch data to identify common patterns across locations, while ANOVA will test for significant differences in spectral features among predefined geographic regions. Additionally, regression analysis will evaluate correlations between geographic position and click characteristics, providing insight into whether spatial separation corresponds to distinct spectral profiles. These analyses will help determine whether regional variation in Risso's dolphin echolocation clicks reflects underlying population structure, contributing to improved stock assessments and informing conservation efforts amid growing offshore wind development in the region (Soldevilla et al., 2008, Soldevilla et al., 2017).

Research Question

How do the spectral features of Risso's dolphin echolocation clicks vary geographically and temporally within the California Current?

Hypothesis

I hypothesize that the spectral features of Risso's dolphin echolocation clicks (specifically peak frequencies and notch depths) exhibit significant geographic variation within the California Current Ecosystem, reflecting potential differences in population structure or habitat use.

Specific Aims

1. Examine and document spatial and temporal variation in Risso's dolphin click types throughout the California Current.
2. Examine and contrast the spectral characteristics of Risso's dolphin clicks in the California Current with those from other regions, such as the Hawaiian Islands, Gulf of Mexico, and U.S. East Coast.

Data Loading

The following datasets are used in this project and are loaded from the `data/` directory:

```r
# Load dataset 1: Detection summary
detection_summary <- read.csv("data/Gg_Detection_Summary.csv")

# Load dataset 2: ADRIFT detections with GPS
adrift_detections <- readxl::read_excel("data/ADRIFT_AllDetections_wGPS.xlsx")

## Load dataset 3: Gg detections from all three surveys with data quality
Gg_detections <- read.csv("data/ADRIFT_GgDetections.csv")

## Load dataset 4: Deployment details and site lists
site_list_details <- read.csv("data/Deployment_Details_Site_List.csv")
```

These datasets represent Risso's dolphin click detections and associated metadata across multiple acoustic surveys conducted between 2016 and 2023. The Gg_Detection_Summary.csv file contains a manually reviewed summary of detection events, including effort and classification notes. The ADRIFT_AllDetections_wGPS.xlsx file includes detailed encounter metadata with GPS coordinates, detection timing, and contextual deployment information from the ADRIFT survey. The ADRIFT_GgDetections.csv file provides timestamped Risso's dolphin detections with associated location data, and is used in spatial and temporal analyses. Finally, the Deployment_Details_Site_List.csv file provides latitudinal ranges and region codes for each deployment site, enabling the geographic classification of detections across the study area.

Statistical Analysis: Seasonal and Temporal Variation in Risso's Dolphin Detections

Temporal Variation: Seasonal

Hypotheses

- **Null hypothesis:** There is no significant difference in the number of Risso's dolphin detections across oceanographic seasons.
- **Alternative hypothesis:** The number of Risso's dolphin detections differs significantly between oceanographic seasons (upwelling, post-upwelling, winter).

