

Extraction des données GDEGeM et mise en forme pour un modèle d'occupancy

Valentin Lauret

13/10/2020

Mise en forme des données SAMM pour l'étude d'occupancy multi-espèces grand dauphins - chalutiers.

Les données

Le 13/10/20 Marine nous envoie les données GDEGeM sur les activités humaines sous la forme de 3 shapefile :

- les transects en effort
- les points avec des activités de plaisance
- les points avec des activités de pêche

```
tr_eoi <- st_read("data_gdegem/Transect en effort GDEGeM RGF93.shp") # données effort EOI
```

```
## Reading layer `Transect en effort GDEGeM RGF93' from data source `/Users/valentinlauret/Google Drive/These/Work/Multispecies Occupancy Data/GDEGeM/Transect en effort GDEGeM RGF93.shp'
## Simple feature collection with 988 features and 17 fields
## geometry type:  MULTILINESTRING
## dimension:      XY
## bbox:           xmin: 721942.7 ymin: 6157794 xmax: 855415.6 ymax: 6270398
## epsg (SRID):    2154
## proj4string:     +proj=lcc +lat_1=49 +lat_2=44 +lat_0=46.5 +lon_0=3 +x_0=700000 +y_0=6600000 +ellps=GRS80 +units=m +no_defs
```

```
peche_eoi <- st_read("data_gdegem/Pts peche GDEGeM RGF93.shp") # données dauphins Matthieu
```

```
## Reading layer `Pts peche GDEGeM RGF93' from data source `/Users/valentinlauret/Google Drive/These/Work/Multispecies Occupancy Data/GDEGeM/Pts peche GDEGeM RGF93.shp'
## Simple feature collection with 680 features and 23 fields
## geometry type:  POINT
## dimension:      XY
## bbox:           xmin: 722971.3 ymin: 6160708 xmax: 843197.4 ymax: 6270307
## epsg (SRID):    2154
## proj4string:     +proj=lcc +lat_1=49 +lat_2=44 +lat_0=46.5 +lon_0=3 +x_0=700000 +y_0=6600000 +ellps=GRS80 +units=m +no_defs
```

La grille d'étude

La grille.

```
load("pays.rdata")
```

```
grid <- st_read("Grid/grid.shp") %>% st_transform(crs = st_crs(pays))
```

```
## Reading layer `grid' from data source `/Users/valentinlauret/Google Drive/These/Work/Multispecies Occupancy Data/Mappe de la zone d'étude/Mappe de la zone d'étude.shp'
## Simple feature collection with 4356 features and 3 fields
## geometry type:  POLYGON
## dimension:      XY
## bbox:           xmin: 701000 ymin: 5886622 xmax: 1467639 ymax: 6390000
## epsg (SRID):    NA
```

