Data_Summary

Jessica French

2023-02-13

This notebook will show the locations that we have nutrient data for and a summary of other data we have aquired.

```
# Load in libraries
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   1.0.1
## v tibble  3.1.8     v dplyr   1.1.0
## v tidyr   1.3.0     v stringr 1.5.0
## v readr   2.1.3     v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(sf)
```

```
## Linking to GEOS 3.10.2, GDAL 3.4.2, PROJ 8.2.1; sf_use_s2() is TRUE
```

```
library(tmap)
library(ncdf4)
library(terra)
```

```
## terra 1.6.17
##
## Attaching package: 'terra'
##
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
library(janitor)
```

```
##
## Attaching package: 'janitor'
##
## The following object is masked from 'package:terra':
##
##     crosstab
```

```
##
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

Read in the boundary data

```
# Read in the shape file of the area of interest
AOI <- st_read("/Users/jfrench/Documents/MEDS/Capstone/DATA/AOI_SBchannel_shp/AOI_SBchannel.shp")
```

```
## Reading layer `AOI_SBchannel' from data source
##   `/Users/jfrench/Documents/MEDS/Capstone/DATA/AOI_SBchannel_shp/AOI_SBchannel.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 1 feature and 1 field
## Geometry type: POLYGON
## Dimension:     XY
## Bounding box:  xmin: -120.5 ymin: 33.83 xmax: -119.45 ymax: 34.49
## Geodetic CRS:  WGS 84
```

```
# Read in expanded AOI for filtering large data sets
Expanded_AOI <- st_read("/Users/jfrench/Documents/MEDS/Capstone/DATA/expanded_AOI/Expaned_AOI_SBchannel
```

```
## Reading layer `Expaned_AOI_SBchannel' from data source
##   `/Users/jfrench/Documents/MEDS/Capstone/DATA/expanded_AOI/Expaned_AOI_SBchannel.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 1 feature and 1 field
## Geometry type: POLYGON
## Dimension:     XY
## Bounding box:  xmin: -123.32 ymin: 30.59 xmax: -115.7 ymax: 36.08
## Geodetic CRS:  WGS 84
```

```
# Read in state and federal boundaries
federal_boundaries <- st_read("/Users/jfrench/Documents/MEDS/Capstone/DATA/maritime_boundaries/federal_
  st_filter(Expanded_AOI)
```

```
## Reading layer `USMaritimeLimitsNBoundaries' from data source
##   `/Users/jfrench/Documents/MEDS/Capstone/DATA/maritime_boundaries/federal_boundaries/USMaritimeLimi
##   using driver `ESRI Shapefile'
## Simple feature collection with 260 features and 16 fields
## Geometry type: MULTILINESTRING
## Dimension:     XY
## Bounding box:  xmin: -180 ymin: -17.555 xmax: 180 ymax: 74.70884
## Geodetic CRS:  WGS 84
```

```
# Read in California MPA boundaries
MPA_boundaries <- st_read("/Users/jfrench/Documents/MEDS/Capstone/DATA/MPAs/mpa_ca_4326 (1)/mpa_ca_4326
  st_filter(Expanded_AOI)
```

```
## Reading layer `mpa_ca_4326Polygon' from data source
```

```
##   `/Users/jfrench/Documents/MEDS/Capstone/DATA/MPAs/mpa_ca_4326 (1)/mpa_ca_4326Polygon.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 146 features and 7 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -124.5326 ymin: 32.53283 xmax: -117.1238 ymax: 42.00003
## Geodetic CRS:  WGS 84
```

Read in the nutrient data

- CalCOFI data has nitrogen species and phosphorous goes back past 2014

```
# Read in CalCOFI cast data
CalCOFI_cast <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/Nutrients/CalCOFI_Database_194903
  filter(Year >= 2014) |>
  select(c("Cst_Cnt", "Quarter", "Year", "Lat_Dec", "Lon_Dec"))
```

```
## Rows: 35376 Columns: 61
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (13): Cruise_ID, Cast_ID, Sta_ID, Sta_Code, Date, Lat_Hem, Lon_Hem, Shi...
## dbl  (43): Cst_Cnt, Cruise, Cruz_Sta, DbSta_ID, Quarter, Distance, Year, Mon...
## time  (5): Time, Inc_Str, Inc_End, PST_LAN, Civil_T
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# Read in the CalCOFI bottle data
CalCOFI_bottle <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/Nutrients/CalCOFI_Database_1949
  right_join(CalCOFI_cast, by = "Cst_Cnt")|>
  st_as_sf(coords = c(lon = "Lon_Dec", lat = "Lat_Dec"), crs = "EPSG: 4326") |>
  st_filter(Expanded_AOI)
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)

