multi-year DaRTS Data

Julia Brown

2025-09-12

Research Question: How does picoplankton concentraion change with temperature?

```
library('tidyverse')
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.2     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.1.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library('ggplot2')
library('tibble')
library('dplyr')
library('lubridate')
```

```
ctdData <- read.csv('data/DaRTS_CTD_data_2012-2024.csv')
discreteData <- read.csv('data/DaRTS_discrete_data_2012-2024.csv')
```

```
head(ctdData)
```

```
##       Date Station Depth_m Conductivity Temperature_C Salinity_PSU Density
## 1 20120911       1   1.363           NA       16.8404      31.4819 22.8529
## 2 20120911       1   1.346           NA       16.8389      31.4812 22.8527
## 3 20120911       1   1.345           NA       16.8371      31.4801 22.8523
## 4 20120911       1   1.354           NA       16.8330      31.4811 22.8540
## 5 20120911       1   1.360           NA       16.8310      31.4815 22.8548
## 6 20120911       1   1.362           NA       16.8287      31.4824 22.8560
##   PAR.Irradiance Fluorescence TurbidityNTU BeamC   O2Conc O2Saturation pH X X.1
## 1           1780       3.0915       1.3585    NA 167.9923     67.05656       NA
## 2           1710       3.1327       1.3913    NA 165.2259     65.95014       NA
## 3           2010       3.1777       1.4420    NA 162.0110     64.66411       NA
## 4           1690       3.2303       1.5023    NA 158.8182     63.38518       NA
## 5           1720       3.1525       1.4535    NA 156.1492     62.31767       NA
## 6           1500       2.9976       1.3543    NA 154.6120     61.70184       NA
##   X.2 X.3 X.4 X.5 X.6 X.7 X.8 X.9 X.10
```

```
## 1  NA  NA  NA  NA  NA  NA  NA  NA   NA
## 2  NA  NA  NA  NA  NA  NA  NA  NA   NA
## 3  NA  NA  NA  NA  NA  NA  NA  NA   NA
## 4  NA  NA  NA  NA  NA  NA  NA  NA   NA
## 5  NA  NA  NA  NA  NA  NA  NA  NA   NA
## 6  NA  NA  NA  NA  NA  NA  NA  NA   NA
```

```
head(discreteData)
```

```
##      Date Cruise Station          Lat         Long Depth pH Alk SiO4_uM
## 1 9/11/12      1       1   43.54.20 N   69.34.66 W     2 NA  NA      NA
## 2 9/11/12      1       1   43.54.20 N   69.34.66 W    10 NA  NA      NA
## 3 9/11/12      1       1   43.54.20 N   69.34.66 W    30 NA  NA      NA
## 4 9/11/12      1       3 43. 48. 65 N 69. 34. 10 W     2 NA  NA      NA
## 5 9/11/12      1       3 43. 48. 65 N 69. 34. 10 W    10 NA  NA      NA
## 6 9/11/12      1       3 43. 48. 65 N 69. 34. 10 W    40 NA  NA      NA
##   NO3.2_uM PO4_uM NH4_uM Pico_Plankton_cells_per_ml Bacteria_cells_per_ml
## 1     0.00   0.49   0.05                      5618               1870000
## 2     0.00   0.56   0.02                      6299               2330000
## 3     0.86   0.61   0.25                      6016               1740000
## 4     0.00   0.14   0.18                      1076               1600000
## 5     0.00   0.32   0.21                     14582               1820000
## 6     9.00   1.16   4.85                     20457               2110000
##   Virus_per_ml Total_Chl_ug_per_l Chl_less_than_20um_ug_per_l
## 1     45100000              9.119                       7.929
## 2     44700000                 NA                          NA
## 3     36600000                 NA                          NA
## 4     31000000              3.615                       2.079
## 5     36000000                 NA                          NA
## 6     51100000                 NA                          NA
##   Chl_less_than_3um_ug_per_l Flowcam_Biomass cyanobacteria_per_ml
## 1                      0.222          224000                  722
## 2                         NA              NA                  574
## 3                         NA              NA                  500
## 4                      0.413          231000                  673
## 5                         NA              NA                 3551
## 6                         NA              NA                 6157
##   picoeukaryotes_per_ml noeukaryotes_per_ml
## 1                  3817                1079
## 2                  4365                1361
## 3                  3618                1898
## 4                   254                 149
## 5                  8847                2184
## 6                 12107                2193
```

```
ctdDates <- ctdData$Date
class(ctdDates[1])
```

```
## [1] "integer"
```

```
discreteDates <- discreteData$Date
discreteDates[1]
```

```
## [1] "9/11/12"
```

```
discreteDates <- lubridate::mdy(discreteData$Date)
discreteDates[1]
```

