```
#Produce groundwater deficit indicator datasets

library(tidync)
library(data.table)
library(tidyverse)
library(lubridate)
library(zyp)
library(sf)
library(raster)
library(terra)
library(RColorBrewer)
library(rasterVis)
library(xts)
library(stars)
library(fasterize)
library(foreign)
library(RcppRoll)

#########################################################################
######################
###Load all the GWS datasets
proj_dir = "~/Dropbox/WB/GRACE_Ensemble/"
pathOut =  "/Users/tejasvi/Dropbox/WB/GRACE-Deficit/"
pathIn = '/Users/tejasvi/Dropbox/gwflagship_typologies/'
pathData = '/Users/tejasvi/Dropbox/gwflagship_GRACEdownscaling/
Downscaled TWS_GWS v2/'

#########################################################################
######################
####Load World Regions
wb_regions =
  st_read(paste0(proj_dir, "Spatial Files/WB_Regions/
WB_countries_Admin0_10m.shp")) %>%
  dplyr::select(WB_NAME, ISO_A2, ISO_A3, ISO_N3, TYPE, REGION_WB) %>%
  filter(TYPE != 'Dependency') %>%
  st_make_valid()

gws_mean_05 =
  tidync(paste0(pathData, 'GWS_mean_05deg.nc')) %>%
  hyper_tibble() %>% as.data.table()

#Create a date sequence
date.seq =
  seq(as.Date("2003/2/1"), as.Date("2021/9/1"), "month") %>%
  as.data.table() %>%
  rownames_to_column() %>%
  .[, rowname := as.integer(rowname)]
colnames(date.seq) = c('rowname', 'date')

#Merge the dates using the rowname column
```

```
gws_mean_05 =
  gws_mean_05 %>%
  merge(date.seq, by.x = 'time', by.y = 'rowname', all.x = T) %>%
  .[, ':='(yearmon = substr(date, 1, 7),
           year = substr(date, 1, 4),
           month = substr(date, 6, 7),
           GWSA_mean = (GWSA_CLSM+GWSA_Noah)/2,
           cell_id = paste0(lat, lon))]

gws.unique =
  gws_mean_05 %>%
  dplyr::select(lat, lon) %>%
  distinct() %>%
  rownames_to_column()



#Load the fishnet
fishnet =
  st_read('/Users/tejasvi/Dropbox/WB/Fishnet_halfdegree/
global_fishnet.shp')

# fishnet.centroid =
#   fishnet %>%
#   filter(!row_number() %in% c(18901, 23763, 26074, 42334, 31737))
%>% #Invalid polygon
#   st_make_valid() %>%
#   st_centroid()

fishnet.r =
  fishnet %>%
  st_drop_geometry() %>%
  as.data.table()

fishnet.r =
  rasterFromXYZ(fishnet.r[,.(Lon, Lat)])

crs(fishnet.r) = crs(wb_regions)

#Load aquifer typology data
typ <-
  read.dta(file=paste0(pathIn, "data_outputs/
aqtyp_gwresource_grid05deg.dta"), convert.factors = TRUE) %>%
  as_tibble %>%
  mutate(aqtyp=factor(aqtyp_max, levels=c("Major
Alluvial","Complex","Karstic","Local/Shallow"))) %>%
  as.data.table()

plot1 =
  rasterFromXYZ(typ[,.(lon, lat, aqtyp_pct_NA)])
```

```
#########################################################################
######################
#Trend Estimates
#########################################################################
######################
trends =
  gws_mean_05 %>%
  filter(year<2021) %>%
  merge(gws.unique, by = c('lat', 'lon'), all.x=T) %>%
  rename(GWS = GWSA_mean) %>%
  group_by(lat, lon, year) %>%
  summarise(GWS.year = mean(GWS, na.rm = T)) %>%
  group_by(lat, lon) %>%
  group_modify(~as.data.frame(t(zyp.yuepilon(.x$GWS.year[1:100])
[c(2,6)]))) %>%
  as.data.table()

trend.neg.sig =
  trends %>%
  mutate(neg_sig = ifelse(trend< 0 & sig<=0.05, 1, 0))

sum(trend.neg.sig$neg_sig, na.rm = T)

