

# “Mooring TH042 Data Exploration”

Eleanor (Ella) Crotty

## Contents

<b>Setup</b>	<b>1</b>
<b>Functions</b>	<b>2</b>
<b>Cleaning</b>	<b>4</b>
<b>Data Exploration</b>	<b>4</b>
Histograms . . . . .	4
Scatterplots . . . . .	10
<b>Oxygen Data Exploration</b>	<b>12</b>
Histograms . . . . .	12
Time Series . . . . .	15
<b>Export</b>	<b>16</b>

## Setup

```
# Warnings and startup messages suppressed
library(tidyverse)
library(patchwork)
library(scales)
library(ggrepel)
library(readxl)
library(here)
```

## Date Formats Understood by Readr

Type	Code	Meaning
Year	%Y	4 digit year
	%y	2 digit year
Month	%m	Number
	%b	Abbreviated (ex. Feb)
Day	%B	Full name
	%d	1-2 digits
Time	%e	2 digits
	%H	24 hour hour
Time	%I	12 hour hour
	%p	am or pm
	%M	Minutes
	%S	Seconds

Type	Code	Meaning
	%OS	Seconds with decimal component
	%Z	Time zone name
	%z	Offset from UTC
Other	%.	Skip one non-digit, (ex. :)
	%*	Skip any number of non-digits

```

wd <- "OCNMS_Hypoxia/Data"
exp <- "OCNMS_Hypoxia/Outputs"
pltpath <- "OCNMS_Hypoxia/Plots/Mooring_Data_Exploration_Plots"

TH042 <- read.csv(here(wd, "MooringData_CTP0_TH042.csv"), na = "NaN")

TH042$time <- as.POSIXct(TH042$time, tryFormats = c("%e-%b-%Y %H:%M:%S"))
TH042 <- TH042 %>%
  mutate(year = as.factor(year(time)))

```

## Functions

```

# Variable to make a base map of OCNMS, which I can then add data points to
mapUC <- map_data("world", region = c("usa", "canada"))
OCNMS_x <- c(-123.5, -125.5)
OCNMS_y <- c(47, 49)

mapOCNMS <- ggplot(mapUC, aes(x = long, y = lat, group = group)) +
  geom_polygon(fill = "gray90", color = "black") +
  coord_sf(xlim = OCNMS_x,
            ylim = OCNMS_y) +
  theme_bw() +
  theme(text = element_text(size=15),
        panel.background = element_rect(fill = "azure1",
                                         colour = "azure1"),
        legend.key = element_rect(fill = "white",
                                   color = "white"),
        # It added in blue behind the dots in the key and I don't want that
        panel.grid.major = element_line(size = 0.5,
                                         linetype = 'solid',
                                         colour = "white"),
        panel.grid.minor = element_line(size = 0.25,
                                         linetype = 'solid',
                                         colour = "white")) # +

```

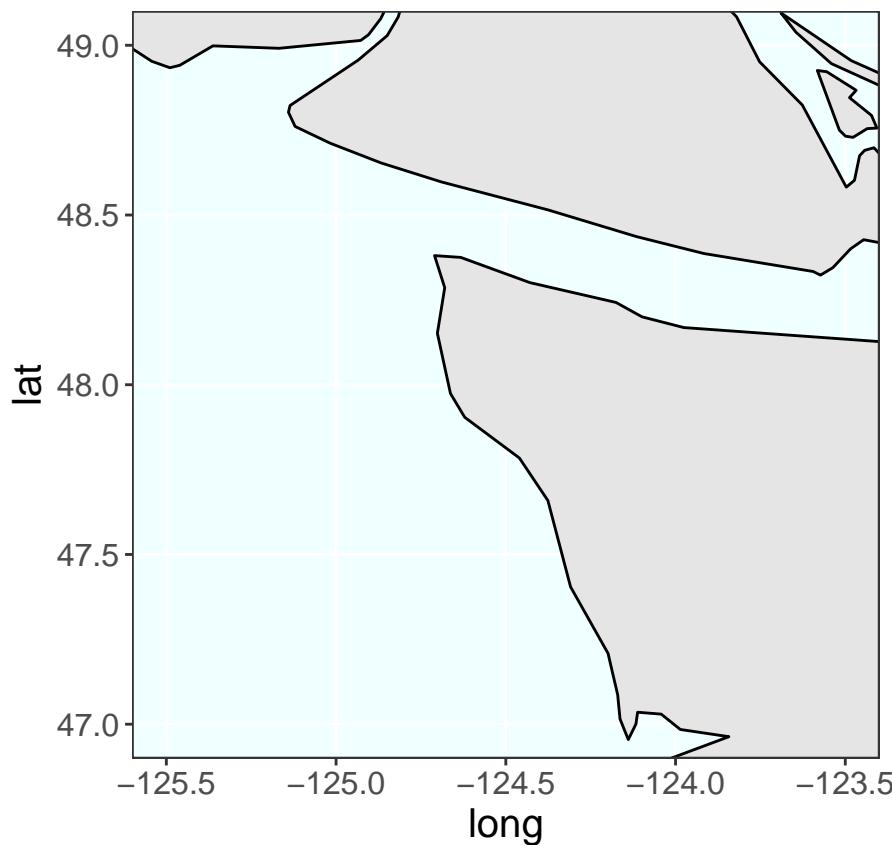
```

## Warning: The `size` argument of `element_line()` is deprecated as of ggplot2 3.4.0.
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

# coord_quickmap() # This will fix the weird stretch usually, can also do
  ↵ coord_fixed(ratio = 1.3)