Lubridate package to handle date values

```
## [1] "2012-09-11"
```

```
class(discreteDates[1])
```

```
## [1] "Date"
```

```
ctdData <- mutate(ctdData, Date = lubridate::ymd(ctdData$Date))

discreteData <- mutate(discreteData, Date = lubridate::mdy(discreteData$Date))

ctdData <- mutate(ctdData,
                  year = year(Date),
                  month = month(Date),
                  day = day(Date))

discreteData <- mutate(discreteData,
                       year = year(Date),
                       month = month(Date),
                       day = day(Date))

ctdData <- mutate(ctdData, doy = yday(Date))
discreteData <- mutate(discreteData, doy = yday(Date))
```

Adding new columns to data frame with the "mutate" function

Depths

```
ctdData$Depth_m[1:10]
```

```
##  [1] 1.363 1.346 1.345 1.354 1.360 1.362 1.355 1.342 1.342 1.356
```

```
discreteData$Depth[1:10]
```

```
##  [1] "2"  "10" "30" "2"  "10" "40" "2"  "10" "40" "2"
```

```
ctdData <- mutate(ctdData, Depth = round(ctdData$Depth_m))

head(ctdData)
```

```
##         Date Station Depth_m Conductivity Temperature_C Salinity_PSU Density
## 1 2012-09-11       1   1.363           NA       16.8404      31.4819 22.8529
## 2 2012-09-11       1   1.346           NA       16.8389      31.4812 22.8527
## 3 2012-09-11       1   1.345           NA       16.8371      31.4801 22.8523
## 4 2012-09-11       1   1.354           NA       16.8330      31.4811 22.8540
## 5 2012-09-11       1   1.360           NA       16.8310      31.4815 22.8548
## 6 2012-09-11       1   1.362           NA       16.8287      31.4824 22.8560
##   PAR.Irradiance Fluorescence TurbidityNTU BeamC   O2Conc O2Saturation pH X X.1
## 1           1780       3.0915       1.3585    NA 167.9923     67.05656       NA
## 2           1710       3.1327       1.3913    NA 165.2259     65.95014       NA
## 3           2010       3.1777       1.4420    NA 162.0110     64.66411       NA
## 4           1690       3.2303       1.5023    NA 158.8182     63.38518       NA
## 5           1720       3.1525       1.4535    NA 156.1492     62.31767       NA
## 6           1500       2.9976       1.3543    NA 154.6120     61.70184       NA
##   X.2 X.3 X.4 X.5 X.6 X.7 X.8 X.9 X.10 year month day doy Depth
## 1  NA  NA  NA  NA  NA  NA  NA  NA   NA 2012     9  11 255     1
## 2  NA  NA  NA  NA  NA  NA  NA  NA   NA 2012     9  11 255     1
## 3  NA  NA  NA  NA  NA  NA  NA  NA   NA 2012     9  11 255     1
## 4  NA  NA  NA  NA  NA  NA  NA  NA   NA 2012     9  11 255     1
## 5  NA  NA  NA  NA  NA  NA  NA  NA   NA 2012     9  11 255     1
## 6  NA  NA  NA  NA  NA  NA  NA  NA   NA 2012     9  11 255     1
```

```
ctdDataGrouped <- group_by(ctdData, Date, Station, Depth)

ctdDataBinned <- summarize_all(ctdDataGrouped,
                               mean, na.rm = TRUE)
```

```
## Warning: There were 31666 warnings in `summarise()`.
## The first warning was:
## i In argument: `pH = (function (x, ...) ...`.
## i In group 1: `Date = 2012-09-11`, `Station = "1"`, `Depth = 1`.
## Caused by warning in `mean.default()`:
## ! argument is not numeric or logical: returning NA
## i Run `dplyr::last_dplyr_warnings()` to see the 31665 remaining warnings.
```

```
discreteData$Station <- as.character(discreteData$Station)
discreteData$Depth <- as.numeric(discreteData$Depth)
```

```
## Warning: NAs introduced by coercion
```

Merging DataFrames

```
combinedData <- full_join(ctdDataBinned, discreteData)
```

```
## Joining with `by = join_by(Date, Station, Depth, pH, year, month, day, doy)`
```