Reading and processing data for use in the LTER-LIFE digital twins: remote sensing data

Karline Soetaert

first version: 01-10-2023; current version: 30 January 2024

This document prepares remote sensing data that are to be used in the LTER life dtPP (digital twin of primary production) package.

# Remote sensing Chlorophyll

Converts monthly chlorophyll data that were obtained from <http://sites.science.oregonstate.edu/ocean.productivity>

NOTE: These data are not the final data to be used; they should be eventually replaced.

## Reading the data

Monthly data are in separate files; they are in HDF format. A function is written to process one file

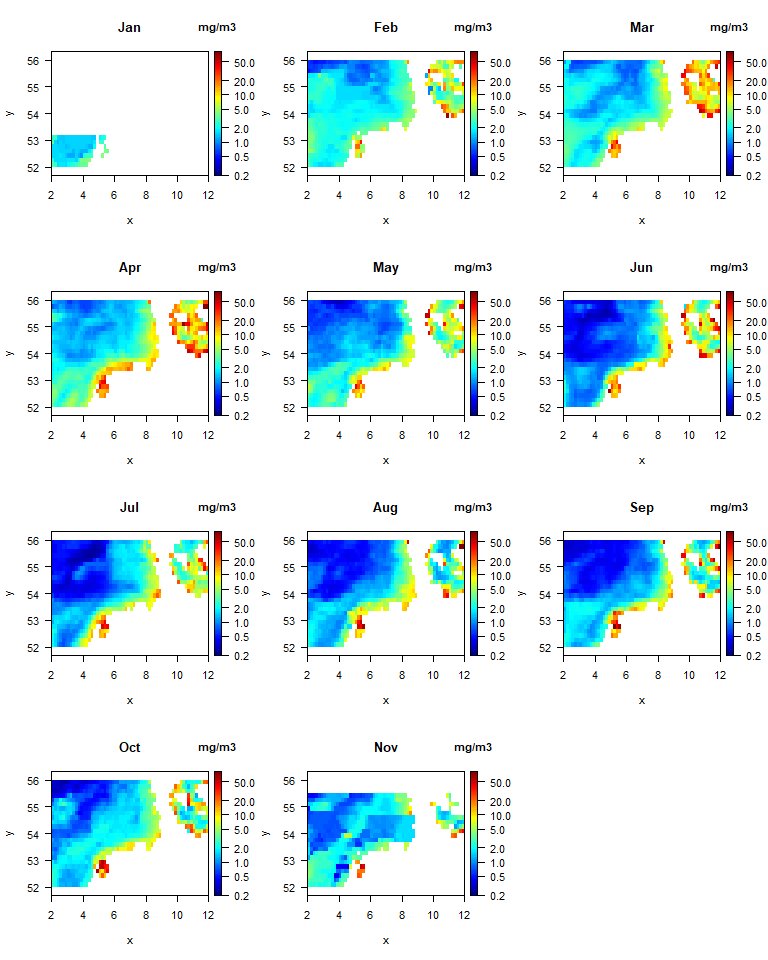
ReadData <- function(file, dir="../raw\_data/chl\_RS", plotit=TRUE){  
   
 chl.2021 <- nc\_open(paste(dir, file, sep="/"))  
   
 # read the chlorophyll variable  
 CHL <- ncvar\_get(chl.2021, "chl")   
   
 # read global attributes  
 tini <- ncatt\_get(chl.2021, varid= 0,   
 attname = "Start Time String")  
 tend <- ncatt\_get(chl.2021, varid= 0,   
 attname = "Stop Time String")  
   
 # Clean up data (flag unavailable data)  
 CHL[CHL == -9999] <- NA  
  
 # Latitude and longitude - they are not given  
 dx <- 180/1080  
 dy <- 360/2160  
   
 LAT <- seq( -90+dx/2, 90-dx/2, length.out=1080)  
 LON <- seq(-180+dy/2, 180-dy/2, length.out=2160)  
  
 # Select a region that includes the Wadden   
 LON.ii <- which(LON >= 2 & LON <= 12)  
 LAT.ii <- which(LAT >= 52 & LAT <= 56)  
  
 # aspect ratio for plotting  
 asp <- 1/cos((mean(LAT[LAT.ii]) \* pi)/180)  
  
 # convert in proper format and select the wadden area  
 Chl <- CHL[, ncol(CHL):1]  
  
 ChlorWad <- Chl[LON.ii, LAT.ii]  
  
 list(x =LON[LON.ii],   
 y =LAT[LAT.ii],   
 z =ChlorWad,   
 tini =tini[[2]],   
 tend =tend[[2]],   
 asp =asp)  
}

The data files are read one by one and put in 3D array

# names of the files  
  
HDFfiles <- c("chl.2021001.hdf", "chl.2021032.hdf", "chl.2021060.hdf",   
 "chl.2021091.hdf", "chl.2021121.hdf", "chl.2021152.hdf",   
 "chl.2021182.hdf", "chl.2021213.hdf", "chl.2021244.hdf",   
 "chl.2021274.hdf", "chl.2021305.hdf", "chl.2021335.hdf")  
nF <- length(HDFfiles)  
  
# read a file  
X <- ReadData ("chl.2021213.hdf")  
DD <- c(dim(X$z), 12) # dimension of the array  
CHLarr <- array(dim=DD) # create an empty array  
  
Ti <- NULL  
Te <- NULL  
  
for (ifile in 1:nF) {  
 X <- ReadData (HDFfiles[ifile])  
   
 Ti <- c(Ti, X$tini)  
 Te <- c(Te, X$tend)  
 CHLarr[,,ifile] <- X$z  
}  
  
frmt <- "%m/%d/%Y %H:%M:%OS" # format of the datetime strings  
  
Chlor2021 <- list(longitude = X$x,   
 latitude = X$y,   
 chlorophyll = CHLarr,   
 datetime = data.frame(ini=as.POSIXct(Ti, format=frmt),   
 end=as.POSIXct(Te, format=frmt)),  
 datasource = "http://orca.science.oregonstate.edu/",   
 asp = X$asp)  
  
save(file="../processed\_data/Chlor2021.rda", Chlor2021)  
  
ChlMean <- apply(Chlor2021$chlorophyll, MARGIN=3, FUN=mean, na.rm=TRUE)

## plot the data

months <- 1:11  
image2D(x = Chlor2021$longitude, y=Chlor2021$latitude,   
 z = Chlor2021$chlorophyll[,,months],   
 main = month.abb[months], clab = "mg/m3",   
 log = "c", # color variable log-transformed  
 asp = Chlor2021$asp,  
 mfrow = c(4,3), clim=c(0.2,80), las=1)



# References

The following R-sources were used for this work:

R-core:

* R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

R-package dtWad,

* Soetaert K (2024). dtWad: Waddensea Digital Twin: general utilities. R package version 0.0.1.

R-package ncdf4

* Pierce D (2023). *ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files*. R package version 1.21, <https://CRAN.R-project.org/package=ncdf4>

R-package plot3D

* Soetaert K (2021). *plot3D: Plotting Multi-Dimensional Data*. R package version 1.4, <https://CRAN.R-project.org/package=plot3D>.