Data mobilisation from GBIF to the EBV Data Portal

Inputs for occurrence cubes rasterisation using the 10 Km EEA vector grid

true

2024-08-20

Introduction

This is a preparatory notebook for mobilising species occurrence cubes from GBIF—the Global Biodiversity Information Facility— to the EBV Data Portal. Specifically, here we use the European Environment Agency (EEA) grid in vector format at 10 km as spatial reference. The subsequent notebooks focus on the generation of species cubes, the calculation of metrics, and the conversion of the results into the EBV cube format for final upload to the EBV data portal.

Note: This series of notebooks is part of the results of Task 3.3 of the Biodiversity Building Blocks for Policy project funded by the European Union's Horizon Europe Research and Innovation Programme (ID No 101059592). Additional notebooks exploring the results and calculating simple metrics are also available in the same repository.

Load Libraries

```
rm(list=ls())
gc()
```

```
##          used (Mb) gc trigger  (Mb) max used  (Mb)
## Ncells 1848789 98.8    9199432 491.4 11612665 620.2
## Vcells 3860036 29.5   43045405 328.5 70740088 539.8
```

```
# Load requiered libraries
library(here)
library(sf)
library(terra)
library(dplyr)
```

Load EEA vector grid

European Union grids are available at different spatial scales in vector format. The EEA vector grids can be downloaded from the Data Hub website. We start by loading the EEA vector grid at the desired spatial resolution. For this, we use the 10 km grid.

```
# Load EEE grid vector file
resx <- "10K" # "10K" or "100K" # define the spatial resolution of the vector file
gridin <- st_read("C:/data/grid/eea_v_3035_10_km_eea-ref-grid-europe_p_2011_v01_r00/Grid_ETRS89-LAEA_10K
```

```
## Reading layer 'Grid_ETRS89-LAEA_10K' from data source
##   'C:\data\grid\eea_v_3035_10_km_eea-ref-grid-europe_p_2011_v01_r00\Grid_ETRS89-LAEA_10K.shp'
##   using driver 'ESRI Shapefile'
## Simple feature collection with 571200 features and 3 fields
## Geometry type: POLYGON
## Dimension:     XY
## Bounding box:  xmin: 0 ymin: 7e+05 xmax: 8400000 ymax: 7500000
## Projected CRS: ETRS89-extended / LAEA Europe
```

We then rasterise the EEA vector grid and save it as a tif file. This file will later be used as a grid reference for rasterisation of the metrics.

```
# Calculate the number of rows and columns corresponding respectively to latitude and longitude
eo <- length(unique(gridin$EofOrigin))
no <- length(unique(gridin$NofOrigin))

# Create raster from vector file
eear <- rast(gridin, nrows=no, ncols=eo, nlyrs=1, crs(gridin), vals=gridin$NofOrigin)
plot(eear)
```
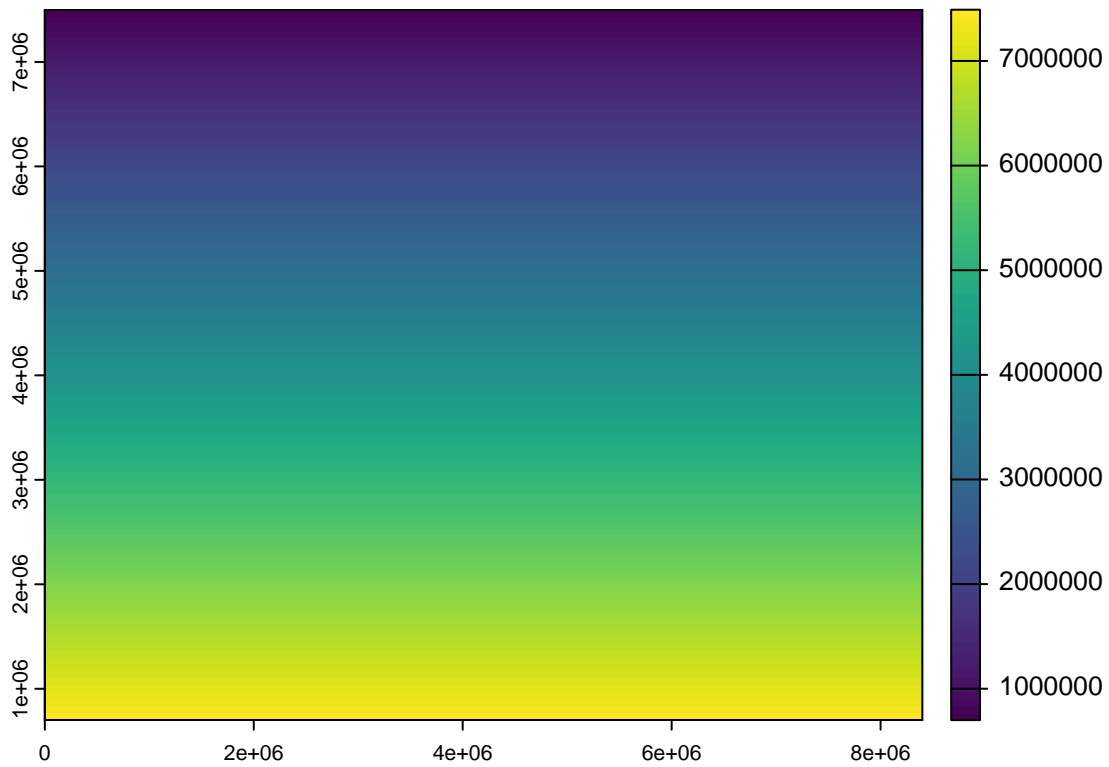
```
# Save raster
# writeRaster(eear, here(paste0("input/grid/eeagrid_", resx, ".tif")), datatype = 'INT4S', overwrite=TR
```

As a last step, we calculate the centroids of the EEA grid to easily link the occurrence cubes, in CSV format, with the corresponding latitude and longitude coordinates of the respective cell. For this we will use the eeacellcode.

```
# Estimate the EEA grid centroids from the vector file
gridxy <- st_coordinates(st_centroid(gridin))

# Add the coordinates to the reference grid
gridin$x <- gridxy[,1]
gridin$y <- gridxy[,2]

# Create a data frame from the polygons attribute table
eeax <- as.data.frame(gridin)

# Substet the dataframe columns to export
eeacoor <- eeax[,c("CellCode", "x", "y")]
colnames(eeacoor) <- c("eeacellcode", "x", "y")
eeacoor[1:5,]
```

```
##   eeacellcode     x      y
## 1   10kmE0N70  5000 705000
## 2   10kmE1N70 15000 705000
## 3   10kmE2N70 25000 705000
## 4   10kmE3N70 35000 705000
## 5   10kmE4N70 45000 705000
```

```
write.csv(eeacoor, here(paste0("input/grid/centroids/eeagrid_centroids_", resx, ".csv"))) # overwrite=T
```