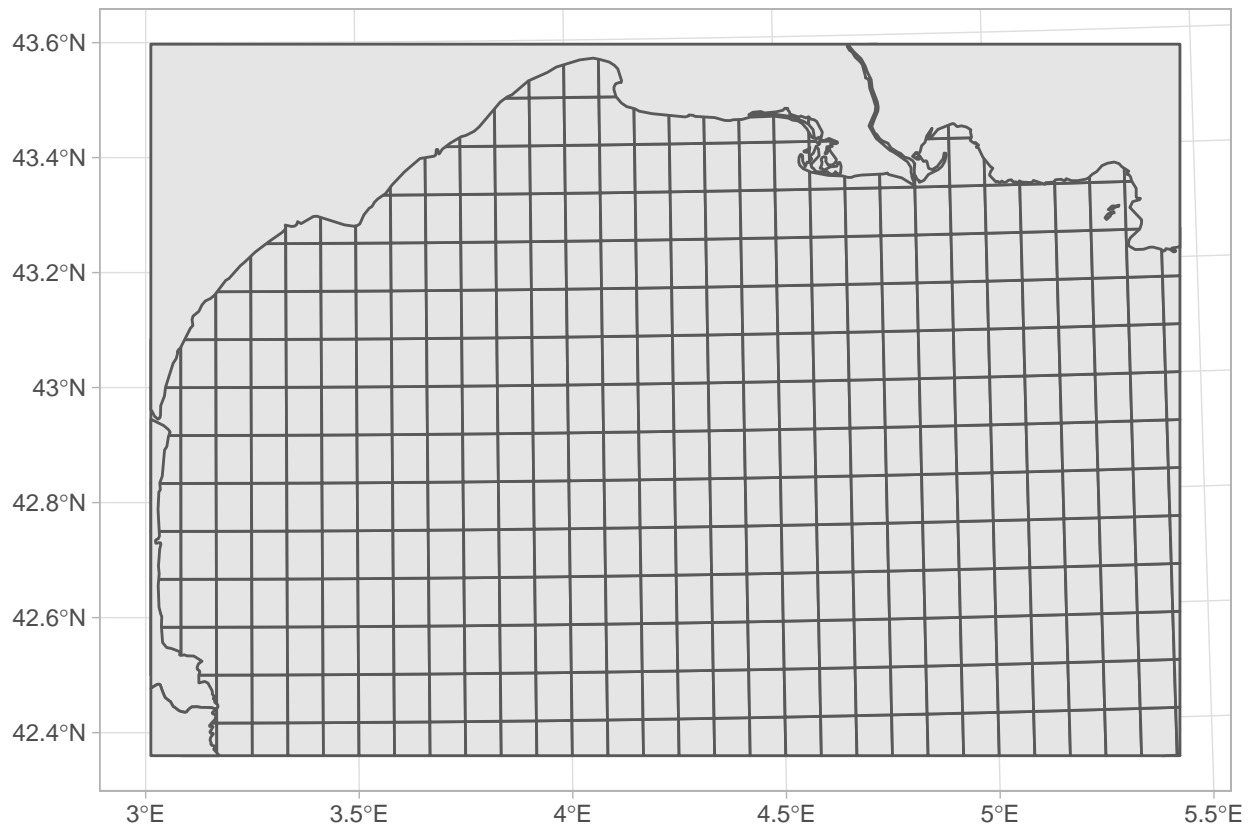
```
## proj4string:      +proj=lcc +lat_1=44 +lat_2=49 +lat_0=46.5 +lon_0=3 +x_0=700000 +y_0=6600000 +ellps=GRS80 +units=m +no_defs
grid %>%
  ggplot() +
  geom_sf() +
  geom_sf(data= pays)
```



On limite l'étude au Golfe du Lion.

```
grid_gdl <- grid %>% st_crop(xmin = 700000, xmax = 900000, ymin = 6140000, ymax = 6300000)
pays_gdl <- pays %>% st_crop(st_bbox(grid_gdl))

grid_gdl %>%
  ggplot() +
  geom_sf() +
  geom_sf(data= pays_gdl)
```