```

mapOCNMS



```
# Histogram function
histogram_fill <- function(df, var, binwidth, fill = "darkgray") {
  label <- rlang::en glue("A histogram of {{var}} in {{df}} with binwidth {binwidth}")

  df |>
    ggplot(aes(x = {{ var }}), fill = {{fill}})) +
    geom_histogram(binwidth = binwidth, color = "black") +
    labs(title = label) +
    theme_bw()
}

histogram2 <- function(df, var, binwidth, fill = NA) { # Not currently functional
  label <- rlang::en glue("A histogram of {{var}} in {{df}} with binwidth {binwidth}")

  if (is.na(fill)) {
    df |>
      ggplot(aes(x = {{ var }})) +
      geom_histogram(binwidth = binwidth, color = "black", fill = "darkgray") +
      labs(title = label) +
      theme_bw()
  } else {
    df |>
      ggplot(aes(x = {{ var }}), fill = {{fill}})) +
      geom_histogram(binwidth = binwidth, color = "black") +
```

```

    labs(title = label) +
    theme_bw()
}

histogram <- function(df, var, binwidth) {
  label <- rlang::englue("A histogram of {{var}} in {{df}} with binwidth {binwidth}")

  df |>
    ggplot(aes(x = {{ var }})) +
    geom_histogram(binwidth = binwidth, color = "black", fill = "darkgray") +
    labs(title = label) +
    theme_bw()
}

scatterplt <- function(df, x, y, shape = 16, alpha = 1) {
  label <- rlang::englue("A scatterplot of {{x}} vs {{y}} in {{df}}")

  df |>
    ggplot(aes(x = {{ x }}, y = {{ y }})) +
    geom_point(color = "blue", shape = shape, alpha = alpha) +
    labs(title = label) +
    theme_bw()
}

scatterplt_fill <- function(df, x, y, color, shape = 16, alpha = 1) {
  label <- rlang::englue("A scatterplot of {{x}} vs {{y}} in {{df}}")

  df |>
    ggplot(aes(x = {{ x }}, y = {{ y }}, color = {{ color }})) +
    geom_point(shape = shape, alpha = alpha) +
    labs(title = label) +
    theme_bw()
}

```

## Cleaning

```

TH042 <- TH042 %>%
  rename(DO = oxy_mg, temperature = temp, salinity = sal, date = time, potential_density
        = dens)

```

## Data Exploration

### Histograms

```

histogram_fill(TH042, depth, 0.5, fill = year) # Range 40-45 m - ok to clip CTD data to
# that?

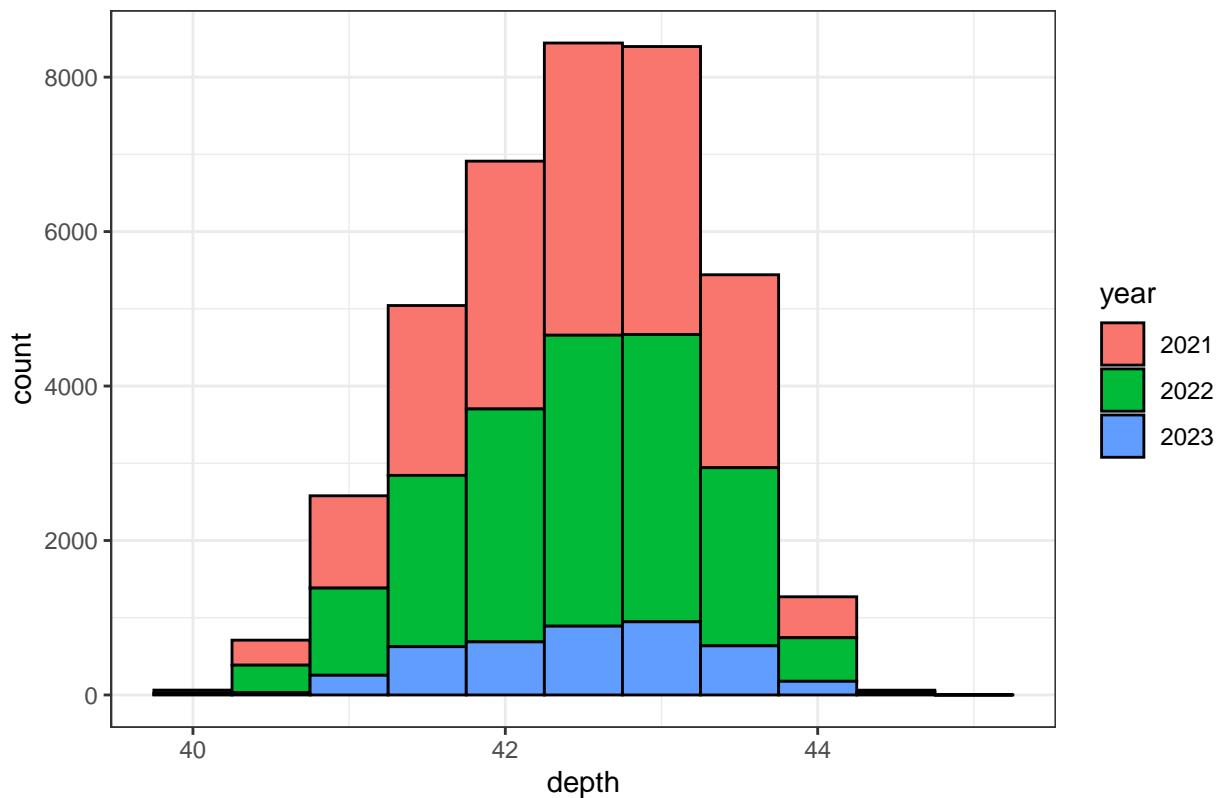
```

```

## Warning: Removed 4 rows containing non-finite outside the scale range
## (`stat_bin()`).