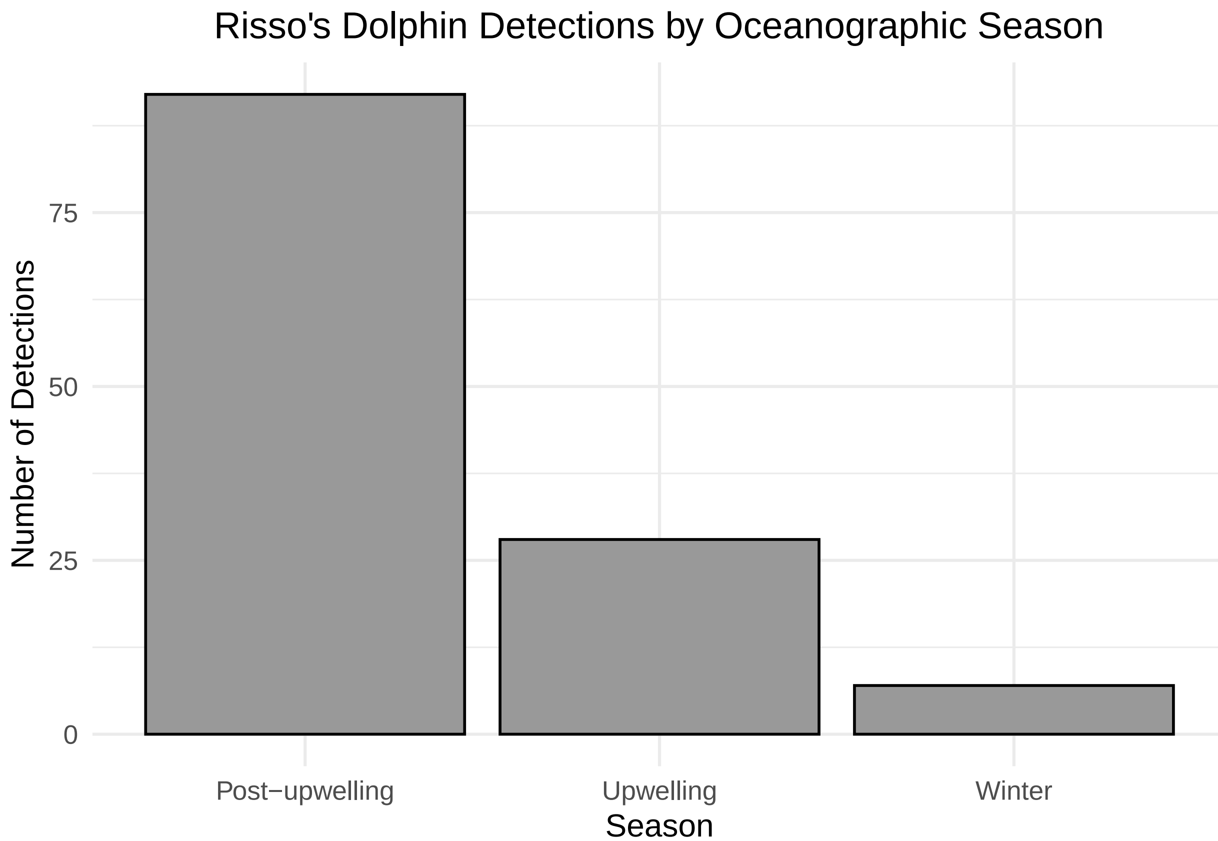
Approach To assess seasonal variation in detections, each detection is categorized based on its date into one of three oceanographic seasons: - Upwelling: March–June - Post-upwelling: July–November - Winter: December–February

Detection counts per season are summarized and a Chi-square goodness-of-fit test is used to test whether detections are evenly distributed across seasons. A bar plot is included for visualization.

```r
# Load libraries
library(tidyverse)
library(lubridate)

# Load and inspect the data
Gg_detections <- read.csv("data/ADRIFT_GgDetections.csv")
Gg_detections$UTC <- ymd_hms(Gg_detections$UTC)
Gg_detections$Date <- as.Date(Gg_detections$UTC)
Gg_detections$Month <- month(Gg_detections$Date)

# Assign oceanographic season
Gg_detections$Season <- case_when(
  Gg_detections$Month %in% 3:6 ~ "Upwelling",
  Gg_detections$Month %in% 7:11 ~ "Post-upwelling",
  Gg_detections$Month %in% c(12, 1, 2) ~ "Winter",
  TRUE ~ NA_character_
)

# Count detections per season
season_counts <- Gg_detections %>%
  count(Season) %>%
  filter(!is.na(Season)) %>%
  arrange(factor(Season, levels = c("Upwelling", "Post-upwelling", "Winter")))

# Perform Chi-square test
chisq_result <- chisq.test(season_counts$n)
chisq_result
```

```
##
##  Chi-squared test for given probabilities
##
## data:  season_counts$n
## X-squared = 92.614, df = 2, p-value < 2.2e-16
```

```r
# Plot
ggplot(season_counts, aes(x = Season, y = n)) +
  geom_col(fill = "grey60", color = "black") +
  labs(
    title = "Risso's Dolphin Detections by Oceanographic Season",
    x = "Season",
    y = "Number of Detections"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(hjust = 0.5)
  )
```



Visualization

Figure 1. Number of Risso's dolphin detections per oceanographic season from the ADRIFT dataset.

Statistical Results The Chi-squared test revealed a statistically significant difference in detection frequencies across seasons (chi-saquared = 92.61, df = 2, $p < 0.001$), indicating that Risso's dolphin detections were not evenly distributed throughout the year. The post-upwelling season had the highest number of detections, suggesting increased presence or detectability during this period in the California Current.

