Annex 3 - R script to construct Table 1A by country

Joel Vigneau, with the collaboration of Matt Elliott, Jon Elson, Kirsten Birch Hakansson, Marie Storr-Paulsen, Nuno Prista, Katja Ringhdal, Lies Vansteenbrugge, Sieto Verver

October 2016

Filling of the EU-MAP table 1A requires to report on country shares of landings and shares of EU TAC when relevant, for all the stocks listed in table 1A of the EU-MAP Regulation (EU Decision 1254/2016). This process necessitates to gather information on landings and EU TAC from an official database, namely EUROSTAT for EU landings and MARE/FIDES for EU TAC.

Two datasets were added to complete the references, (1) the Nephrops FU landings provided by ICES and (2) the Mediterranean and Black Sea landings fisgures put together by 2016 RCM Mediterranean and Black Sea.

First of all, the datasets listed above contain information from all EU Member States, which means that the script has the potential to be used by all Member States, and by STECF for control of the NWP submitted for 2017.

Setting the parameters for the analysis

The variables needed for the work are the working directory, the country code (2-letter code) and the reference years

```
library(tidyr)
library(stringr)
library(reshape)
```

```
##
## Attaching package: 'reshape'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, smiths
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:reshape':
##
##     rename
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
CTRY <- 'DK'
refYears <- 2016:2018
```

Importing the datasets

The list of datasets are the following :

1. Landings and TAC shares files:

- EUROSTAT landings files: http://ec.europa.eu/eurostat/web/fisheries/data/database
- MARE/FIDES TAC file: https://webgate.ec.europa.eu/fides/index.cfm
- ICES Nephrops fishery units landings per country for 2015
- RCM Mediterranean and Black Sea 2016 landings compilation

```
DF27  <- read.table(file.path(input_path_common, "fish_ca_atl27.tsv"),header=TRUE, sep='\t', as.is=TRUE)
DF21  <- read.table(file.path(input_path_common, "fish_ca_atl21.tsv"),header=TRUE, sep='\t', as.is=TRUE)
DF34  <- read.table(file.path(input_path_common, "fish_ca_atl34.tsv"),header=TRUE, sep='\t', as.is=TRUE)
DF41  <- read.table(file.path(input_path_common, "fish_ca_atl41.tsv"),header=TRUE, sep='\t', as.is=TRUE)
DF47  <- read.table(file.path(input_path_common, "fish_ca_atl47.tsv"),header=TRUE, sep='\t', as.is=TRUE)
DF51  <- read.table(file.path(input_path_common, "fish_ca_ind51.tsv"),header=TRUE, sep='\t', as.is=TRUE)
DF37  <- read.table(file.path(input_path_common, "fish_ca_atl37.tsv"),header=TRUE, sep='\t', as.is=TRUE)

DF <- rbind.data.frame(DF27, DF21)
DF <- rbind.data.frame(DF, DF34)
DF <- rbind.data.frame(DF, DF37)
DF <- rbind.data.frame(DF, DF41)
DF <- rbind.data.frame(DF, DF47)
DF <- rbind.data.frame(DF, DF51)

TAC <- read.csv(file.path(input_path_common,'EU opening quota 2018.csv'), header=TRUE, sep=';',as.is=TRU
NEP <- read.csv(file.path(input_path_common,'Nephrops landings 2015.csv'), header=TRUE, sep=';', as.is=
patch_land_codIIIa <- read.table(file.path(input_path_common, 'patch_codIIIa_20191009.txt'), sep = '\t'
MED <- read.csv(file.path(input_path_common,'RCM MED landings.csv'), header=TRUE, sep=';', as.is=TRUE)
```