## Rows: 889500 Columns: 62
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (2): Sta_ID, Depth_ID
## dbl  (54): Cst_Cnt, Btl_Cnt, Depthm, T_degC, Salnty, O2ml_L, STheta, O2Sat, ...
## lgl   (5): DIC2, TA2, pH1, pH2, DIC Quality Comment
## time  (1): IncTim
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

- The LTER Biomass and environmental drivers dataset only has nitrogen and wave action
- Goes back past 2014

```r
# Read in the LTER Biomass and Nitrogen Data Set
LTER_nutrients_biomass <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/Nutrients/Nitrate_LTER/
  filter(year >= 2014)
```

```
## Rows: 47652 Columns: 6
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (6): site_id, year, quarter, kelp, no3, waves
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
# Read in the LTER sampling locations
LTER_locations <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/Nutrients/Nitrate_LTER/locations
```

```
## Rows: 361 Columns: 3
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (3): site_id, lat, lon
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
# Combine the locations to the nutrient data
lter_data_geom <- left_join(LTER_nutrients_biomass, LTER_locations, by = "site_id") |>
  st_as_sf(coords = c(lon = "lon", lat = "lat"), crs = "EPSG:4326")
```

- Plumes and Blooms Data has nitrogen species and phosphorous
- From before 1995 to 2016

```r
# Read in the plumes and blooms data
plumes_blooms <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/Nutrients/plumes_blooms/Plumes_a
  mutate("DATE" = lubridate::ymd(DATE),
         "LON" = LON*-1)|>
  filter(DATE >= 2014-01-01,
         LON < 100,
         LAT >= 30 & LAT <= 40) |>
  st_as_sf(coords = c(lon = "LON", lat = "LAT"), crs = "EPSG:4326")
```

```
## Rows: 2964 Columns: 54
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (1): DATA_FILE_NAME
## dbl (53): ID, STATION, DATE, LAT, LON, DEPTH, SAL, POTEMP068, SIG00, SIGT00,...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

- USGS Bottle Data from OCT and NOV 2018 and 2019

```
# Read in USGS bottle data 2018
USGS_2018 <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/usgs_bottle_fall/SH-18-12_BTL_CTD_da
  st_as_sf(coords = c("Longitude", "Latitude"),
           crs = 4326) |>
  st_filter(Expanded_AOI)
```

```
## Rows: 161 Columns: 36
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (5): Cruise, Site Name, Date, Ammonium, Conventional Radiocarbon Age
## dbl  (30): CTD Station, TimeJ, Temperature, Conductivity, Pressure, Sbeox0, ...
## time  (1): Time
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# Read in USGS bottle data 2019
USGS_2019 <- read_csv("/Users/jfrench/Documents/MEDS/Capstone/DATA/usgs_bottle_fall/RL-19-05_BTL_CTD_da
  st_as_sf(coords = c("Longitude", "Latitude"),
           crs = 4326) |>
  st_filter(Expanded_AOI)
```

```
## Rows: 131 Columns: 29
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (3): Cruise, Site Name, Date
## dbl  (25): CTD Station, TimeJ, Temperature, Conductivity, Pressure, Sbeox0, ...
## time  (1): Time
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