```

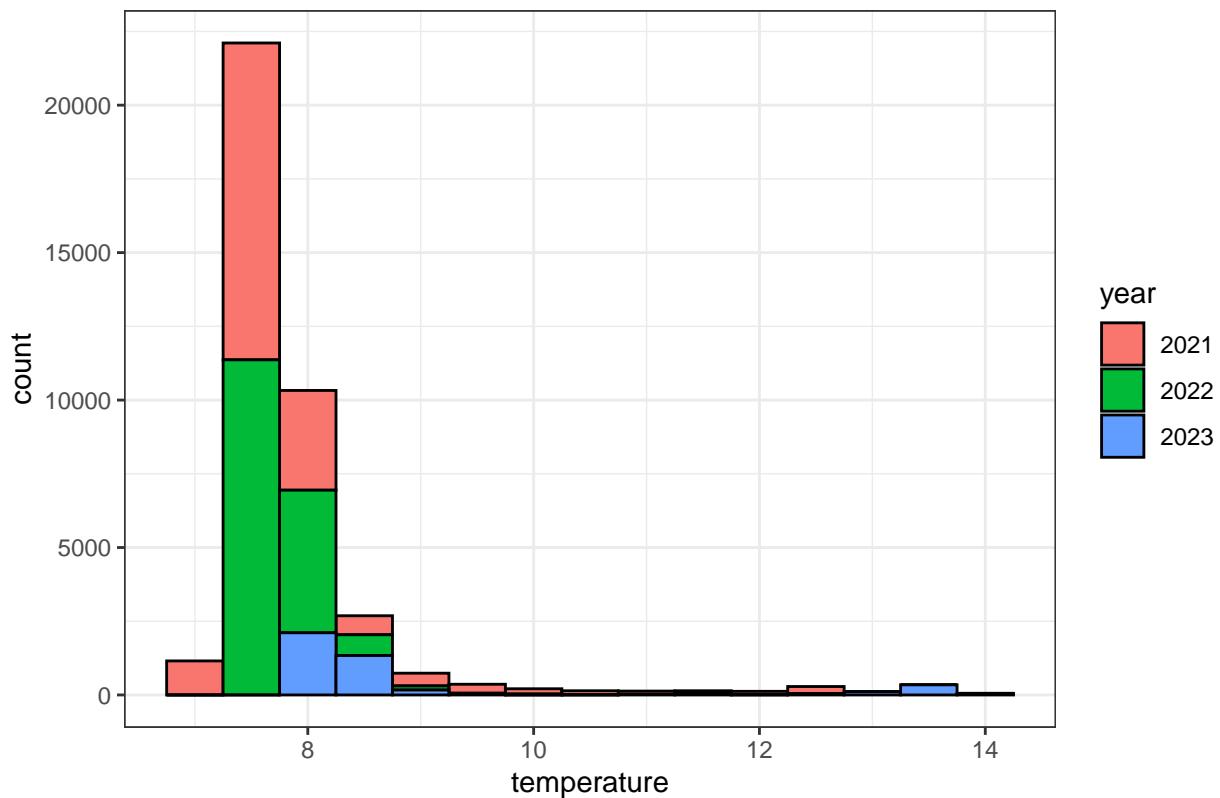
A histogram of depth in TH042 with binwidth 0.5



```
histogram_fill(TH042, temperature, 0.5, fill = year)
```

```
## Warning: Removed 4 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

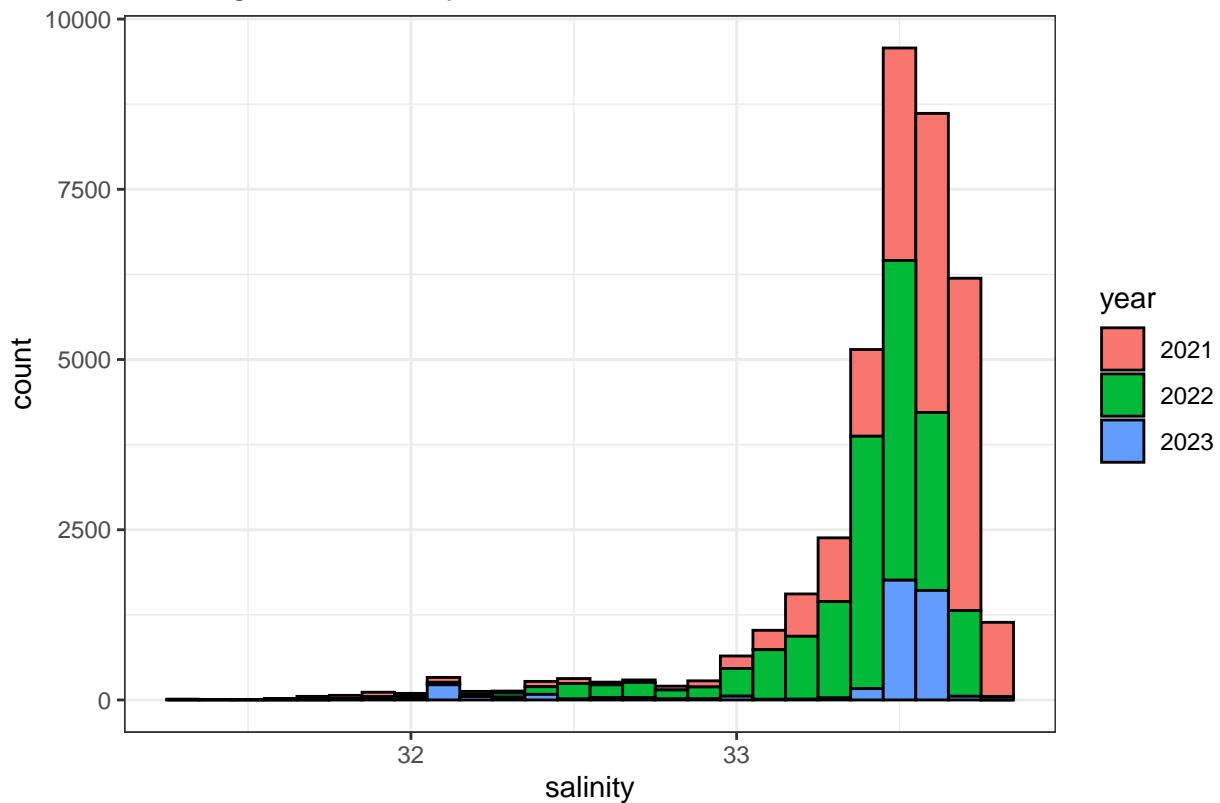
A histogram of temperature in TH042 with binwidth 0.5