2. Reference tables:

- EuroStat Geo.def: full names of countries
- ASFIS file : FAO species naming and coding
- Linkage table mirroring EU-MAP Table 1A naming of species and stock area, and lining to EUROSTAT and MARE/FIDES species and area naming

```
GEO     <- read.table(file.path(input_path_common,'geo.def'),header=TRUE,sep=";", as.is=TRUE)
ASFIS   <- read.table(file.path(input_path_common,'ASFIS_sp_Feb_2016.txt'), header=TRUE, sep="\t", as.is
table1A <- read.csv(file.path(input_path_dnk,'EUMAP_Table1A_Linkage_EUROSTAT and EC_TAC_dnk.csv'), sep=
```

data.frame preparation

The country names are matching between GEO and TAC data.frame, except for UK, so the following lines enables the full match.

```
TAC$StockID <- paste(TAC$Species.Code,TAC$Area.Code,sep="")
TAC$Level.Description[substring(TAC$Level.Description,1,3) %in% 'U.K'] <- 'United Kingdom'
```

The TAC dataset is well structured and thus ready for the analysis

```
head(TAC,3)
```

```
##   Load_ind Definition.Year Species.Code
## 1     INIV            2018      AGO_CSQ
## 2     INIV            2018      AGO_CSQ
## 3     INIV            2018      AGO_CSQ
##                                                 Species.Name Area.Code
## 1 Angola direct agreement fishing category (Coastal State quota)    DIR_AG
## 2 Angola direct agreement fishing category (Coastal State quota)    DIR_AG
## 3 Angola direct agreement fishing category (Coastal State quota)    DIR_AG
```

```
##    Area.Description Level.Code Level.Description Initial.Quantity
## 1 Direct Agreements        BEL           Belgium               NA
## 2 Direct Agreements        BGR          Bulgaria               NA
## 3 Direct Agreements        CYP            Cyprus               NA
##   Adapted.Quota Eurlex.Ref OJ.Ref Publication.Date Page.Number
## 1            NA 32017R2403   L347       2017-12-28          81
## 2            NA 32017R2403   L347       2017-12-28          81
## 3            NA 32017R2403   L347       2017-12-28          81
##   In.regulation Compute.uptake       StockID
## 1             Y              Y AGO_CSQDIR_AG
## 2             Y              Y AGO_CSQDIR_AG
## 3             Y              Y AGO_CSQDIR_AG
```

```
names(GEO)[2] <- "Country"
GEO$geo <- toupper(GEO$geo) #2-letter code should be in capitals
SRG <- strsplit(as.character(DF$species.fishreg.unit.geo.time),split=",")
SRG.m <- matrix(unlist(SRG), ncol=4, byrow=TRUE)
coln <- sapply(refYears, function(x) which(grepl(x,names(DF))))
DFT <- data.frame(X3A_CODE = toupper(SRG.m[,1]), area = toupper(SRG.m[,2]), geo = SRG.m[,4],
Y1 = DF[,coln[1]], Y2 = DF[,coln[2]], Y3 = DF[,coln[3]])
DFM <- merge(DFT, GEO, all.x=TRUE)
DFM$Y1 <- as.numeric(str_remove_all(as.character(DFM$Y1), "[bcdefinpzsu]")) #Removing characthers in am
DFM$Y2 <- as.numeric(str_remove_all(as.character(DFM$Y2), "[bcdefinpzsu]")) #Removing characthers in am
DFM$Y3 <- as.numeric(str_remove_all(as.character(DFM$Y3), "[bcdefinpzsu]")) #Removing characthers in am
DFM <- DFM[!is.na(DFM$Country),]
DFM <- merge(DFM, ASFIS[,c(3:6)], all.x=TRUE)
```

Let's have a look at the workable structure of EuroStat dataset. Note that Y1, Y2 and Y3 are the 3-year period demanded, and the presence of NA's. The assumption made here (further in the Construction of the table section) is to exclude NA from the average, i.e. like if MS had omitted to report, instead of a NA which would mean 0. The confusion comes because lots of 0 are reported in EuroStat (implicitely meaning that NA is not a 0). This point may be subject of a STECF agreement or suggestion for modification.

```
head(DFM,3)
```

```
##   X3A_CODE  geo      area   Y1 Y2 Y3                Country
## 1      AAS EU28 27_3_C_22   NA NA NA European union (28 MS)
## 2      AAS EU28    27_4_B 0.03 NA NA European union (28 MS)
## 3      AAS   DK      27_4 0.03 NA NA                Denmark
##   Scientific_name   English_name              French_name
## 1 Astacus astacus Noble crayfish Écrevisse à pieds rouges
## 2 Astacus astacus Noble crayfish Écrevisse à pieds rouges
## 3 Astacus astacus Noble crayfish Écrevisse à pieds rouges
```

```
NEP <- merge(NEP, GEO, all.x=TRUE)
NEP$geo[is.na(NEP$geo)] <- 'UK'
NEP2 <- data.frame(X3A_CODE='NEP', geo=NEP$geo, area=NEP$Stock, Y1=round(NEP$TotalLanding.in.kg/1000,0)
```