#########################################################################
######################
#GGDI steps
#########################################################################
######################
GGDI =
  gws_mean_05 %>%
  filter(year<2021) %>%
  merge(gws.unique, by = c('lat', 'lon'), all.x=T) %>%
  rename(GWS = GWSA_mean) %>%
  group_by(rowname, month) %>% #Group by month and cell to get
climatology
  mutate(GWS.month.mean = mean(GWS, na.rm = T)) %>%
  rowwise() %>%
  mutate(GWS.climatology = GWS - GWS.month.mean) %>%
  group_by(rowname) %>%
  mutate(grid.mean = mean(GWS.climatology, na.rm = T),
         grid.sd = sd(GWS.climatology, na.rm = T)) %>%
  mutate(GWS.deficit = (GWS.climatology-grid.mean)/grid.sd) %>%
  mutate(GWS.def.roll12 = roll_mean(GWS.deficit, 12, align = 'right',
fill = NA),
          GWS.def.roll18 = roll_mean(GWS.deficit, 18, align = 'right',
fill = NA),
          GWS.def.roll24 = roll_mean(GWS.deficit, 24, align = 'right',
fill = NA),
```

```
          GWS.def.roll30 = roll_mean(GWS.deficit, 30, align = 'right',
fill = NA)) %>%
  as.data.table()

plot(GGDI[cell_id == '23.7588.75']$GWS.def.roll24) #West Bengal/
Bangladesh
plot(GGDI[cell_id == '11.2579.25']$GWS.def.roll24) #South India
(Chennai)
plot(GGDI[cell_id == '13.2577.75']$GWS.def.roll24) #South India
(Kolar/Bangalore)
plot(GGDI[cell_id == '-23.7516.75']$GWS.def.roll24) #Southern Africa



#Basic Indicator -  Is the GW Deficit value between 2019-2020 less
than -1.5
GGDI.binary1 =
  GGDI %>%
  filter(year>=2019) %>%
  dplyr::select(rowname, GWS.deficit) %>%
  group_by(rowname) %>%
  summarise(GWS.deficit.mean = mean(GWS.deficit, na.rm = T)) %>%
  mutate(Def.19_20_neg15 = ifelse(GWS.deficit.mean > -1.5, 0, 1),
         Def.19_20_neg1 = ifelse(GWS.deficit.mean > -1, 0, 1)) %>%
  as.data.table()

#Indicator2 -  Is there any period between 2002 and 2017 where the GW
Deficit value is less than -1.5
GGDI.binary2 =
  GGDI %>%
  filter(year>=2010 & year<=2017) %>%
  dplyr::select(rowname, GWS.def.roll12, GWS.def.roll18,
GWS.def.roll24, GWS.def.roll30) %>%
  mutate(GWS.binary12 = ifelse(GWS.def.roll12 > -1.25, 0, 1),
         GWS.binary18 = ifelse(GWS.def.roll18 > -1.5, 0, 1),
         GWS.binary24_150 = ifelse(GWS.def.roll24 > -1.5, 0, 1),
         GWS.binary24_100 = ifelse(GWS.def.roll24 > -1, 0, 1)) %>%
  group_by(rowname) %>%
  summarise(Def.total12 = sum(GWS.binary12, na.rm = T),
            Def.total18 = sum(GWS.binary18, na.rm = T),
            Def.total24_150 = sum(GWS.binary24_150, na.rm = T),
            Def.total24_100 = sum(GWS.binary24_100, na.rm = T)) %>%
  mutate(Def.bin12 = ifelse(Def.total12>0, 1, 0),
         Def.bin18 = ifelse(Def.total18>0, 1, 0),
         Def.bin24_150 = ifelse(Def.total24_150>0, 1, 0),
         Def.bin24_100 = ifelse(Def.total24_100>0, 1, 0)) %>%
  as.data.table()


#Merge indicators
```

```
GGDI.out =
  GGDI.binary1 %>%
  merge(GGDI.binary2, by = 'rowname', all.x = T) %>%
  merge(gws.unique, by = 'rowname', all.x = T) %>%
  merge(typ[,c('lat', 'lon', 'aqtyp_max')], by.x = c('lat', 'lon'),
by.y = c('lat', 'lon'), all.x = T) %>%
  #Basically, making sure that hotspots in 2019 remain hotspots in the
rolling mean indicator
  mutate(Def.bin24_150 = ifelse(Def.19_20_neg15==1, 1, Def.bin24_150),
         Def.bin24_100 = ifelse(Def.19_20_neg1==1, 1, Def.bin24_100))