```
histogram_fill(TH042, salinity, 0.1, fill = year)
```

```
## Warning: Removed 55 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

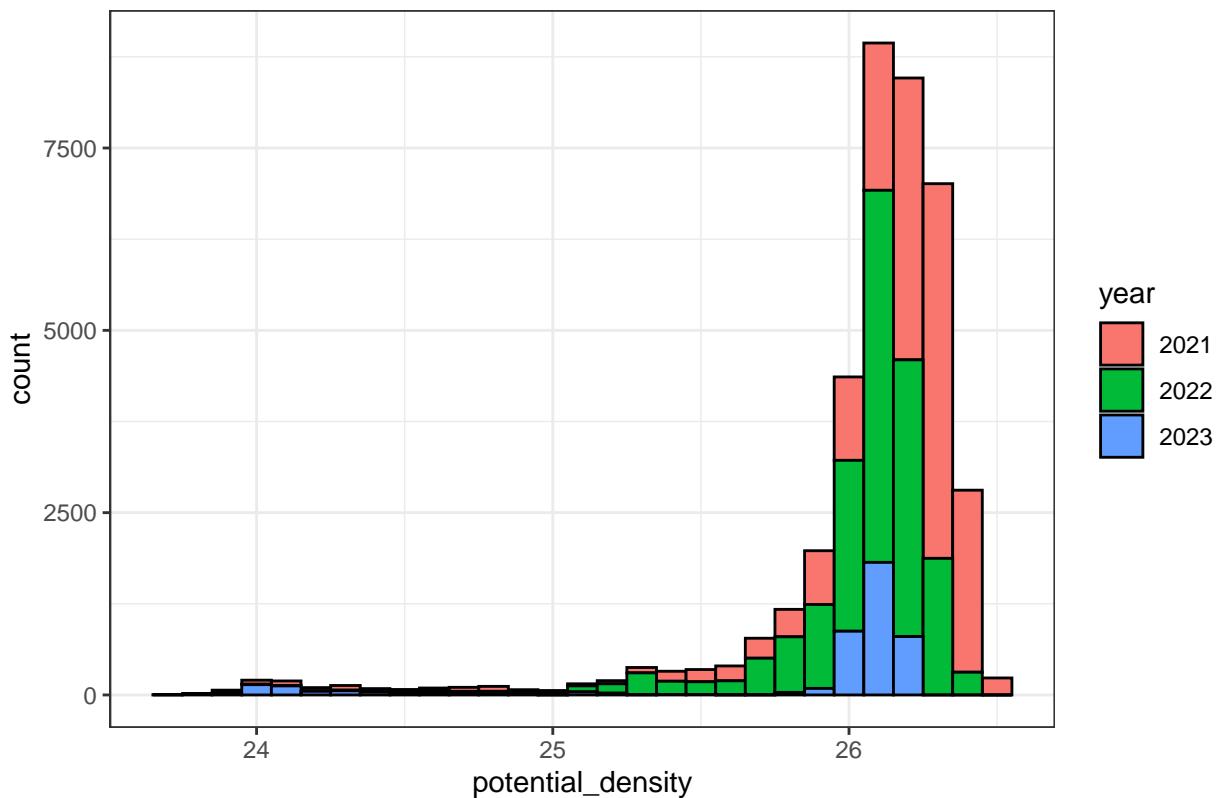
A histogram of salinity in TH042 with binwidth 0.1



```
histogram_fill(TH042, potential_density, 0.1, fill = year)
```

```
## Warning: Removed 55 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

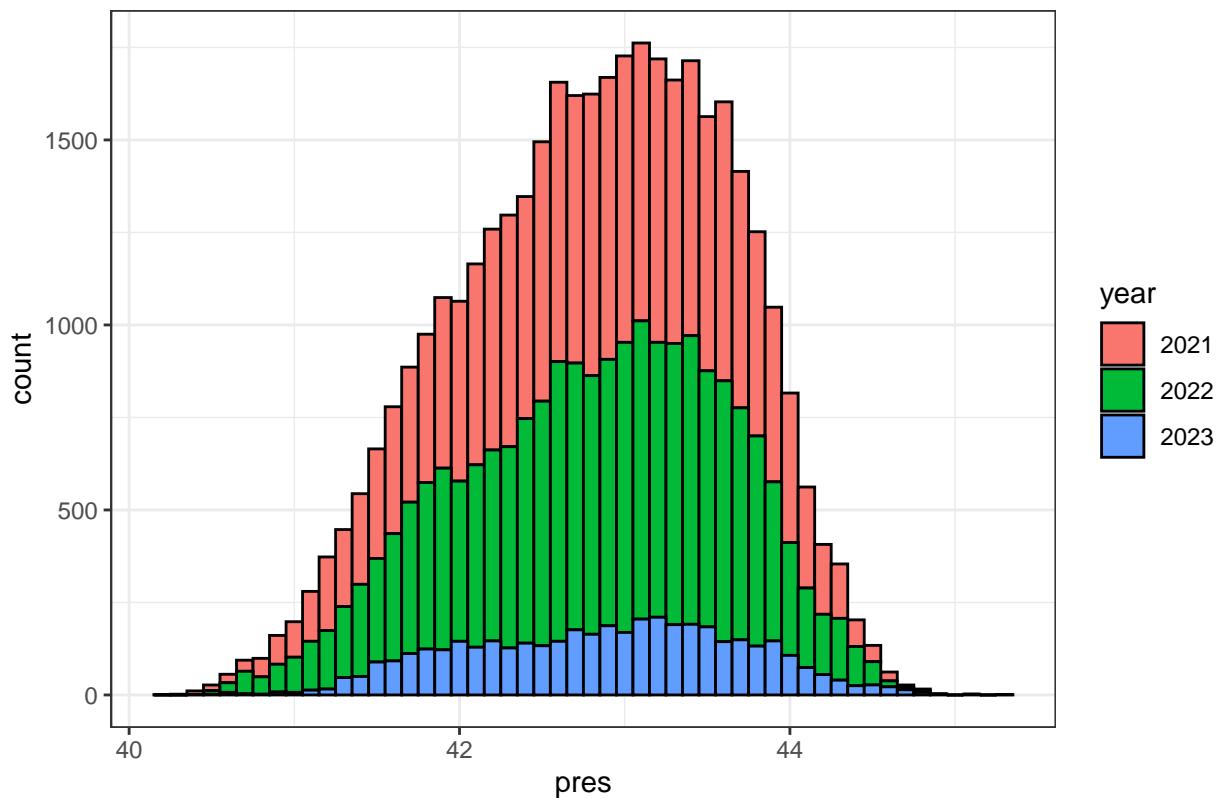
A histogram of potential\_density in TH042 with binwidth 0.1



```
histogram_fill(TH042, pres, 0.1, fill = year)