- LTER Water Chemistry data

```
# Read in the Bottle file from LTER

LTER_bottle <- read_csv2("/Users/jfrench/Documents/MEDS/Capstone/DATA/Nutrients/LTER_bottle/LTER_monthly
  clean_names() |>
  filter(longitude_e < 0) |>
  select(po4_umol_l, no2_no3_umol_l, latitude_n, longitude_e)
```

```
## i Using "',' as decimal and "'.'" as grouping mark. Use `read_delim()` for more control.

## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)

## Rows: 3379 Columns: 51
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ";"
## chr  (29): Cruise, Station, Type, Actual Depth Z (m), Bottle ID, PO4 (umol/l...
## dbl  (12): Bot_Depth (m), Target Depth (m), Consecutive Sample #, Mean TCO2 ...
## num   (8): Longitude (E), Latitude (N), decimal year, TDN (umol/L), POC (umo...
## date  (1): yyyy-mm-dd
## time  (1): hh:mm
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

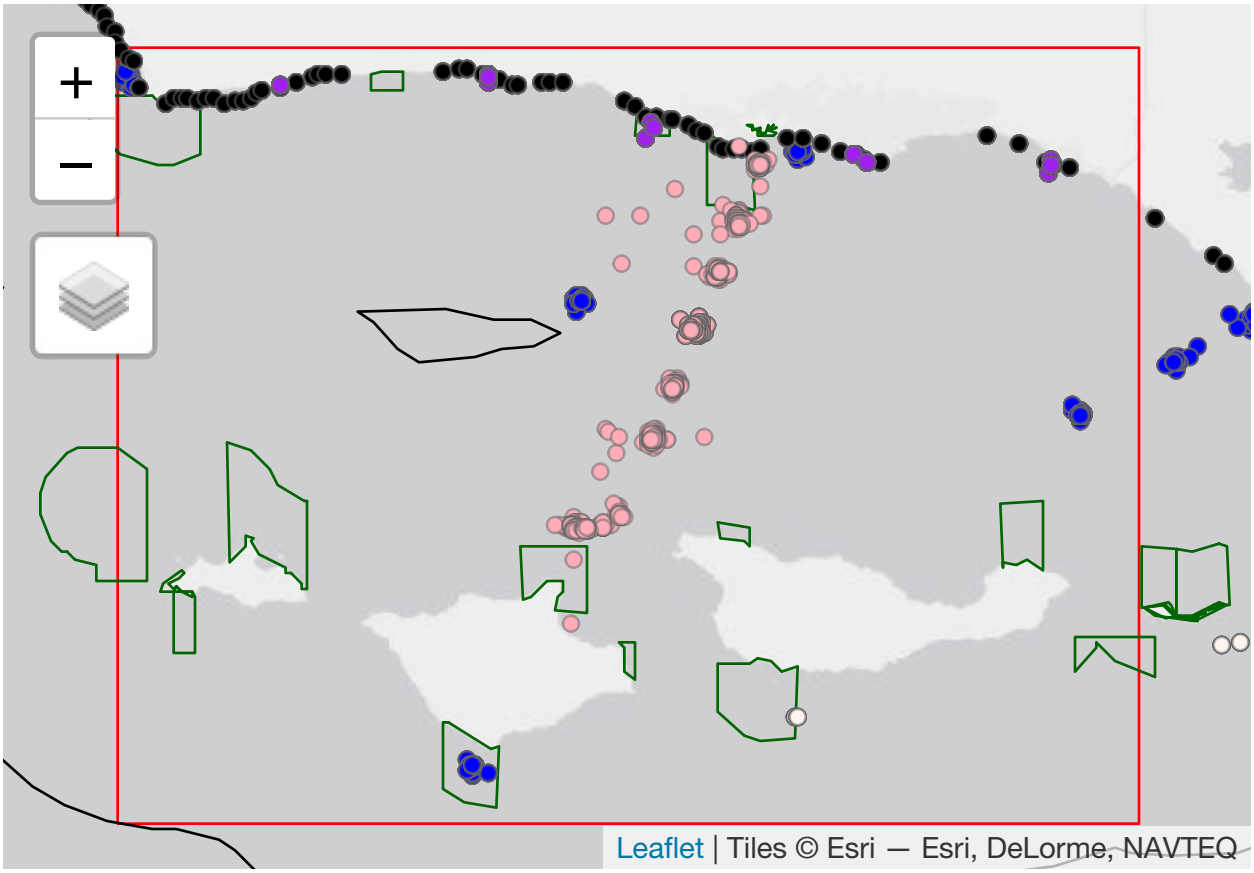
```
stringi::stri_sub(LTER_bottle$latitude_n, 3, 2) = '.'
stringi::stri_sub(LTER_bottle$longitude_e, 5, 4) <- '.'
#Encoding(LTER_bottle$comments <- "UTF-8")

LTER_bottle <- LTER_bottle |>
  mutate("latitude_n" = as.numeric(latitude_n),
         "longitude_e" = as.numeric(longitude_e)) |>
  st_as_sf(coords = c("longitude_e", "latitude_n"),
           crs = 4326)
```

```
# Test map
tmap_mode(mode = "view")
```

```
## tmap mode set to interactive viewing
```

```
tm_shape(AOI) +
  tm_borders(col = "red") +
  tm_shape(federal_boundaries) +
  tm_lines() +
  tm_shape(MPA_boundaries) +
  tm_borders(col = "darkgreen") +
  tm_shape(CalCOFI_bottle) +
  tm_dots(col = "blue") +
  tm_shape(lter_data_geom) +
  tm_dots() +
  tm_shape(plumes_blooms) +
  tm_dots(col = "lightpink1") +
  tm_shape(USGS_2018) +
  tm_dots(col = "seashell") +
  tm_shape(USGS_2019) +
  tm_dots(col = "seashell") +
  tm_shape(LTER_bottle) +
  tm_dots(col = "purple")
```


Leaflet | Tiles © Esri — Esri, DeLorme, NAVTEQ