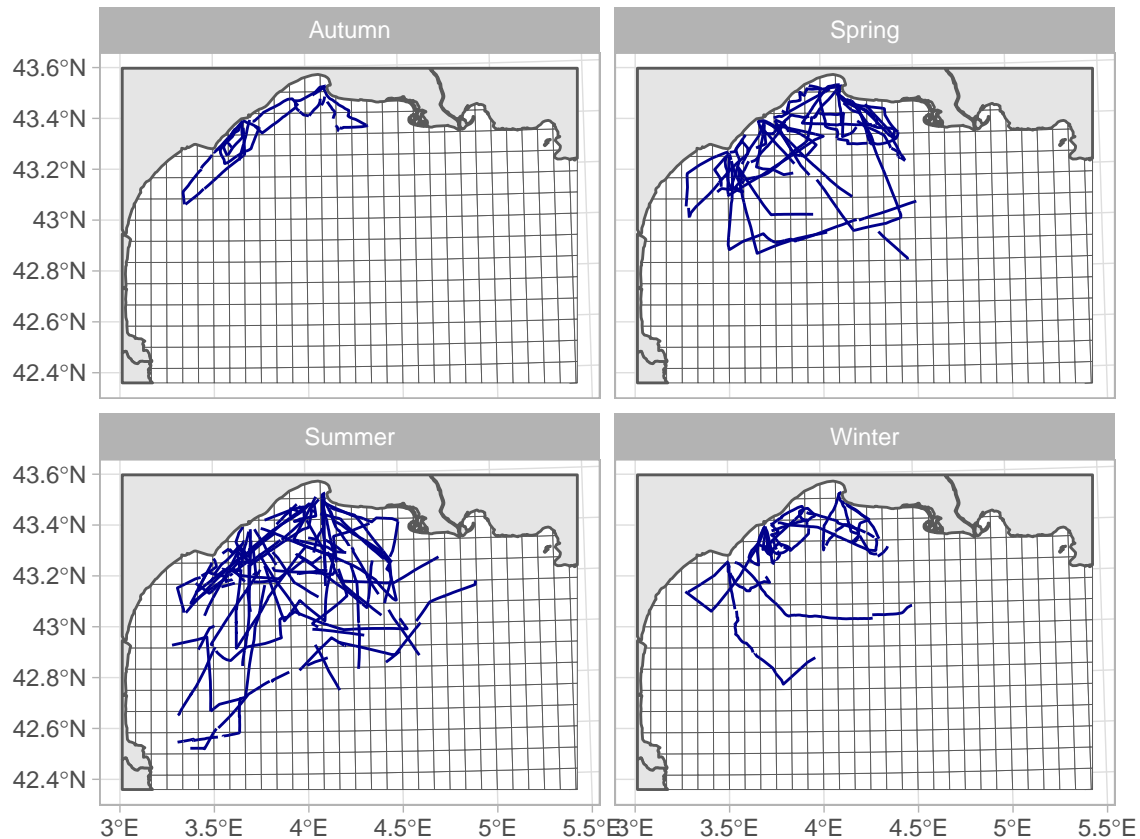


Les transects

Visualisation des transects par saison.

```
tr_eoi2 <- tr_eoi %>% mutate(m = month(tr_eoi$Date)) %>% mutate(season = case_when(
  m %in% c(1,2,3) ~ "Winter",
  m %in% c(4,5,6) ~ "Spring",
  m %in% c(7,8,9) ~ "Summer",
  m %in% c(10,11,12) ~ "Autumn"
))

grid_gdl %>%
  ggplot() +
  geom_sf(fill = "white", lwd = 0.1) +
  geom_sf(data= tr_eoi2, color = "darkblue")+
  geom_sf(data= pays_gdl) +
  facet_wrap(~season, ncol = 2)
```



Pas beaucoup de transect en Automne.

