Earth System
Science
Data

Supplement of

Global dataset on seagrass meadow structure, biomass and production

Simone Strydom et al.

Correspondence to: Oscar Serrano (oserrano@ceab.csic.es)

###--- Aim: create global maps & boxplots for Seagrass Global Database project ---###
(https://doi.pangaea.de/10.1594/PANGAEA.929968)
 # Code written by Charlie M. Phelps & Roisin McCallum, figures produced by Simone Strydom

```r
library(tidyverse)
require(leaflet)
library(leaflet)
library(RColorBrewer)
library(dplyr)
library(leaflet.minicharts)
library(raster)

setwd("C:/Users/simonestrydom/OneDrive - Department of Biodiversity, Conservation and Attractions/ECU/ECU
backup/Documents/My Papers/CMER macrophyte manual/ Updated2022/Update Graphs")

###--- Leaflet maps---###

data<- read.csv("CMERpoints.csv",  na.strings = "0")
tilesURL <-
"http://server.arcgisonline.com/ArcGIS/rest/services/Canvas/World_Light_Gray_Base/MapServer/tile/{z}/{y}/{x}"
basemap <- leaflet(width = "100%", height = "700px") %>% addTiles(tilesURL); basemap
pal <- colorFactor(c( "honeydew4","burlywood4", "yellowgreen", "magenta4",
            "aquamarine4",  "black",  "darkcyan",
            "lightcoral",  "darkblue",  "darkorange2",  "purple1","darkgreen"
), domain = c('Amphibolis',      'Cymodocea',  'Enhalus',      'Halodule',      'Halophila',
        'Phyllospadix',  'Posidonia',     'Ruppia',       'Syringodium',
        'Thalassia','Thalassodendron','Zostera'))

str(data)
#check, lat lng needs to be numeric
data$lat <- as.numeric(data$lat)
data$lng <- as.numeric(data$lng)

# Map of points coloured by Genus type (fig3)
map<-leaflet(data=data) %>% addTiles(tilesURL) %>%
  addCircles(lat = ~lat, lng = ~lng, radius=101100,color = ~pal(Genera),stroke = FALSE, fillOpacity = 0.5) %>%
  addLegend (position = 'bottomleft',pal = pal, values = ~Genera);map

cata <- data %>%
  mutate(
    Biomass = Biomass,
    Structure =  Structure,
    Production = Production) %>%
  group_by(lat, lng) %>% ungroup()

cata1 <- cata %>% dplyr::select(Biomass, Production, Structure)
colors <- c("#3838bd", "#099903", "#c810be",)

### Subset
catB = subset(cata, select = -c(Production, Structure, Total)) #getting object that just has biomass
catB2 <- catB[!with(catB,is.na(Biomass)),]  #remove all NA rows (otherwise each map will be identical)
catP = subset(cata, select = -c(Biomass, Structure, Total)) #getting object that just has production
catP2 <- catP[!with(catP,is.na(Production)),]  #remove all NA rows
catS = subset(cata, select = -c(Biomass, Production, Total)) #getting object that just has structure
catS2 <- catS[!with(catS,is.na(Structure)),]  #remove all NA rows
```

```r
# Three maps for Biomass, Production & Structure by number of datapoints (fig4)
mapB<-leaflet(data=catB2) %>% addTiles(tilesURL) %>%
  addCircleMarkers(lat = ~lat, lng = ~lng, radius=~Biomass * 1.5,color = "#3838bd",stroke = FALSE, fillOpacity =
0.5);mapB  #Biomass map, blue

mapP<-leaflet(data=catP2) %>% addTiles(tilesURL) %>%
  addCircleMarkers(lat = ~lat, lng = ~lng, radius=~Production * 1.5,color = "#099903",stroke = FALSE, fillOpacity =
0.5);mapP  #Production map, green

mapS<-leaflet(data=catS2) %>% addTiles(tilesURL) %>%
  addCircleMarkers(lat = ~lat, lng = ~lng, radius=~Production * 1.5,color = "#c810be",stroke = FALSE, fillOpacity =
0.5);mapP  #Structure map, purple

###--- Histogram & Boxplots ---###

library(reshape)
library(ggrepel)

publication <- theme(plot.title = element_text(face = "bold",
                              size = rel(1.2), hjust = 0.5),
        text = element_text(),
        panel.background = element_rect(colour = NA),
        plot.background = element_rect(colour = NA),
        axis.title = element_text(face = "bold",size = rel(1.5)),
        axis.title.y = element_text(angle=90,vjust =2),
        axis.title.x = element_text(vjust = -0.2),
        axis.text.x = element_text(angle = 45, hjust = 1),
        axis.text = element_text(size=15),
        axis.line = element_line(colour="black"),
        axis.ticks = element_line(),
        axis.line.y.right = element_line(colour="black"),
        axis.line.x.top = element_line(colour="black"),
        panel.grid.major = element_line(colour="#FFFFFF"),
        panel.grid = element_line(colour="black"),
        panel.grid.minor = element_blank(),
        panel.border=  element_rect(colour="black", fill = NA, size=1),
        legend.position = "bottom",
        legend.direction = "horizontal",
        legend.key.size= unit(0.7, "cm"),
        legend.spacing = unit(0.7, "cm"),
        legend.title = element_text(NA),
        legend.text= element_text(size=18),
        plot.margin=unit(c(10,5,5,5),"mm"),
        strip.background=element_rect(colour="black",fill="#FFFFFF"),