A look at the Nephrops dataset on the same format as EuroStat dataset, so they can be merged

```
head(NEP2)
```

```
##   X3A_CODE geo   area  Y1 Y2 Y3 Country     Scientific_name   English_name
## 1      NEP  BE nep-22   5 NA NA Belgium Nephrops norvegicus Norway lobster
## 2      NEP  BE nep-15   0 NA NA Belgium Nephrops norvegicus Norway lobster
## 3      NEP  BE nep-33 299 NA NA Belgium Nephrops norvegicus Norway lobster
## 4      NEP  BE  nep-5 146 NA NA Belgium Nephrops norvegicus Norway lobster
```

```
## 5      NEP  BE nep-14   0 NA NA Belgium Nephrops norvegicus Norway lobster
## 6      NEP  BE  nep-6   0 NA NA Belgium Nephrops norvegicus Norway lobster
##   French_name
## 1 Langoustine
## 2 Langoustine
## 3 Langoustine
## 4 Langoustine
## 5 Langoustine
## 6 Langoustine
```

```
DFM <- rbind.data.frame(DFM, NEP2)
MEDA <- merge(MED, ASFIS[,c(3,4,5,6)], by.x='Species', by.y='Scientific_name', all.x=TRUE)
MEDA <- tidyr::gather(MEDA,"Country","n",4:13)
MEDAG <- merge(MEDA, GEO, all.x=TRUE)
```

and a look at the Mediterranean dataset

```
head(MEDAG,3)
```

```
##    Country           Species                 Area  RefYears
## 1 Bulgaria  Alopias vulpinus All areas in the Med 2013-2015
## 2 Bulgaria Anguilla anguilla all areas in the Med 2013-2015
## 3 Bulgaria      Aphia minuta   GSA 9,10,16 and 19 2013-2015
##   Total.average.landings..t. X3A_CODE     English_name       French_name n
## 1                        9.0      ALV         Thresher            Renard 0
## 2                      308.0      ELE     European eel Anguille d'Europe 0
## 3                       50.7      FIM Transparent goby            Nonnat 0
##   geo
## 1  BG
## 2  BG
## 3  BG
```

```
MED <- data.frame(X3A_CODE=MEDAG$X3A_CODE, geo=MEDAG$geo, area=MEDAG$Area, Y1=round(MEDAG$n,0),
                  Y2=NA, Y3=NA, Country=MEDAG$Country, Scientific_name=MEDAG$Species, English_name=MEDA
                  French_name=NA)
DFM <- rbind.data.frame(DFM, MED)
```