Temportal Analysis: Time of Day

Hypotheses

- **Null hypothesis:** Risso's dolphin detections occur equally across all times of day.
- **Alternative hypothesis:** Detection rates vary by time of day, indicating a diel pattern.

Approach To assess diel variation in detections, the hour is extracted from each detection timestamp and it is assigned to one of six 4-hour bins:

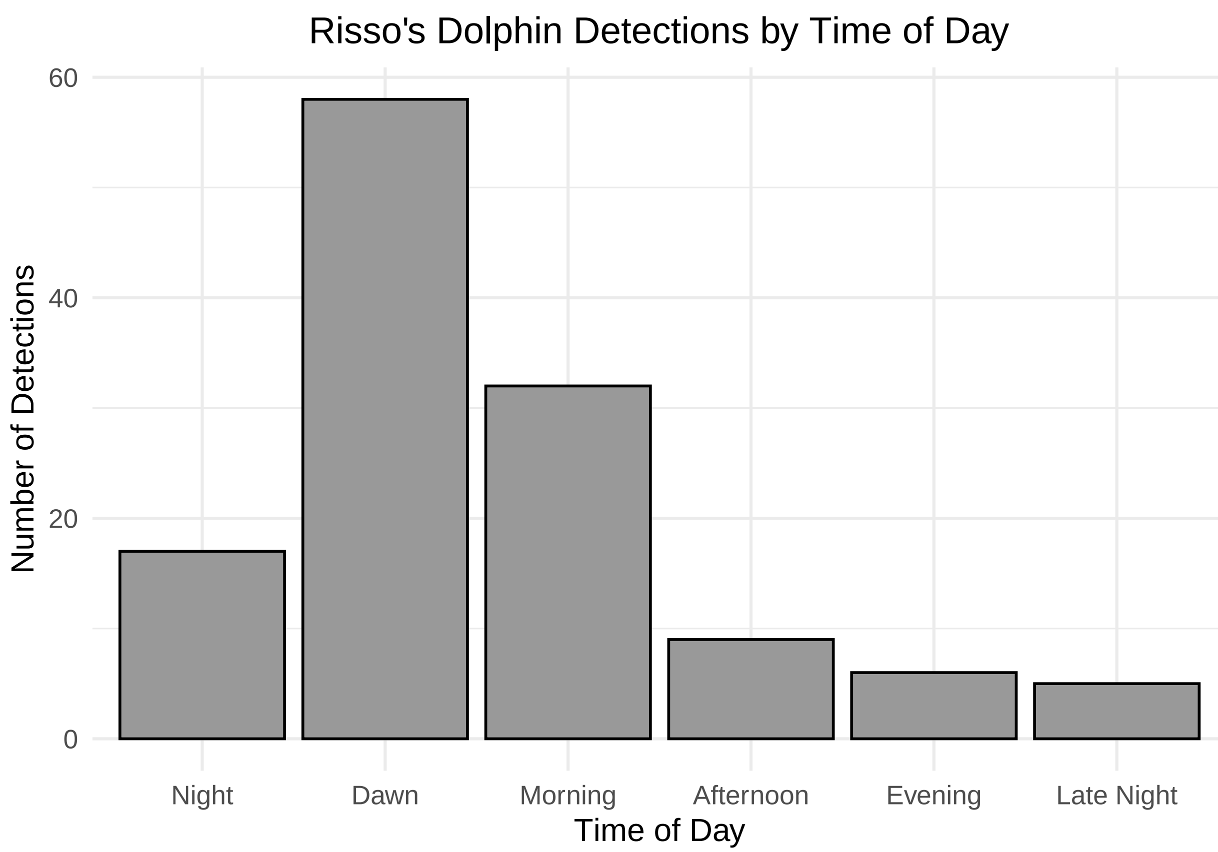
- Night: 00:00–03:59
- Dawn: 04:00–07:59
- Morning: 08:00–11:59
- Afternoon: 12:00–15:59
- Evening: 16:00–19:59
- Late Night: 20:00–23:59