Les observations de grand dauphin

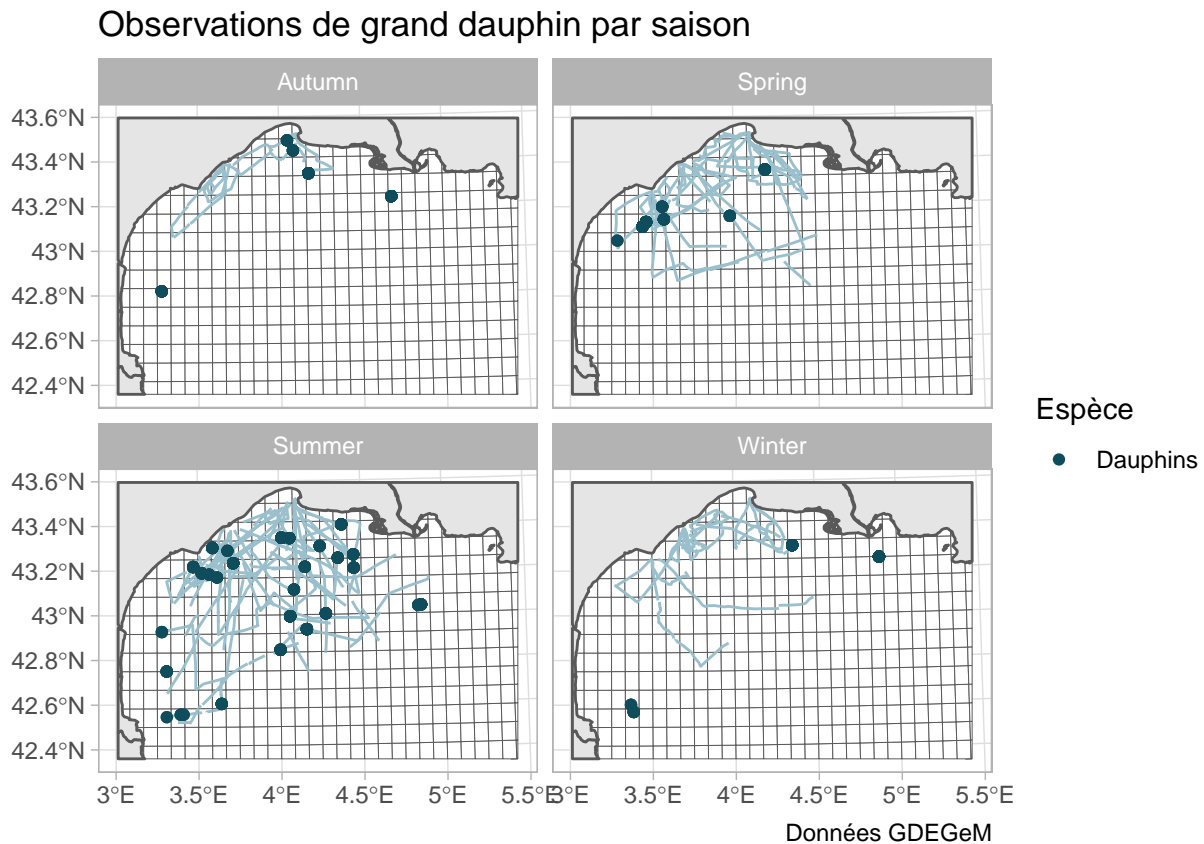
J'utilise la table `tt_obs` en filtrant les detections qui ont été faites par GDEGeM et dans la bonne zone.

```
load("tt_obs.rdata")

obs <- tt_obs %>% filter(suivi == "GDEGeM") %>%
  st_crop(tr_eoi2) %>%
  mutate(m = month(as_date(dates))) %>%
  mutate(season = case_when(
    m %in% c(1,2,3) ~ "Winter",
    m %in% c(4,5,6) ~ "Spring",
    m %in% c(7,8,9) ~ "Summer",
    m %in% c(10,11,12) ~ "Autumn"
  ))

grid_gdl %>%
  ggplot() +
  geom_sf(fill = "white", lwd = 0.1) +
  geom_sf(data= tr_eoi2, color = "lightblue3")+
  geom_sf(data= pays_gdl) +
  geom_sf(data = obs,aes(color = "Dauphins")) +
  facet_wrap(~season, ncol = 2) +
  scale_color_manual(name = "Espèce", values = c("Dauphins" = "#0F4C5C")) +
```

```
labs(title = "Observations de grand dauphin par saison", caption = "Données GDEGeM")
```



Attention il y a des obs qui ont été faites hors transects, par BREACH probablement. On s'en occupera plus tard.