Construction of the table

```
T1A <- data.frame()
for (i in 1:nrow(table1A)) {

    ctry2 <- GEO$Country[GEO$geo %in% CTRY]
    reg <- strsplit(as.character(table1A$areaBis[i]), split=',')

    if (table1A$region[i] %in% 'Mediterranean and Black Sea') {
        DT <- DFM[DFM$Scientific_name %in% table1A$latinName[i] & tolower(DFM$area) %in% tolower(paste(
      } else {

        ting<-unlist(strsplit(table1A$latinName[i], split=","))
        DT <- DFM[DFM$Scientific_name %in% ting & tolower(DFM$area) %in% tolower(reg[[1]]),]

              }

    if (table1A$latinName[i]=="Gadus morhua" & (table1A$area[i] %in% c("IIIaN","IIIaS")) & sum(refYears
```

```
{
 a<-melt(DT[c("geo","Y1","Y2","Y3")])
 a$id<-paste(a$geo,a$variable)
 patch_land_codIIIa$id <- paste0(patch_land_codIIIa$geo, " Y", patch_land_codIIIa$year_seq)
 a<-data.frame(a, b=patch_land_codIIIa[,table1A$area[i]][match(a$id, patch_land_codIIIa$id)]*a$valu
 DT[,4:6]<-matrix(data=a$b, ncol=3)
}

DT$MOY <- apply(DT[,4:6],1,mean,na.rm=TRUE)
#RFMO <- 'ICES'
if (substring(table1A$region[i],1,3) %in% 'Med') RFMO <- 'GFCM'

T1 <- data.frame(MS=CTRY, refYears=paste(min(refYears),'-',max(refYears),sep=""),spp=table1A$latinN
    RFMO=table1A$RFMO[i], area = table1A$area[i],select=NA, landings=NA, TAC=NA,shareLanding=NA,Thr
ind <- which(DT$geo %in% CTRY)
if (length(ind)>0) {
    T1$landings <- sum(DT$MOY[DT$geo %in% CTRY],na.rm=TRUE)
    T1$shareLanding <- T1$landings/sum(DT$MOY[DT$geo != 'EU28'], na.rm=TRUE)
    #T1$shareLanding <- T1$landings/sum(DT$MOY, na.rm=TRUE)
    } else {
    T1$landings <- 0
    T1$shareLanding <- 0
    }

## TAC
ind.ct<-NULL
if (!(table1A$FIDES_stockID[i] %in% 'No TAC')) {
    aa<-strsplit(as.character(table1A$FIDES_stockID[i]),split=',')[[1]]
    TACi <- TAC[TAC$StockID %in% aa,]
    if (length(aa)>1)
        TACi <- aggregate(list(Initial.Quantity = TACi$Initial.Quantity),
            by=list(Level.Code=TACi$Level.Code, Level.Description=TACi$Level.Description), sum)
    ind.ct <- TACi$Initial.Quantity[which(TACi$Level.Description %in% ctry2)]
    ind.eu <- TACi$Initial.Quantity[which(TACi$Level.Code %in% 'EEC')]
    if (length(ind.ct) == 1) T1$TAC <- ind.ct/ind.eu
    T1$Comments<-NA
    TT <- tapply(TACi$Initial.Quantity, TACi$Level.Description,sum,na.rm=TRUE)/TACi$Initial.Quantit
    TT <-  TT[names(TT) %in% GEO$Country]  #Keep only the EU countries to calculate the 25% rule
    if (!(is.na(T1$TAC)) & T1$TAC <0.1 & T1$TAC>0)  T1$Comments <- sum(TT[which(TT<0.1)])
    if (!(is.na(T1$Comments)) & T1$Comments >=.25)  {
        print(T1)
        print(TT[TT<.1])
        cat('\n')
    }
}
#Add-on Sept 2019 Joel
if (length(ind.ct)>0) {
  if (!is.na(ind.ct))  {
  T1$Comments2 <- paste('FIDES Initial.Quantity =',ind.ct)
}}
##
T1A <- rbind.data.frame(T1A, T1)
T1A$Thresh <- as.character(T1A$Thresh)
```

```
    #Threshold ruling
    # T1A$Thresh[T1A$TAC >=.1 & T1A$landings >=200] <- 'M'  #rule (a) & (c)
    # T1A$Thresh[is.na(T1A$TAC) & T1A$shareLanding >=.1 & T1A$landings >=200] <- 'M' #rule (b) & (c)