Detection counts are summarized per time-of-day bin and a Chi-square goodness-of-fit test is used to assess whether detections are evenly distributed across the diel cycle. A bar plot is included for visualization.

```r
# Extract hour from timestamp
Gg_detections$Hour <- hour(Gg_detections$UTC)

# Define time-of-day bins
Gg_detections$TimeOfDay <- case_when(
  Gg_detections$Hour >= 0 & Gg_detections$Hour < 4 ~ "Night",
  Gg_detections$Hour >= 4 & Gg_detections$Hour < 8 ~ "Dawn",
  Gg_detections$Hour >= 8 & Gg_detections$Hour < 12 ~ "Morning",
  Gg_detections$Hour >= 12 & Gg_detections$Hour < 16 ~ "Afternoon",
  Gg_detections$Hour >= 16 & Gg_detections$Hour < 20 ~ "Evening",
  Gg_detections$Hour >= 20 & Gg_detections$Hour < 24 ~ "Late Night",
  TRUE ~ NA_character_
)

# Count detections per time bin
time_counts <- Gg_detections %>%
  count(TimeOfDay) %>%
  filter(!is.na(TimeOfDay)) %>%
  arrange(factor(TimeOfDay, levels = c("Night", "Dawn", "Morning", "Afternoon", "Evening", "Late Night"

# Chi-squared test
time_chisq_result <- chisq.test(time_counts$n)
time_chisq_result
```

```
##
##  Chi-squared test for given probabilities
##
## data:  time_counts$n
## X-squared = 100.67, df = 5, p-value < 2.2e-16
```

```r
# Count detections per time bin and set proper order
time_counts <- Gg_detections %>%
  count(TimeOfDay) %>%
  filter(!is.na(TimeOfDay)) %>%
  mutate(TimeOfDay = factor(TimeOfDay, levels = c(
    "Night", "Dawn", "Morning", "Afternoon", "Evening", "Late Night"
  )))

# Plot
ggplot(time_counts, aes(x = TimeOfDay, y = n)) +
  geom_col(fill = "grey60", color = "black") +
  labs(
    title = "Risso's Dolphin Detections by Time of Day",
    x = "Time of Day",
    y = "Number of Detections"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(hjust = 0.5)
  )
```



Visualization

Figure 2. Number of Risso's dolphin detections by time of day from the ADRIFT dataset.

Statistical Results The Chi-squared test revealed a statistically significant difference in detection frequencies across time-of-day bins (chi-squared = 100.67, df = 5, $p < 0.001$), indicating that Risso's dolphin detections are not evenly distributed throughout the diel cycle. The highest number of detections occurred during the Dawn and Morning periods, suggesting that Risso's dolphins may be more vocally active or more detectable during early-day hours in the California Current.

Spatial Variation: Mapping Detection Locations

To visualize the geographic distribution of Risso's dolphin detections, latitude and longitude of each detection are plotted on a map of the U.S. West Coast using base map data from the `maps` package.

```
library(maps)
library(ggplot2)

# Get a base map of the west coast
us_map <- map_data("world")

# Plot
ggplot() +
  geom_polygon(data = us_map, aes(x = long, y = lat, group = group),
               fill = "grey95", color = "grey70") +
  geom_point(data = Gg_detections, aes(x = Longitude, y = Latitude),
             color = "grey20", size = 2, alpha = 0.7) +
  coord_fixed(xlim = c(-130, -120), ylim = c(30, 50)) +
  labs(title = "Risso's Dolphin Detections Along the U.S. West Coast",
```

```r
      x = "Longitude",
      y = "Latitude") +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(hjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1)
  )
```