Nombre d'observations de dauphins (sur-estimé)

```
obs %>% nrow() # 673 observations
```

```
## [1] 673
```

Données de pêche

Charge les données sur les activités humaines à partir de la base `peche_eoi`.

```
names(peche_eoi)
```

```
## [1] "campagne" "methode" "ID.TRANSEC" "Plate.form" "Nom"
## [6] "Lat..decim" "Lat..dir." "Long..Deci" "Long..dir" "Saison"
## [11] "Date" "JJ" "MM" "AA" "heureTU"
## [16] "HeureMinlo" "pechart" "fileyeur" "palangrier" "pechprof"
## [21] "chalutier" "senneur" "Engin.dorm" "geometry"
```

Pour l'étude, on ne garde que les chalutiers. On filtre sur `is.na(chalutier) == FALSE`.

```
chalut_eoi <- peche_eoi %>%
  filter(chalutier > 0) %>%
  mutate(m = month(as_date(Date))) %>%
  mutate(season = case_when(
```

```

m %in% c(1,2,3) ~ "Winter",
m %in% c(4,5,6) ~ "Spring",
m %in% c(7,8,9) ~ "Summer",
m %in% c(10,11,12) ~ "Autumn"
))

```

Nombre d'observation de chalutiers

```
chalut_eoi %>% nrow()
```

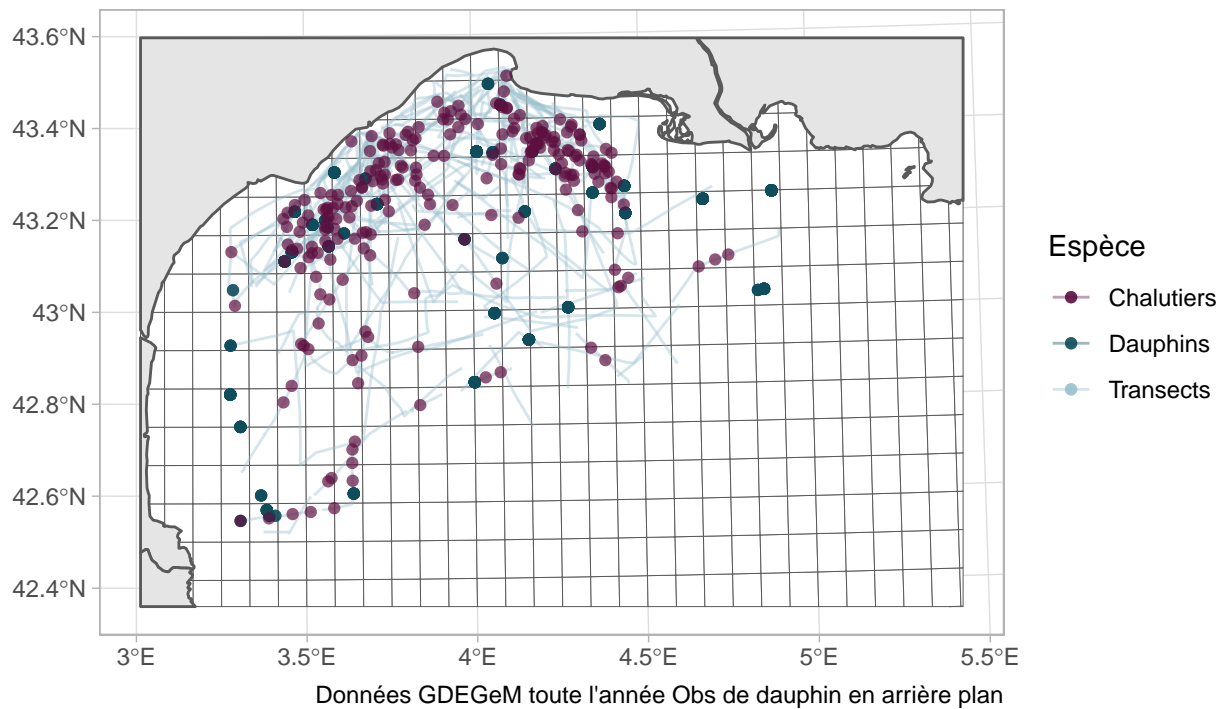
```
## [1] 233
```

```

grid_gdl %>% ggplot() + geom_sf(lwd = 0.1, fill = "white")+
  geom_sf(data = pays_gdl) +
  geom_sf(data= tr_eoi2, aes(color = "Transects"), alpha = 0.4)+
  geom_sf(data = obs, aes(color = "Dauphins"), alpha = 0.7) +
  geom_sf(data = chalut_eoi, aes(color = "Chalutiers"), alpha = 0.7 )+
  scale_color_manual(name = "Espèce", values = c("Dauphins" = "#0F4C5C", "Chalutiers" = "#5f0f40", "Transects" = "#5f0f40"))+
  labs(title = "Observations de chalutiers", caption = "Données GDEGeM toute l'année Obs de dauphin en")