    # T1A$Thresh[T1A$TAC <.1 & T1A$Comments >=.25] <- 'C'   # 25% rule, sampling to be coordinated betw
    #Threshold ruling specified like the EU Reg
    T1A$Thresh[T1A$TAC <.1] <- 'Y'  #rule (a)
    T1A$Thresh[is.na(T1A$TAC) & T1A$shareLanding <.1] <- 'Y' #rule (b)
    T1A$Thresh[T1A$landings < 200] <- 'Y'  #rule (c)
    T1A$Thresh[T1A$TAC <.1 & T1A$Comments >=.25] <- 'N'   # 25% rule, sampling to be coordinated betwee
}
```

```
## Using geo as id variables
## Using geo as id variables
```

```
##   MS  refYears               spp                        region RFMO area
## 1 DK 2016-2018 Macrourus berglax North Sea and Eastern Arctic ICES   IV
##   select landings TAC shareLanding Thresh Comments Comments2
## 1     NA        0 0.1            0      N      0.3        NA
##        Denmark        Germany United Kingdom
##            0.1            0.1            0.1
##
##   MS  refYears              spp         region RFMO
## 1 DK 2016-2018 Scomber scombrus North Atlantic ICES
##                          area select landings        TAC shareLanding
## 1 II,IIIa,IV,V,VI,VII,VIII,IX     NA 37314.99 0.08336916   0.08313482
##   Thresh Comments Comments2
## 1      N 0.290335        NA
##      Belgium      Denmark      Estonia       France      Germany
## 1.412206e-03 8.336916e-02 4.680145e-04 4.208049e-02 5.791407e-02
##       Latvia    Lithuania  Netherlands       Poland        Spain
## 3.455688e-04 3.455688e-04 8.678675e-02 3.972681e-03 5.986232e-05
##       Sweden
## 1.358058e-02
##
##   MS  refYears               spp     region RFMO  area select landings
## 1 DK 2016-2018 Sprattus sprattus Baltic Sea ICES 22-32     NA 23341.25
##          TAC shareLanding Thresh  Comments Comments2
## 1 0.09864283   0.09847156      N 0.2628188        NA
##    Denmark    Finland    Germany  Lithuania
## 0.09864283 0.05163738 0.06249476 0.05004384
##
##   MS  refYears                 spp         region RFMO
## 1 DK 2016-2018 Trachurus trachurus North Atlantic ICES
##                               area select landings        TAC
## 1 IIa,IVa,Vb,VIa,VIIa-c,e-k,VIIIabde     NA 6839.608 0.09913542
##   shareLanding Thresh  Comments Comments2
## 1    0.1060762      N 0.3265306        NA
##        Belgium        Denmark         France        Germany         Latvia
##    0.000000000    0.099135418    0.039810998    0.077349955    0.000000000
##      Lithuania       Portugal         Sweden United Kingdom
##    0.000000000    0.010163869    0.006785966    0.093284407
##
##   MS  refYears                 spp                        region RFMO area
## 1 DK 2016-2018 Trachurus trachurus North Sea and Eastern Arctic ICES  IIa
```

```
##   select landings        TAC shareLanding Thresh  Comments Comments2
## 1     NA        0 0.09913542            0      N 0.3265306        NA
##        Belgium        Denmark         France        Germany         Latvia
##    0.000000000    0.099135418    0.039810998    0.077349955    0.000000000
##      Lithuania       Portugal         Sweden United Kingdom
##    0.000000000    0.010163869    0.006785966    0.093284407
```