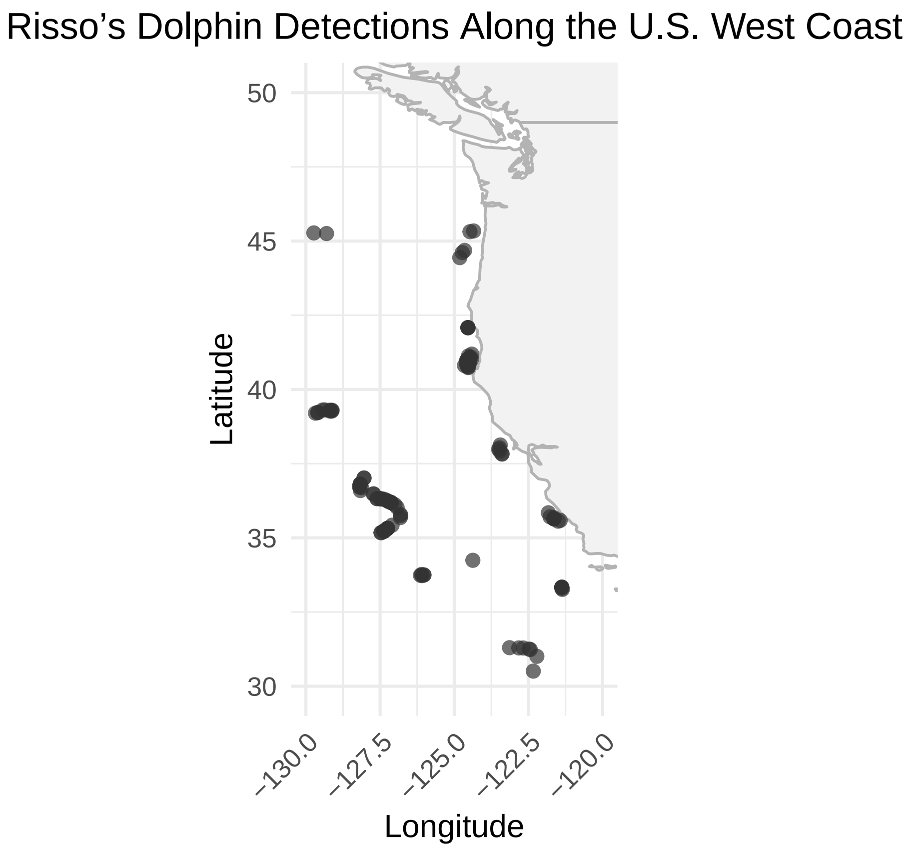


Figure 3. Geographic locations of Risso's dolphin detections from the ADRIFT dataset.

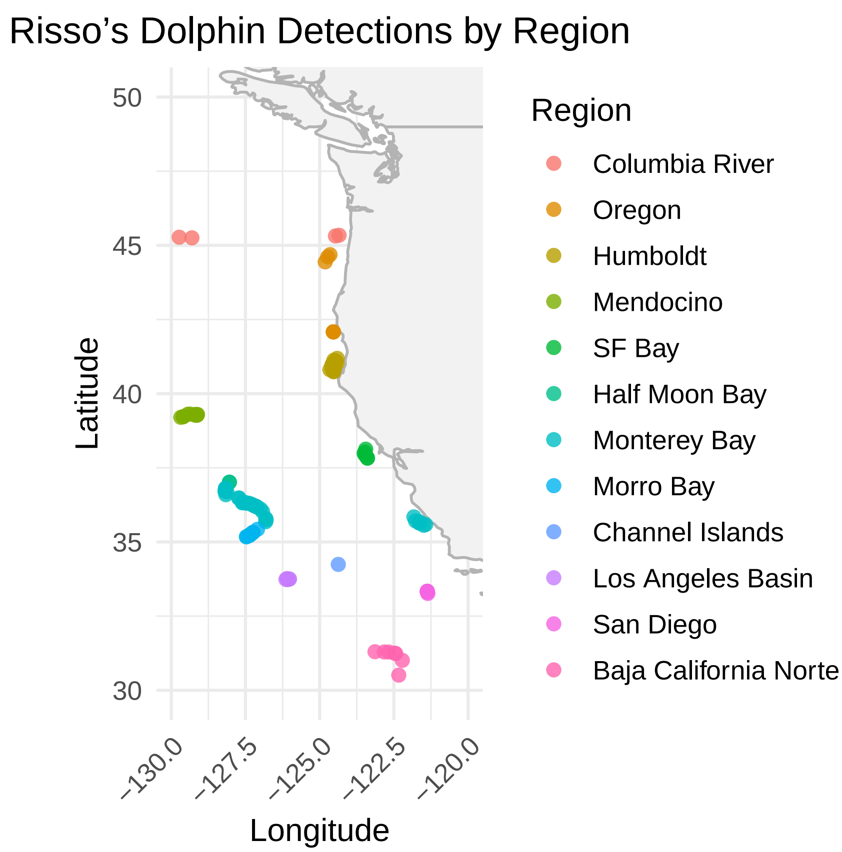
Spatial Variation: Regional Identification of Detections

Hypotheses

- **Null hypothesis:** There is no significant difference in the number of Risso's dolphin detections across regions.
- **Alternative hypothesis:** The number of Risso's dolphin detections differs significantly between regions.