## Warning: Removed 4 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

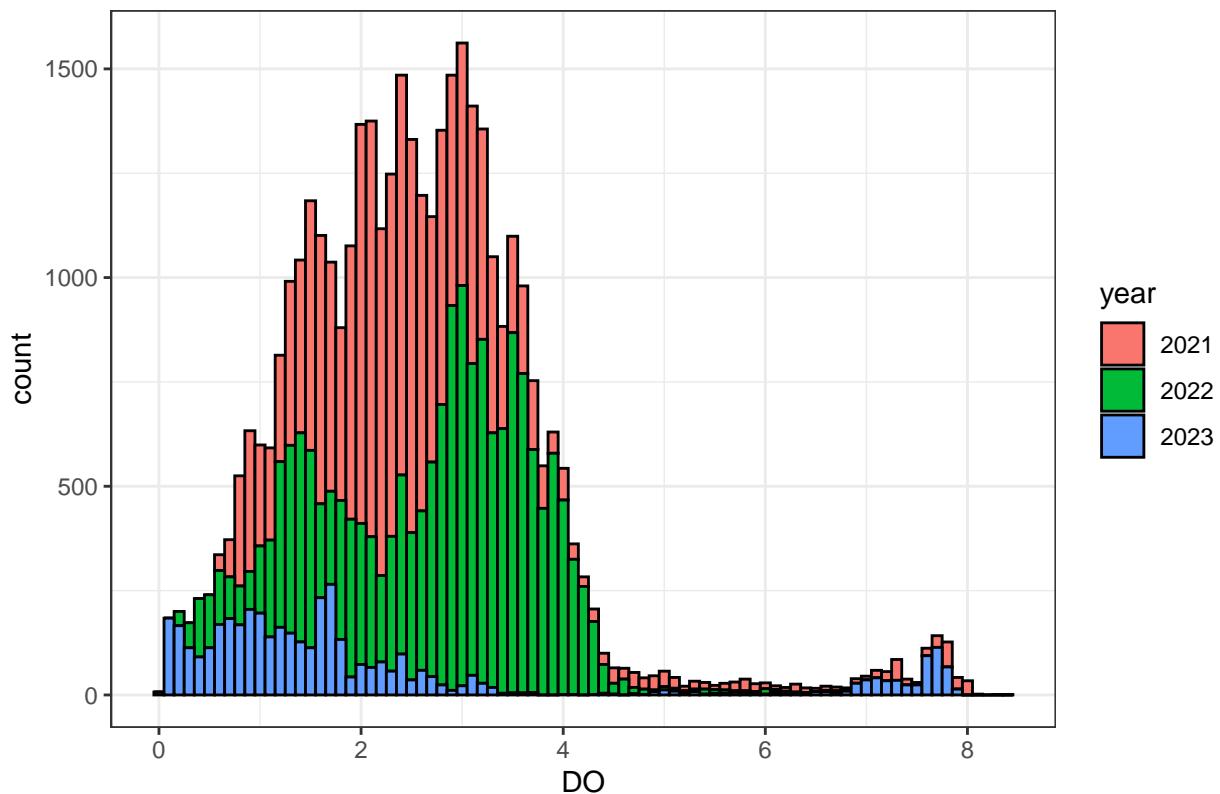
A histogram of pres in TH042 with binwidth 0.1



```
histogram_fill(TH042, D0, 0.1, fill = year)
```

```
## Warning: Removed 238 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

A histogram of DO in TH042 with binwidth 0.1

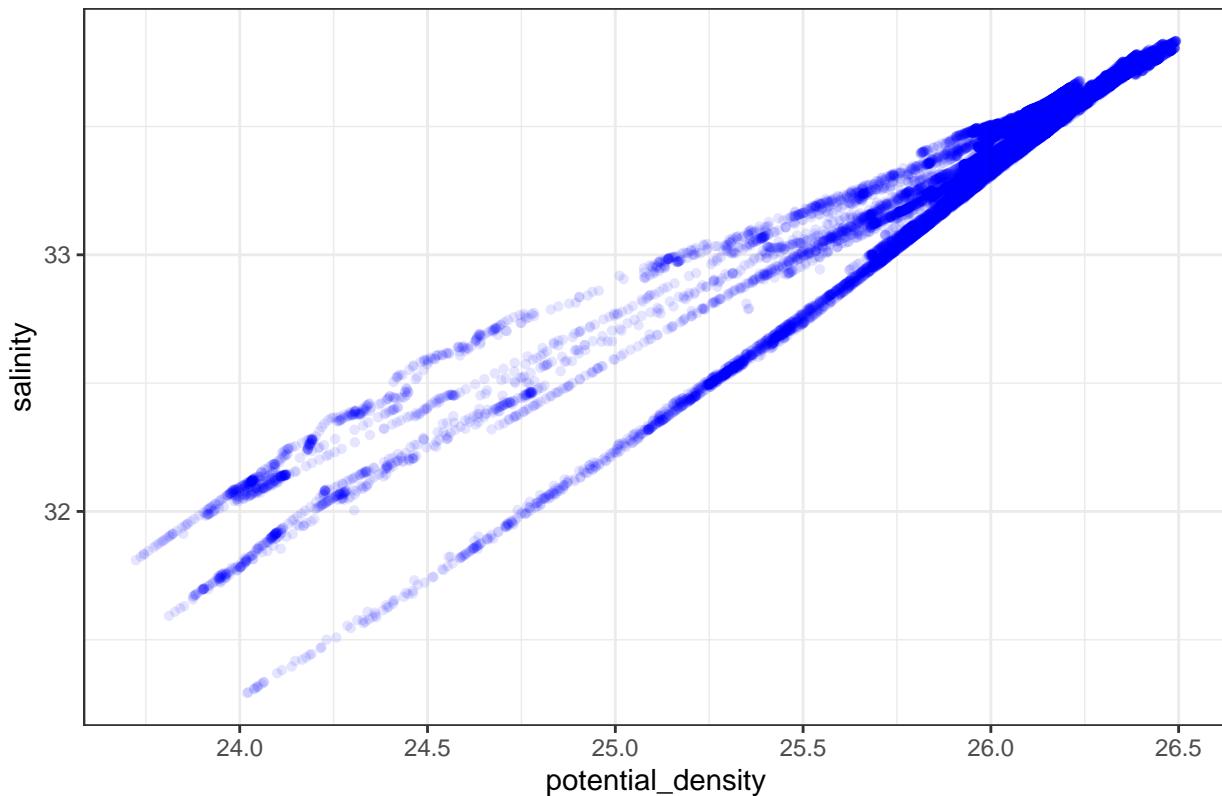


## Scatterplots