```

Observations de chalutiers



Création des matrices pour l'occupancy

On va créer 3 matrices de dimensions nb_cellules x nb_occasions, i.e. 397 x 4 :

- `effort_gd` pour stocker l'effort comme la longueur de transect par cellule et par occasion.
- `yd_gd` pour stocker les détections de **grand dauphin** par cellule et par occasion.

- `yp_gd` pour stocjer les détection de **chalutier** par cellule et par occasion.

Observation de grand dauphin

Intersection des observations de grand dauhpin avec les cellules de la grille occupancy

```
int <- obs %>%
  st_intersection(grid_gdl) %>%
  select(dates, m, season, objectid, geometry)

int %>% st_intersects(grid_gdl) %>% # nombre de cellule de la grille avec une detection
  unlist() %>%
  unique() %>%
  length()
```

```
## [1] 34
```

Sépare chaque observation en fonction du mois de l'année pour l'assigner à une occasion.

```
# cree la matrice
yd_gd <- grid_gdl %>% mutate(obs = 0, autumn = 0, winter = 0, spring = 0, summer = 0)

# remplit la matrice (il y a peut-être un version tidy pour reste ça rapidos)
for(i in 1:nrow(int)){

  index <- which(yd_gd$objectid == int$objectid[i])
  occ <- int$season[i]

  yd_gd$obs[index] <- 1

  if( occ == "Winter" ) yd_gd$winter[index] <- 1
  if( occ == "Spring" ) yd_gd$spring[index] <- 1
  if( occ == "Summer" ) yd_gd$summer[index] <- 1
  if( occ == "Autumn" ) yd_gd$autumn[index] <- 1

}

# verif
sum(yd_gd$obs)
```

```
## [1] 34
```

La matrice observation de grand dauphin.

```
head(yd_gd)
```

```
## Simple feature collection with 6 features and 8 fields
## geometry type: POLYGON
## dimension: XY
## bbox: xmin: 701000 ymin: 6155576 xmax: 706859.8 ymax: 6211187
## epsg (SRID): NA
## proj4string: +proj=lcc +lat_1=44 +lat_2=49 +lat_0=46.5 +lon_0=3 +x_0=700000 +y_0=6600000 +ellps=G
## objectid ligne colonne geometry obs autumn winter
## 1 22269 DV 133 POLYGON ((701000 6164846, 7... 0 0 0
## 2 22270 DU 133 POLYGON ((701000 6174115, 7... 0 0 0
## 3 22271 DT 133 POLYGON ((701000 6183383, 7... 0 0 0
## 4 22272 DS 133 POLYGON ((701000 6192651, 7... 0 0 0
## 5 22273 DR 133 POLYGON ((701000 6201918, 7... 0 0 0
```

```
## 6      22274      DQ      133 POLYGON ((701000 6211184, 7... 0      0      0
##      spring summer
## 1      0      0
## 2      0      0
## 3      0      0
## 4      0      0
## 5      0      0
## 6      0      0
```