GGDI.out.sub =
  GGDI.out %>%
  dplyr::select(1:6, Def.bin24_100, Def.bin24_150, aqtyp_max) %>%
  merge(trend.neg.sig, by = c('lat', 'lon'), all.x = T)


#########################################################################
######################
#Plots
#########################################################################
######################

plot1 =
  rasterFromXYZ(GGDI.out[,.(lon, lat, Def.19_20_neg1)])

crs(plot1) = crs(fishnet.r)

plot1 =
  crop(plot1, extent(-180, 180, -60, 60)) %>%
  rast()

cls <- data.frame(id=0:1, cover=c("No Deficit", "Deficit"))
levels(plot1) <- cls


#color_pal =
c('#b2182b','#d6604d','#f4a582','#fddbc7','#d1e5f0','#92c5de','#4393c3
','#2166ac')

#Save the plot for future reference
plot_name =
  paste0("~/Dropbox/WB/GRACE-Deficit/Figures/",
"GW_Deficit_19_20_dscl.png")

png(plot_name, width = 1250, height = 500)

plot(plot1,
     col = c('#4393c3', '#d6604d'), plg=list(cex=1.2))
```

```
dev.off()

#####
plot2 =
  rasterFromXYZ(GGDI.out.sub[,.(lon, lat, neg_sig)])

crs(plot2) = crs(fishnet.r)

plot2 =
  crop(plot2, extent(-180, 180, -60, 60)) %>%
  rast()

cls <- data.frame(id=0:1, cover=c("No Deficit", "Deficit"))
levels(plot2) <- cls


#color_pal =
c('#b2182b','#d6604d','#f4a582','#fddbc7','#d1e5f0','#92c5de','#4393c3
','#2166ac')

#Save the plot for future reference
plot_name =
  paste0("~/Dropbox/WB/GRACE-Deficit/Figures/",
"GW_Trend_ind_dscl.png")

png(plot_name, width = 1250, height = 500)

plot(plot2,
     col = c('#4393c3', '#d6604d'), plg=list(cex=1.2))

dev.off()


#####

plot3 =
  rasterFromXYZ(GGDI.out[,.(lon, lat, Def.bin24_100)])

crs(plot3) = crs(fishnet.r)

plot3 =
  crop(plot3, extent(-180, 180, -60, 60))  %>%
  rast()

cls <- data.frame(id=0:1, cover=c("No Deficit", "Deficit"))
levels(plot3) <- cls


#color_pal =
c('#b2182b','#d6604d','#f4a582','#fddbc7','#d1e5f0','#92c5de','#4393c3
```

```
','#2166ac')


#Save the plot for future reference
plot_name =
  paste0("~/Dropbox/WB/GRACE-Deficit/Figures/",
"GW_Deficit_24_month_binary_dscl_10_20.png")

png(plot_name, width = 1250, height = 500)

plot(plot3,
     col = c('#4393c3', '#d6604d'), plg=list(cex=1.2))


dev.off()


#########################################################################
######################
#Distribution of hotspots by aq typ
table(GGDI.out[!aqtyp_max %in% c("", NA) & Def.19_20==1,]$aqtyp_max)
%>% prop.table()
table(GGDI.out[!aqtyp_max %in% c("", NA) &
Def.bin24_150==1,]$aqtyp_max, useNA = 'no') %>% prop.table()

#Overall
table(GGDI.out$Def.19_20,
      GGDI.out$Def.bin24_150,
      dnn = c("Def.19_20", "Def.bin24_150"), useNA = 'no')

#Without NA
table(GGDI.out[!WB_REGION %in% c("Other"),]$Def.19_20,
      GGDI.out[!WB_REGION %in% c("Other"),]$Def.bin24_150,
      dnn = c("Def.19_20", "Def.bin24_150"), useNA = 'no') %>%
prop.table()


#########################################################################
######################
#Write Files
#########################################################################
######################


pathOut =  "/Users/tejasvi/Dropbox/WB/GRACE-Deficit/"

fwrite(GGDI.out.sub, paste0(pathOut, 'GGDI_output_dscl_230128.csv'))

#
# st_write(GGDI.fishnet,
```

```
#           paste0(pathOut, 'GWS_Deficit_nonDownscaled_05degree.shp'),
#           delete_layer = T)
#
```