##Formatting

```
T1B <- T1A
T1B$landings <- round(T1B$landings,0)
T1B$landings[T1B$landings == 0] <- '-'
T1B$TAC <- paste(round(100*T1B$TAC,0),'%',sep='')
T1B$TAC[T1B$TAC %in% c('NA%','NaN%','Inf%')] <- '-'
T1B$shareLanding <- paste(round(100*T1B$shareLanding,0),'%',sep='')
T1B$shareLanding[T1B$shareLanding %in% c('NA%','NaN%','Inf%')] <- '-'
T1B$Thresh[T1B$landings %in% '-' & T1B$TAC %in% '-'] <- T1B$shareLanding[T1B$landings %in% '-' & T1B$TA
ind <- which(T1B$Comments>.0) #Changed, so I know what the figures means
T1B$Comments <- paste(round(100*T1B$Comments,0),'%',sep='')
T1B$Comments[T1B$Comments<.25] <- '-'
T1B$Comments[T1B$Comments %in% c('NA%','NaN%','Inf%')] <- '-'
T1B$Comments[ind] <- paste('Sum of MS shares <10% = ',T1B$Comments[ind],sep='')
T1B$Comments[!ind]<-'-'
T1B$select <- '-'
T1B$select[T1B$Thresh %in% c('N')] <- 'Y'
T1B$select[T1B$Thresh %in% 'Y'] <- 'N'
T1B[T1B$RFMO %in% c('ICCAT','IOTC','WCPFC') & T1B$landings>0,c('select','Thresh')] <- c('Y','N')
T1B[T1B$spp %in% 'Anguilla anguilla' & T1B$landings>0,c('select','Thresh')] <- c('Yes','No')
T1B[T1B$spp %in% 'Nephrops norvegicus' & !(grepl('TAC', T1B$area)) & T1B$landings < 0,'TAC'] <- '-' #Ch
levels(T1B$refYears) <- c(levels(T1B$refYears), '2015')
T1B[T1B$spp %in% 'Nephrops norvegicus' & !(grepl('TAC', T1B$area)) & T1B$landings < 0,'refYears'] <- '2
T1B[T1B$RFMO %in% 'GFCM', 'refYears'] <- '2015'
```

```
#Added so we can use the output to create table 1B & C
names(T1B)
```

```
##  [1] "MS"           "refYears"     "spp"          "region"
##  [5] "RFMO"         "area"         "select"       "landings"
##  [9] "TAC"          "shareLanding" "Thresh"       "Comments"
## [13] "Comments2"
```

```
names(table1A)
```

```
##  [1] "region"               "sppName"              "latinName"
##  [4] "RFMO"                 "RFMO_Stock_ID"        "FIDES_stockID"
##  [7] "area"                 "areaBis"              "TAC.area.description"
## [10] "Comments"             "reportingName"        "latinName_old"
## [13] "FIDES_stockID_old"
```

```
T1B$Comments_comb <- ifelse(T1B$Comments != "-" & !(is.na(T1B$Comments2)), paste(T1B$Comments, T1B$Comm
                            ifelse(T1B$Comments == "-" & !(is.na(T1B$Comments2)), T1B$Comments2,
                                   ifelse(T1B$Comments != "-" & is.na(T1B$Comments2), T1B$Comments, "")


T1C <- merge(select(distinct(T1B), -Comments, -Comments2), select(distinct(table1A), -Comments), by.x =
names(T1C)
```

```
##  [1] "region"               "spp"                  "area"
##  [4] "RFMO"                 "MS"                   "refYears"
##  [7] "select"               "landings"             "TAC"
## [10] "shareLanding"         "Thresh"               "Comments_comb"
## [13] "sppName"              "RFMO_Stock_ID"        "FIDES_stockID"
## [16] "areaBis"              "TAC.area.description" "reportingName"
## [19] "latinName_old"        "FIDES_stockID_old"
```

```
T1C <- mutate(T1C, speciesIncluded = spp, spp = reportingName, Comments = Comments_comb)
names(T1C)
```

```
##  [1] "region"               "spp"                  "area"
##  [4] "RFMO"                 "MS"                   "refYears"
##  [7] "select"               "landings"             "TAC"
## [10] "shareLanding"         "Thresh"               "Comments_comb"
## [13] "sppName"              "RFMO_Stock_ID"        "FIDES_stockID"
## [16] "areaBis"              "TAC.area.description" "reportingName"
## [19] "latinName_old"        "FIDES_stockID_old"    "speciesIncluded"
## [22] "Comments"
```

```
T1C<-T1C[,c("MS","refYears","spp","region","RFMO","area","select","landings","TAC","shareLanding","Thres
```

Export of Table 1A

the rule sum of quotas for coutries <10% (less or more than 25%) is noted in the comments column

```
ind <- order(as.character(T1B$region), as.character(T1B$RFMO), as.character(T1B$spp), as.character(T1B$a
write.table(T1C[ind,], file=paste(output_path, CTRY,'_table1A_filled_dnk.csv',sep=''), sep=';',row.names
```