Observation de chalutiers

Intersection des observations de chalutiers avec les cellules de la grille occupancy

```
int <- chalur_eoi %>%
  st_intersection(grid_gdl) %>%
  select(Date, m, chalutier, season, objectid, geometry)

int %>% st_intersects(grid_gdl) %>% # nombre de cellule de la grille avec une detection
  unlist() %>%
  unique() %>%
  length()
```

```
## [1] 71
```

Sépare chaque observation en fonction du mois de l'année pour l'assigner à une occasion.

```
# cree la matrice
yp_gd <- grid %>% mutate(obs = 0, autumn = 0, winter = 0, spring = 0, summer = 0)

# remplit la matrice (il y a peut-être un version tidy pour reste ça rapides)
for(i in 1:nrow(int)){

  index <- which(yp_gd$objectid == int$objectid[i])
  occ <- int$season[i]

  yp_gd$obs[index] <- 1

  if( occ == "Winter" ) yp_gd$winter[index] <- 1
  if( occ == "Spring" ) yp_gd$spring[index] <- 1
  if( occ == "Summer" ) yp_gd$summer[index] <- 1
  if( occ == "Autumn" ) yp_gd$autumn[index] <- 1
}

# verif
sum(yp_gd$obs)
```

```
## [1] 71
```

La matrice observation de chalutier.

```
head(yp_gd)
```

```
## Simple feature collection with 6 features and 8 fields
## geometry type: POLYGON
## dimension: XY
## bbox: xmin: 701000 ymin: 5886622 xmax: 707143.7 ymax: 5942119
## epsg (SRID): NA
## proj4string: +proj=lcc +lat_1=44 +lat_2=49 +lat_0=46.5 +lon_0=3 +x_0=700000 +y_0=6600000 +ellps=G
```



```
##      objectid ligne colonne          geometry obs autumn winter
## 1      22240    EY      133 POLYGON ((701000 5895635, 7...    0      0      0
## 2      22241    EX      133 POLYGON ((701000 5904934, 7...    0      0      0
## 3      22242    EW      133 POLYGON ((701000 5914231, 7...    0      0      0
## 4      22243    EV      133 POLYGON ((701000 5923527, 7...    0      0      0
## 5      22244    EU      133 POLYGON ((701000 5932822, 7...    0      0      0
## 6      22245    ET      133 POLYGON ((701000 5942116, 7...    0      0      0
##      spring summer
## 1          0      0
## 2          0      0
## 3          0      0
## 4          0      0
## 5          0      0
## 6          0      0
```

Effort d'échantillonnage

Longueur totale des transect GDEGeM toutes saisons confondues

```
st_length(tr_eoi2) %>% sum()
```

```
## 4977300 [m]
```

Intersection entre tous les transects et les cellules de la grille + calcul de la longueur de chaque bout de transect

```
int <- tr_eoi2 %>%
  st_intersection(grid_gdl)# %>%
```

```
int <- int %>% select(Date, m, season, objectid, geometry) %>% mutate(eff = st_length(int))
```

Dans plusieurs cas, il y a différents transects qui sont passés dans une même cellule au cours de la même saison. Il faut regrouper.

```
int2 <- int %>%
  group_by(objectid, season) %>%
  summarise(sum(eff))
```

```
# nb de sites différents échantillonnés
length(unique(int2$objectid))
```

```
## [1] 148
```

```
# effort total
sum(int2$`sum(eff)`)
```

```
## 4977300 [m]
```

Maintenant on peut séparer chaque transect en fonction du mois de l'année pour l'assigner à une occasion.

```
# cree la matrice
effort_gd <- grid_gdl %>% mutate(eff.tot = 0, autumn = 0, winter = 0, spring = 0, summer = 0)

# remplit la matrice (il y a peut-être un version tidy pour reste ça rapides)
for(i in 1:nrow(int2)){

  index <- which(effort_gd$objectid == int2$objectid[i])
  occ <- int2$season[i]
```

```

if( occ == "Winter" )   effort_gd$winter[index] <- int2$`sum(eff)`[i]
if( occ == "Spring" )   effort_gd$spring[index] <- int2$`sum(eff)`[i]
if( occ == "Summer" )   effort_gd$summer[index] <- int2$`sum(eff)`[i]
if( occ == "Autumn" )   effort_gd$autumn[index] <- int2$`sum(eff)`[i]
}

effort_gd$eff.tot <- effort_gd$autumn + effort_gd$winter + effort_gd$summer + effort_gd$spring

# verif
sum(effort_gd$eff.tot)