```
scatterplt(TH042, potential_density, salinity, alpha = 0.1)
```

```
## Warning: Removed 55 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

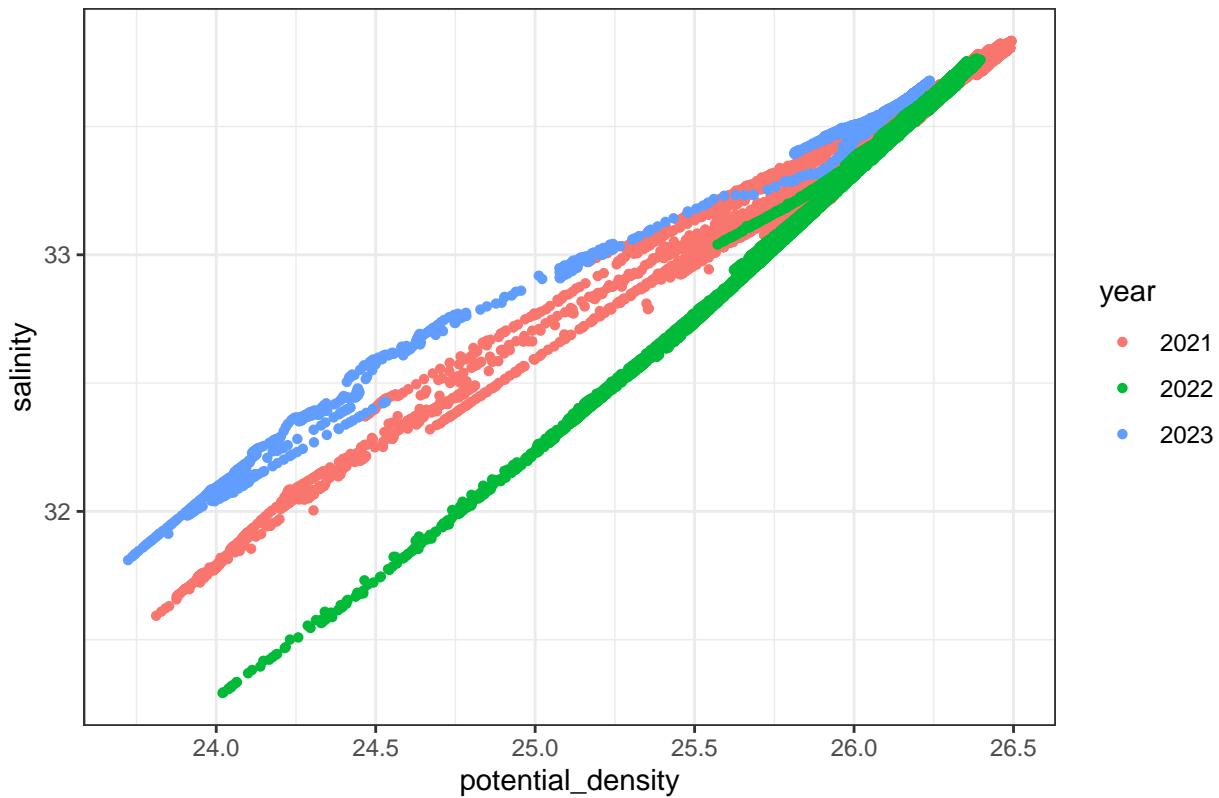
A scatterplot of potential\_density vs salinity in TH042



```
scatterplt_fill(TH042, potential_density, salinity, color = year)
```

```
## Warning: Removed 55 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

A scatterplot of potential\_density vs salinity in TH042



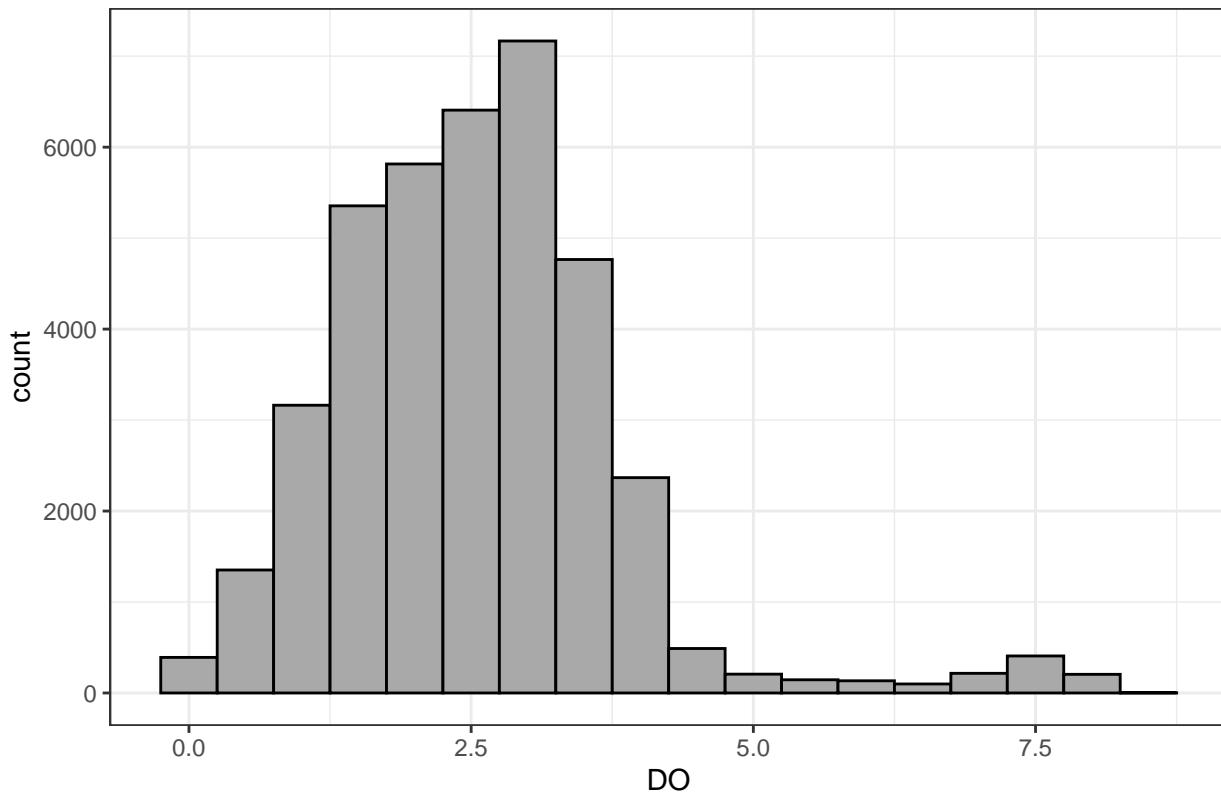
## Oxygen Data Exploration

### Histograms