        strip.placement = "outside",
        strip.text = element_text(face="bold", size=18))

### Histogram with data points by study type
data<- read.csv("CMERpoints_PubYear.csv")
colnames(data)
# subset the data to only include 4 key variables and year
DPoints <- data %>%
  dplyr::select(Reference, PubYear, Biomass, Structure, Production, Total)
colnames(DPoints)

# for each unique reference, sum Biomass, Structure, Production, total values
# so there's ONE row per reference. Make new column showing unique number pubs per year
```

```
tally<-DPoints %>%
  group_by(PubYear) %>%
  summarise(
    unique_sp = n_distinct(Reference))

dP <- melt(DPoints, id.vars=c("PubYear"),
       measure.vars=c( "Biomass" ,"Structure","Production"))
str(dP)
dP$PubYear <- as.factor(dP$PubYear) #remove NA's if needed

# Assign colours to match maps   Note: #blue "#3838bd"  #magenta #c810be     #green #099903
c2 <- c("#099903","#3838bd","#c810be")
names(c2)=c("Production","Biomass","Structure")

# Histogram (fig2)
g2 <- ggplot(dP2,aes(x=PubYear, y=value, group=variable,fill=variable))+
  geom_bar(stat="identity") + scale_fill_manual(values=c2)+
  xlab("Publication Year") + ylab("Number of Data Points")+theme_bw()+publication;g2

### Boxplots by latitude & by genera

# Subset the data to only include 4 key variables by bioregion
df1 <- data %>%
  dplyr::select(Bioregion, Geotype, Genera, Total.biomass..g.DW.m.2., Shoot.density..no..shoots.m.2.,
         Total.production..g.DW.m.2.day..1.,Percent.cover....)

# Remove rows that have all NA in the 4 key variable columns
df2 <- df1[!with(df1,is.na(Total.biomass..g.DW.m.2.)& is.na(Shoot.density..no..shoots.m.2.)&
is.na(Total.production..g.DW.m.2.day..1.)& is.na(Percent.cover....)),]

df <- melt(df2, id.vars=c("Bioregion", "Geotype", "Genera"),
       measure.vars=c("Total.biomass..g.DW.m.2.",  "Shoot.density..no..shoots.m.2.",
              "Total.production..g.DW.m.2.day..1.",
              "Percent.cover...."))
levels(df$variable) #check names & recode to look neat
df$variable <-recode(df$variable, Total.biomass..g.DW.m.2.="Total Biomass", Shoot.density..no..shoots.m.2.="Shoot
Density", Total.production..g.DW.m.2.day..1.= "Total Production", Percent.cover....= "Percent Cover")

df$Bioregion <- as.factor(df$Bioregion)
level_order2 <- c("Temperate North Atlantic","Temperate North Pacific","Mediterranean",
          "Tropical Atlantic" , "Tropical Indo Pacific","Temperate Southern")

# bioregion count
df2 <- merge(aggregate(value ~Bioregion + variable  , df, length),
        aggregate(value~ Bioregion + variable, df, median),
        by=c("Bioregion", "variable"))
names(df2) <- c("Bioregion","variable","count","value")

# Box/Whisker with n values on top for 4 variables by Bioregion (fig5)
Bioregion <- ggplot(df,aes(x=factor(Bioregion,level=level_order2),y=value, fill=variable)) +
  geom_boxplot(aes(fill=variable),position = position_dodge(width=1.0)) +
  geom_text_repel(data=df2, aes(label = count, group=variable),
          vjust=-30, hjust=1.25,segment.color = 'transparent')+
  facet_wrap(~variable, scales="free_y", ncol = sqrt(length(levels(df$variable))))+
  scale_fill_manual(values=k2) +
  labs(fill = "Study Type") +
  xlab("") + ylab(expression(paste("Number of Flowers (m"^"-2",")")))+ theme_bw()+publication+
```

```
   guides(fill='none');Bioregion

df$Genera <- as.factor(df$Genera)
levels(df$Genera) #check names
level_order<- c('Ruppia', 'Halophila','Halodule', 'Zostera', 'Syringodium','Cymodocea',
          'Phyllospadix', 'Amphibolis','Thalassodendron', 'Thalassia', 'Posidonia','Enhalus')

# Box/Whisker with n values on top for 4 variables by Genera (fig6)
df3 <- merge(aggregate(value ~Genera + variable, df, length),
        aggregate(value~ Genera + variable, df, median),
        by=c("Genera", "variable"))
names(df3) <- c("Genera","variable","count","value")

Genera <- ggplot(df,aes(x=factor(Genera, level = level_order), y=value,)) +
 geom_boxplot(aes(fill=variable),position = position_dodge(width=1.0)) +
 geom_text_repel(data=df3, aes(label = count, group=variable),
       vjust=-30, hjust=1,segment.color = 'transparent')+
 facet_wrap(~variable, scales="free_y", ncol = sqrt(length(levels(df$variable))))+
 scale_fill_manual(values=k2) +
 labs(fill = "Study Type") +
 xlab("") + ylab("") + theme_bw()+publication + guides(fill='none');Genera

# save figures using ggsave at 300 DPI
ggsave("BoxplotByGernera.tiff", plot = Genera, device = "tiff", height = 16, width = 21,
     units="cm", path = "C:/Users/simonestrydom/OneDrive - Department of Biodiversity, Conservation and
Attractions/ECU/ECU backup/Documents/My Papers/CMER macrophyte manual/ Updated2022/Update Graphs",
     dpi = 300)

###---###
```