```

```
## [1] 4977300
```

Attention, maintenant on va dégager les observations faites hors effort.

```

# autumn
notsampled_a <- which(effort_gd$autumn == 0)
yd_gd$autumn[notsampled_a] <- 0

# winter
notsampled_w <- which(effort_gd$winter == 0)
yd_gd$winter[notsampled_w] <- 0

# spring
notsampled_sp <- which(effort_gd$spring == 0)
yd_gd$spring[notsampled_sp] <- 0

# summer
notsampled_s <- which(effort_gd$summer == 0)
yd_gd$summer[notsampled_s] <- 0

# all
notsampled <- which(effort_gd$eff.tot == 0)
yd_gd$obs[notsampled] <- 0

# verif
sum(yd_gd$obs)

```

```
## [1] 30
```

La matrice de l'effort d'échantillonnage.

```
head(effort_gd)
```

```

## Simple feature collection with 6 features and 8 fields
## geometry type: POLYGON
## dimension: XY
## bbox: xmin: 701000 ymin: 6155576 xmax: 706859.8 ymax: 6211187
## epsg (SRID): NA
## proj4string: +proj=lcc +lat_1=44 +lat_2=49 +lat_0=46.5 +lon_0=3 +x_0=700000 +y_0=6600000 +ellps=GRS80
## objectid ligne colonne geometry eff.tot autumn winter
## 1 22269 DV 133 POLYGON ((701000 6164846, 7... 0 0 0
## 2 22270 DU 133 POLYGON ((701000 6174115, 7... 0 0 0
## 3 22271 DT 133 POLYGON ((701000 6183383, 7... 0 0 0
## 4 22272 DS 133 POLYGON ((701000 6192651, 7... 0 0 0
## 5 22273 DR 133 POLYGON ((701000 6201918, 7... 0 0 0

```

```
## 6      22274      DQ      133 POLYGON ((701000 6211184, 7...      0      0      0
##      spring summer
## 1        0        0
## 2        0        0
## 3        0        0
## 4        0        0
## 5        0        0
## 6        0        0
```

Résumé

Regroupe les trois matrices dans une liste et sauve les résultats

```
multioccu_gd <- list(dauphins = yd_gd, chalut = yp_gd, effort = effort_gd)

save(multioccu_gd, file = "msoccu_gd.rdata")
```

Visualisation des occurrences des deux espèces

```
# co-occurrence
cooc <- which(multioccu_gd$chalut$obs == 1 & multioccu_gd$dauphins$obs == 1)

# plot
multioccu_gd$dauphins[] %>% filter(obs == 1) %>%
  ggplot() + geom_sf(aes(fill = "Dauphins"), lwd = 0.1) +
  geom_sf(data = multioccu_gd$chalut[multioccu_gd$chalut$obs == 1,], aes(fill = "Chalutiers"), lwd = 0.1) +
  geom_sf(data = grid_gdl[cooc,], aes(fill = "Co-occurrence"), lwd = 0.1) +
  scale_fill_manual(name = "Espèce", values = c("Dauphins" = "#0F4C5C", "Chalutiers" = "#5f0f40", "Co-occurrence" = "#f0f0f0")) +
  geom_sf(data = pays_gdl) +
  labs(title = "Occurrences des dauphins et chalutiers à partir des données GDEGeM",
        subtitle = "sur la grille utilisée pour l'occupancy")
```

Occurrences des dauphins et chalutiers à partir des données GDEGeM
sur la grille utilisée pour l'occupancy