```
histogram(TH042, D0, 0.5)
```

```
## Warning: Removed 238 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

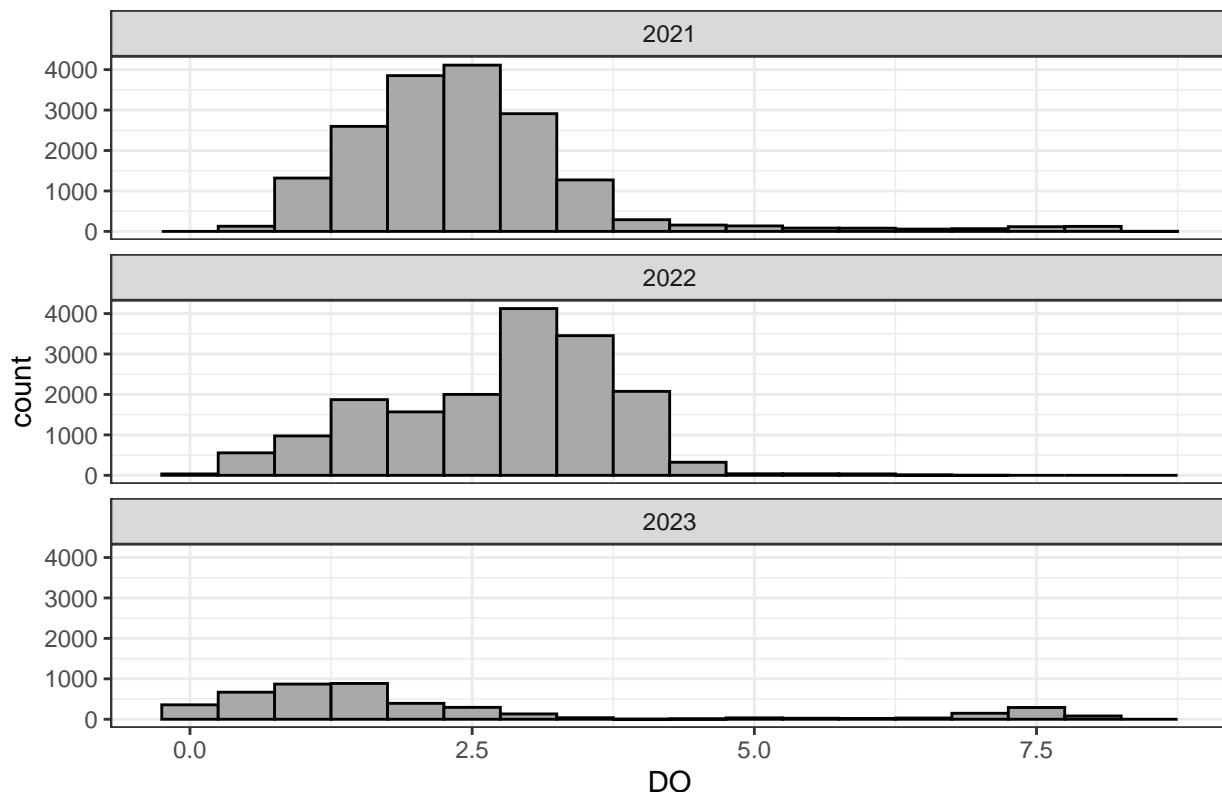
A histogram of DO in TH042 with binwidth 0.5



```
histogram(TH042, DO, 0.5) +
  facet_wrap(facets = vars(year), ncol = 1)

## Warning: Removed 238 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

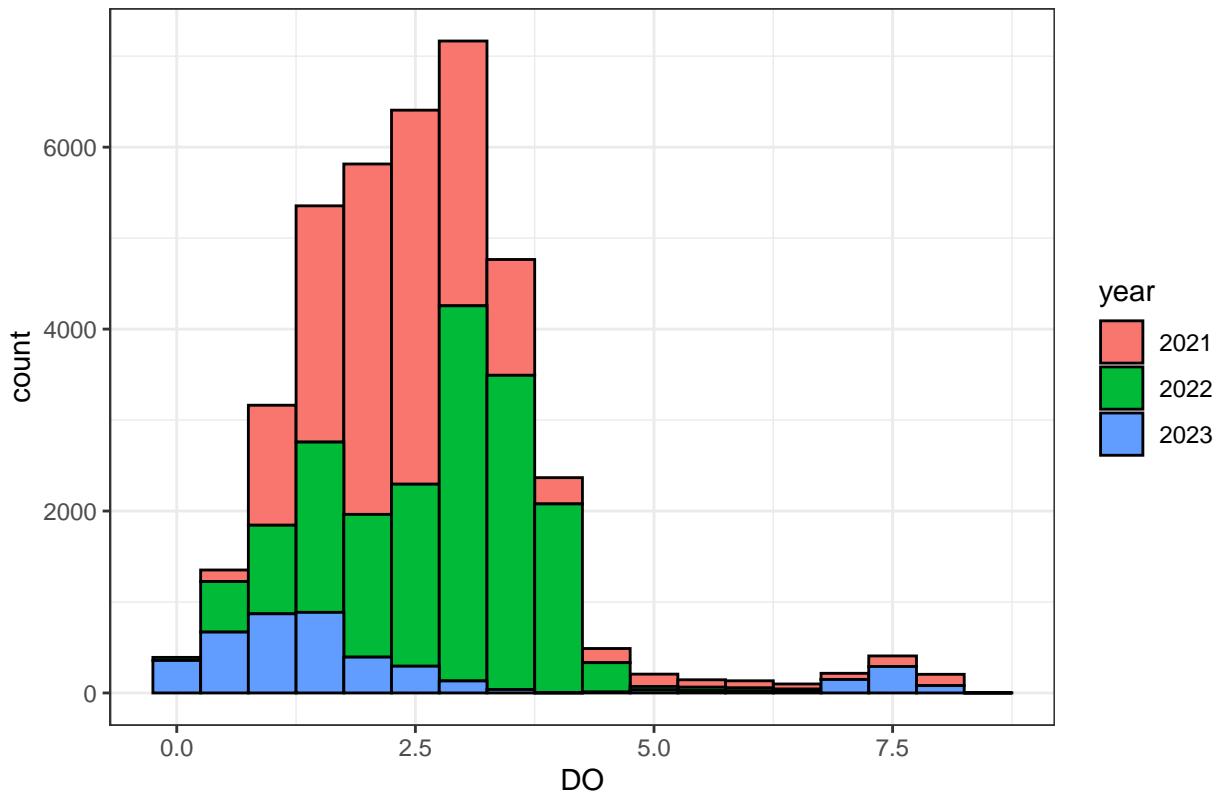
A histogram of DO in TH042 with binwidth 0.5



