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)
setwd('.')
CTRY <- 'ES'
refYears <- 2013:2015
```

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

```
DF  <- read.table("fish_ca_atl27.tsv",header=TRUE, sep='\t', as.is=TRUE)  # Atlantic NE
TAC <- read.csv('EU opening quota.csv', header=TRUE, sep=';',as.is=TRUE)
NEP <- read.csv('Nephrops landings 2015.csv', header=TRUE, sep=';', as.is=TRUE)
MED <- read.csv('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('geo.def',header=TRUE,sep=";", as.is=TRUE)
ASFIS   <- read.table('ASFIS_sp_Feb_2012.txt', header=TRUE, sep="\t", as.is=TRUE)
table1A <- read.table('EUMAP_Table1A_Linkage_EUROSTAT and EC_TAC.csv', sep=';', header=TRUE, as.is=TRUE)
```

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$Level.Description[substring(TAC$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 Species.Name Area.Code
## 1     INIV            2015          ALB     Albacore     AN05N
## 2     INIV            2015          ALB     Albacore     AN05N
## 3     INIV            2015          ALB     Albacore     AN05N
##                         Area.Description Level.Code Level.Description
## 1 Atlantic ocean, north of latitude 5° N        ESP             Spain
## 2 Atlantic ocean, north of latitude 5° N        FRA            France
## 3 Atlantic ocean, north of latitude 5° N        GBR    United Kingdom
##   Initial.Quantity Adapted.Quota Eurlex.Ref OJ.Ref Publication.Date
## 1         17690.58      15690.59 32015R0104    L22       28/01/2015
## 2          4421.71       6451.71 32015R0104    L22       28/01/2015
## 3           195.89        165.89 32015R0104    L22       28/01/2015
##   Page.Number In.regulation Compute.uptake  StockID X.ECTAC
## 1         112             Y              Y ALBAN05N     66%
## 2         112             Y              Y ALBAN05N     16%
## 3         112             Y              Y ALBAN05N      1%
```

```
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(as.character(DFM$Y1))
DFM$Y2 <- as.numeric(as.character(DFM$Y2))
DFM$Y3 <- as.numeric(as.character(DFM$Y3))
DFM <- DFM[!is.na(DFM$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 Scientific_name   English_name
## 1      AAS  DK 27_3_C_22  0 NA NA Denmark Astacus astacus Noble crayfish
## 2      AAS  DK    27_3_A NA NA NA Denmark Astacus astacus Noble crayfish
## 3      AAS  DK      27_3  0 NA NA Denmark Astacus astacus Noble crayfish
##                French_name
## 1 Écrevisse à pieds rouges
## 2 Écrevisse à pieds rouges
```

```
## 3 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 Regions 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$areaBis[i]), split=',')
    DT <- DFM[DFM$Scientific_name %in% table1A$latinName[i] & DFM$area %in% reg[[1]],]
    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='2013-2015',spp=table1A$latinName[i],region=table1A$region[i],
        RFMO=RFMO, area = table1A$area[i],select=NA, landings=NA, TAC=NA,shareLanding=NA,Thresh='M', Co
    ind <- which(DT$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, na.rm=TRUE)
        } else {
        T1$landings <- 0
        T1$shareLanding <- 0
        }

    ## TAC
    if (!(table1A$stockID[i] %in% 'No TAC')) {
        aa<-strsplit(as.character(table1A$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$Initial.Quantity
        TT <-  TT[names(TT) %in% GEO$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')
        }
    }
    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] <- 'O'  #rule (a)
    T1A$Thresh[is.na(T1A$TAC) & T1A$shareLanding <.1] <- 'O' #rule (b)
    T1A$Thresh[T1A$landings < 200] <- 'O'  #rule (c)
    T1A$Thresh[T1A$TAC <.1 & T1A$Comments >=.25] <- 'C'   # 25% rule, sampling to be coordinated betwee
}
```

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
T1B$Comments <- paste(round(100*T1B$Comments,0),'%',sep='')
T1B$Comments[T1B$Comments %in% c('NA%','NaN%','Inf%')] <- '-'
T1B$select <- '-'
T1B$select[T1B$Thresh %in% c('C','M')] <- 'Yes'
T1B$select[T1B$Thresh %in% 'O'] <- 'No'
T1B[T1B$spp %in% 'Nephrops norvegicus' & !(grepl('TAC', T1B$area)),'TAC'] <- '-'
levels(T1B$refYears) <- c(levels(T1B$refYears), '2015')
T1B[T1B$spp %in% 'Nephrops norvegicus' & !(grepl('TAC', T1B$area)),'refYears'] <- '2015'
T1B[T1B$RFMO %in% 'GFCM', 'refYears'] <- '2015'
```

Export of Table 1A

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

```
write.table(T1B, file=paste(CTRY,'_table1A_filled.csv',sep=''), sep=';',row.names=FALSE, quote=FALSE)
```