Approach To assess spatial variation in detections, each detection is categorized by its assigned region using latitude boundaries provided in the deployment site metadata. Detection counts per region are summarized and a Chi-square goodness-of-fit test is used to test whether detections are evenly distributed across regions. A bar plot and map are included for visualization.

```r
library(dplyr)
library(purrr)
library(maps)
library(ggplot2)

# Clean and prepare site_list_details
site_list_details <- site_list_details %>%
  filter(!is.na(Region.Code), !is.na(Lat_min), !is.na(Lat_max))

# Function to match region by latitude
assign_region <- function(lat) {
  match <- site_list_details %>%
    filter(lat >= Lat_min & lat <= Lat_max) %>%
    pull(Region)

  if (length(match) > 0) return(match[1]) else return(NA)
}

# Apply to Gg_detections
Gg_detections <- Gg_detections %>%
  mutate(Region = map_chr(Latitude, assign_region))

# Create ordered list of regions by midpoint latitude (north to south)
region_order <- site_list_details %>%
  mutate(Lat_mid = (Lat_min + Lat_max) / 2) %>%
  arrange(desc(Lat_mid)) %>%
  distinct(Region) %>%
  pull(Region)

# Convert Region column to ordered factor
Gg_detections <- Gg_detections %>%
  mutate(Region = factor(Region, levels = region_order))

# Plot
ggplot() +
  geom_polygon(data = us_map, aes(x = long, y = lat, group = group),
               fill = "grey95", color = "grey70") +
  geom_point(data = Gg_detections %>% filter(!is.na(Region)),
             aes(x = Longitude, y = Latitude, color = Region),
             size = 2, alpha = 0.8) +
  coord_fixed(xlim = c(-130, -120), ylim = c(30, 50)) +
  labs(title = "Risso's Dolphin Detections by Region",
       x = "Longitude",
       y = "Latitude",
       color = "Region") +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(hjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1)
  )
```



Visualization

Figure 4. Geographic locations of Risso's dolphin detections from acoustic surveys conducted between 2016 and 2023, colored by region.

```r
# Summarize detection counts per region
region_counts <- Gg_detections %>%
  count(Region) %>%
  filter(!is.na(Region)) %>%
  arrange(desc(n))

# Plot
ggplot(region_counts, aes(x = factor(Region, levels = region_order), y = n)) +
  geom_col(fill = "grey60", color = "black") +
  labs(
    title = "Risso's Dolphin Detections by Region",
    x = "Region (North to South)",
    y = "Number of Detections"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(hjust = 0.5),
    axis.text.x = element_text(angle = 45, hjust = 1)
  )
```