```
histogram_fill(TH042, DO, 0.5, fill = year)
```

```
## Warning: Removed 238 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

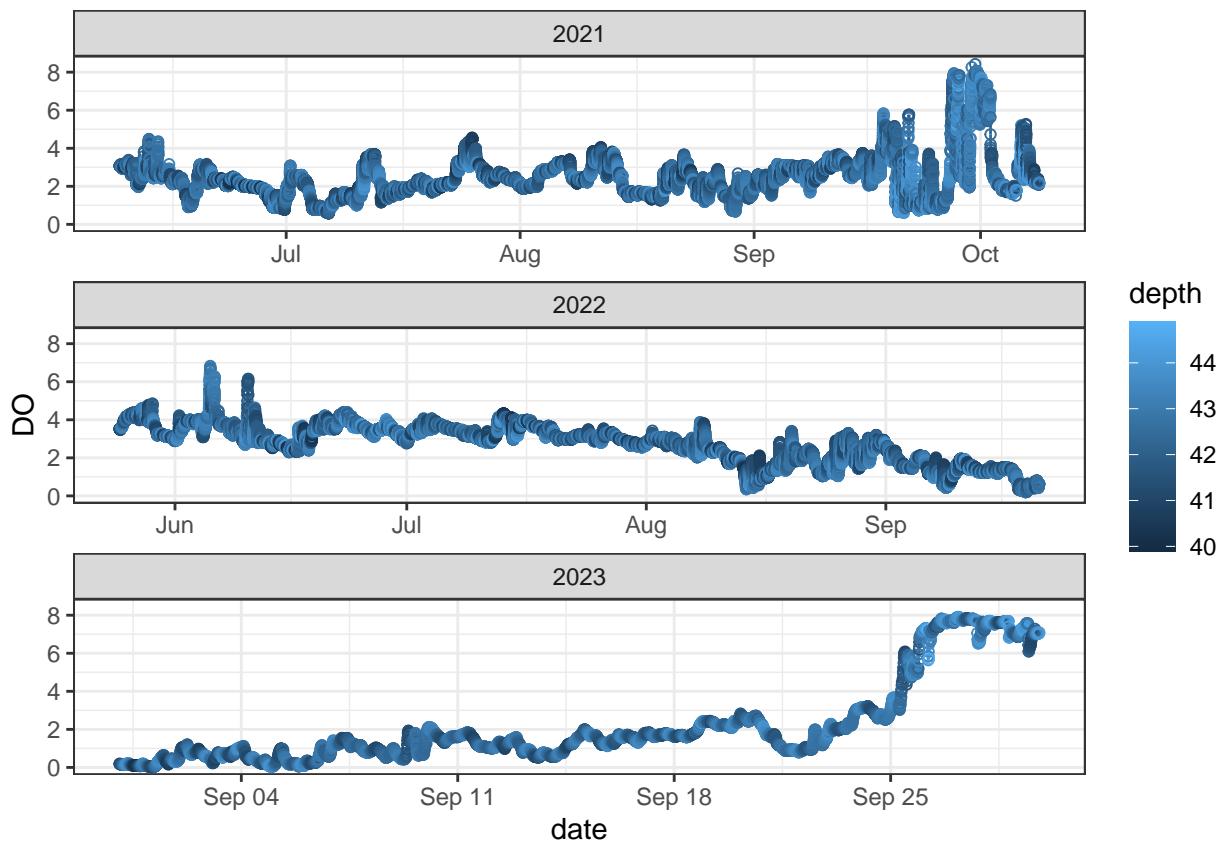
A histogram of DO in TH042 with binwidth 0.5



## Time Series

```
# Graph of oxygen levels around 42 meters for each year
TH042 %>%
  ggplot(aes(x = date, y = DO, color = depth)) +
  geom_line() +
  geom_point(shape = 1) +
  theme_bw() +
  facet_wrap(facets = vars(year), scales = "free_x", ncol = 1)

## Warning: Removed 238 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



```
ggsave("OCNMS_Mooring_Oxygen_FacetYear.png", path = here(pltpath), width = 3000, height =
  ↪ 4000, units = "px")
```

```
## Warning: Removed 238 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

## Export

```
# Add + reorder columns so that it can bind with the CTD data
TH042$sampleID = NA
TH042$source = "Mooring"
TH042 <- TH042 %>% select("source", "date", "year", "depth", "temperature", "DO",
  ↪ "salinity", "potential_density", "pres", "cond")

write.csv(TH042, file = here("OCNMS_Hypoxia", "Outputs", "OCNMS_Mooring_CleanData.csv"))
```