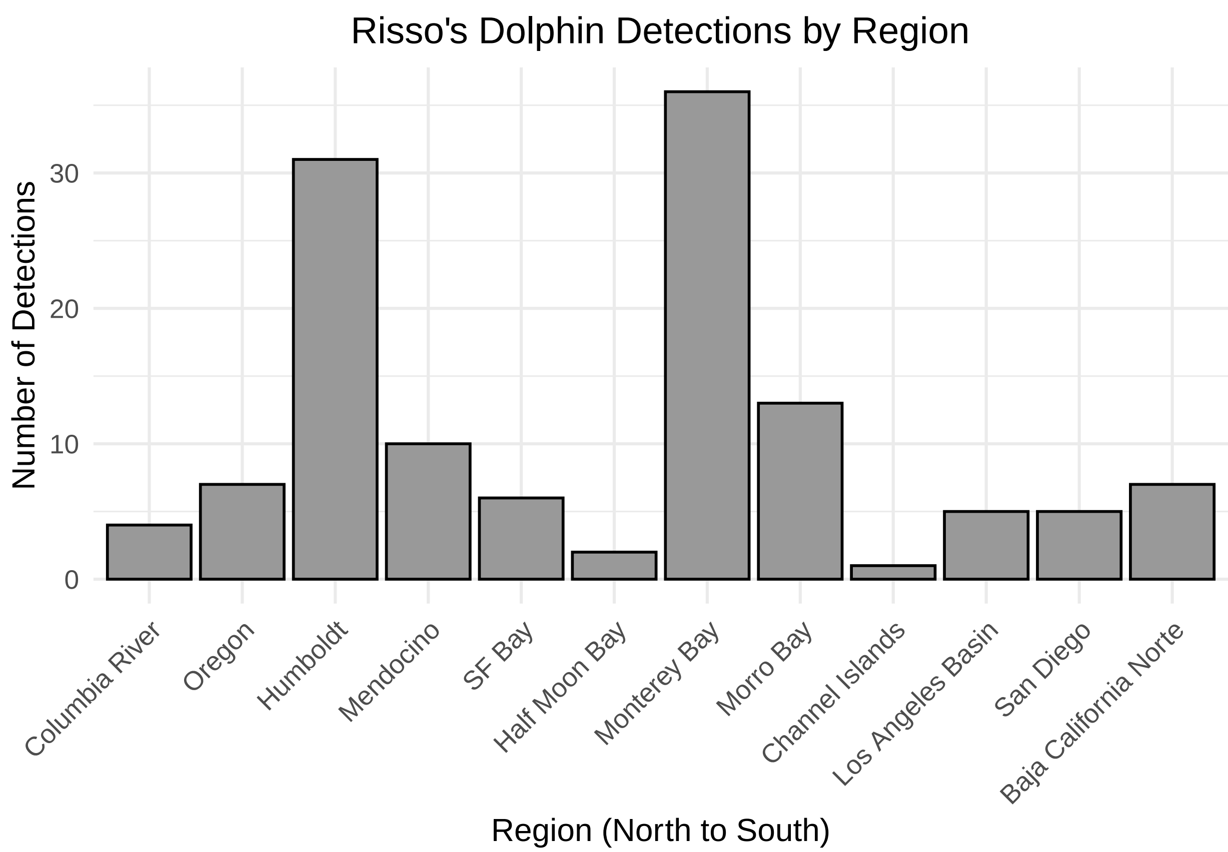


Figure 5. Number of Risso's dolphin detections by region from acoustic surveys.

```r
# Perform Chi-square test
region_chisq_result <- chisq.test(region_counts$n)
region_chisq_result
```

```
##
##  Chi-squared test for given probabilities
##
## data:  region_counts$n
## X-squared = 131.05, df = 11, p-value < 2.2e-16
```

Statistical Results The Chi-squared test revealed a statistically significant difference in detection frequencies across geographic regions (chi-squared = 131.05, df = 11, $p < 0.001$), indicating that Risso's dolphin detections were not evenly distributed along the U.S. West Coast. The highest number of detections occurred in the Monterey Bay and Humboldt regions, suggesting localized hotspots of vocal activity or presence. Conversely, regions such as Half Moon Bay and the Channel Islands showed notably fewer detections, highlighting spatial variability in dolphin distribution or detectability within the California Current Ecosystem.

References

Arranz, P., DeRuiter, S. L., Stimpert, A. K., Neves, S., Friedlaender, A. S., Goldbogen, J. A., Visser, F., Calambokidis, J., Southall, B. L., & Tyack, P. L. (2016). Discrimination of fast click-series produced by tagged Risso's dolphins (Grampus griseus) for echolocation or communication. Journal of Experimental Biology, 219(18), 2898–2907. https://doi.org/10.1242/jeb.144295

Rankin, S., Burger, K., Hom-Weaver, C., Palmer K., Sakai, ., & Simonis, A. (2024). Final Report for Adrift in the California Current Survey: Passive Acoustic Monitoring in the California Current Using Drifting Recorders. Camarillo, CA: U.S. Department of the Interior, Bureau of Ocean Energy Management. Report No. OCS Study BOEM 2024-047. Contract No. G14 M20PG00013.

Sakai, T. (2023). PAMpal: Load and Process Passive Acoustic Data. R package version 1.0.0, https://CRAN.R-project.org/package=PAMpal.

Soldevilla, M. S., Baumann-Pickering, S., Cholewiak, D., Hodge, L. E. W., Oleson, E. M., & Rankin, S. (2017). Geographic variation in Risso's dolphin echolocation click spectra. The Journal of the Acoustical Society of America, 142(2), 599–617. https://doi.org/10.1121/1.4996002

Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., & Roch, M. A. (2008). Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks. The Journal of the Acoustical Society of America, 124(1), 609–624. https://doi.org/10.1121/1.2932059

Van Parijs, S., Clark, C., Sousa-Lima, R., Parks, S., Rankin, S., Risch, D., & Van Opzeeland, I. (2009). Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales. Marine Ecology Progress Series, 395, 21–36. https://doi.org/10.3354/meps08123