BIODIVERSITY BUILDING BLOCKS FOR POLICY

M25 Code to calculate the impact indicator

08/11/2024

Author(s): Mukhtar Muhammed Yahaya, Sabrina Kumschick, Sandra MacFadyen, Pietro Landi & Cang Hui

Funded by the European Union

This project receives funding from the European Union's Horizon Europe Research and Innovation Programme (ID No 101059592). Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them.

Table of contents

Summary

We report the GitHub repository containing the code to compute impact indicators of alien taxa (https://github.com/mmyahaya/impact.indicator) and demonstrate its implementation using Acacia species in South Africa. This code adopts methods of computing impact indicators discussed at the Mathematical Biosciences Hub, Stellenbosch University workshop titled "Quantifying potential impact of biological invasions using occurrence cubes and consensus-based impact categories" held on 21 October 2024.

The impact of invasive alien species is recognised as one of the leading drivers of biodiversity loss, disrupting native ecosystems and threatening conservation efforts. To understand the consequences of invasive species on our ecosystem, studies are being conducted to assess the extent of the damage caused by invasive species. A commonly used assessment tool, the Environmental Impact Classification for Alien Taxa (EICAT), has been adopted by the International Union for Conservation of Nature (IUCN) to categorise the severity of impacts across species. EICAT classifies species with sufficient data into five impact levels, ranging from minimal concern (MC) to massive (MV). Given the mechanisms of alien species impacts such as competition, predation, or parasitism, species may receive various impact ratings depending on their mechanisms of impact and geographic contexts Also, each site can have multiple species with different impact categories which requires a method to aggregate the impact for the site. Furthermore, sites in a region can have different impact scores aggregated across multiple species in each site which also requires aggregating for a study region.

This code incorporates several methods of aggregating impact scores within and across species and across sites which have been used and proposed by various studies to arrive at an impact indicator of alien species in a region. Each aggregation method is tailored to address specific research questions and contextual needs. The code utilises GBIF occurrence data and EICAT information for user-selected taxa to produce impact risk maps and impact indicators of alien species.

List of abbreviations

EU	European Union
EICAT	Environmental Impact Classification for Alien Taxa
IUCN	International Union for Conservation of Nature
MC	Minimal Concern
MV	Massive

1. Demonstration of code with R markdown

1.1.Description

This R Markdown demonstrates the computation of an impact indicator for biological invasions using the `impact_indicator()`. The `impact_indicator()` feeds in species occurrence cube from the `b3gbi::process_cube()` using `taxaFun()` and processed Environmental Impact Classification of Alien Taxa (EICAT) impact score of species using `impact_cat()`.

1.2.Process occurrence cube

```
# load GBIF occurrence data for taxa
taxa_Acacia<-readRDS(paste0(getwd(),"/Data/taxa_Acacia.rds"))
acacia_cube<-taxaFun(taxa=taxa_Acacia,
                     country.sf=SA.sf,
                     res=0.25,
                     first_year=2010)

acacia_cube$cube
```

```
##
## Simulated data cube for calculating biodiversity indicators
##
## Date Range: 2010 - 2024
## Number of cells: 369
## Grid reference system: custom
## Coordinate range:
## [1] "Coordinates not provided"
##
## Total number of observations: 5558
## Number of species represented: 25
## Number of families represented: Data not present
##
## Kingdoms represented: Data not present
##
## First 10 rows of data (use n = to show more):
##
## # A tibble: 5,558 × 7
##    scientificName   taxonKey iucnRedListCategory
minCoordinateUncertaint…¹  year
##    <chr>               <dbl> <chr>                                <dbl>
<dbl>
##  1 Acacia implexa    2979232 LC                                       1
2010
##  2 Acacia cyclops    2980425 LC                                     122
2010
##  3 Acacia saligna    2978552 LC                                       1
2010
##  4 Acacia pycnantha  2978604 LC                                       1
2010
```

```
##  5 Acacia mearnsii   2979775 NE                                        110
2010
##  6 Acacia mearnsii   2979775 NE                                          1
2010
##  7 Acacia mearnsii   2979775 NE                                          8
2010
##  8 Acacia saligna    2978552 LC                                          1
2011
##  9 Acacia saligna    2978552 LC                                         15
2011
## 10 Acacia mearnsii   2979775 NE                                          1
2011
## # ℹ 5,548 more rows
## # ℹ abbreviated name: ¹minCoordinateUncertaintyInMeters
## # ℹ 2 more variables: cellCode <int>, obs <dbl>
```

```
head(acacia_cube$coords)
```

```
##   siteID        X       Y
## 1      1 16.60833 -34.697
## 2      2 16.85833 -34.697
## 3      3 17.10833 -34.697
## 4      4 17.35833 -34.697
## 5      5 17.60833 -34.697
## 6      6 17.85833 -34.697
```

1.3. Aggregate impact scores for each species

There are often several impact records per species in different mechanisms and regions. There is a need to aggregate these impact records for each species. The `impact_cat()` aggregates impact using ***max***, ***mean*** and ***max_mech*** as metrics as used by different studies.

- ***max***: maximum impact score across all records for the species
- ***mean***: mean impact score across all records
- ***max_mech***: sum of the maximum impact per mechanisms

```
eicat_data<-readRDS(paste0(getwd(),"/Data/eicat_data.rds"))
full_species_list<-sort(unique(acacia_cube$cube$data$scientificName))
eicat_data %>% select(scientific_name,impact_region,impact_mechanism,
                      impact_category) %>%
  head(10)
```

```
## # A tibble: 10 × 4
##    scientific_name     impact_region            impact_mechanism
impact_category
##    <chr>               <chr>                    <chr>            <chr>
##  1 Dryophytes cinereus <NA>                     <NA>             DD - Data
defi…
##  2 Acacia cyclops      Cape Province            (9) Chemical Im… DD -
Data defi…
```

```
##  3 Acacia cyclops      Melkbosstrand            (9) Chemical Im... DD -
Data defi...
##  4 Acacia cyclops      Agulhas Plain            (9) Chemical Im... MR -
Major
##  5 Acacia cyclops      South-western part of t... (11) Structural... MO -
Moderate
##  6 Acacia cyclops      Cape of Good Hope Natur... (11) Structural... MO -
Moderate
##  7 Acacia cyclops      Millers Point            (9) Chemical Im... MC -
Minimal c...
##  8 Acacia cyclops      Cape Province            (10) Physical I... MN -
Minor
##  9 Acacia cyclops      Western Cape             (11) Structural... MO -
Moderate
## 10 Acacia dealbata     Ourense (NW of Spain)    (9) Chemical Im... MR -
Major
```

```
agg_impact<-impact_cat(impact_data=eicat_data,
                       species_list=full_species_list,
                       col_impact=NULL,
                       col_name=NULL)
agg_impact
```

```
##                         max     mean max_mech
## Acacia acinacea          NA       NA       NA
## Acacia adunca            NA       NA       NA
## Acacia baileyana         NA       NA       NA
## Acacia binervata         NA       NA       NA
## Acacia crassiuscula      NA       NA       NA
## Acacia cultriformis      NA       NA       NA
## Acacia cyclops            3 1.500000        6
## Acacia dealbata           3 1.812500       18
## Acacia decurrens          3 1.500000        6
## Acacia elata             NA       NA       NA
## Acacia falciformis       NA       NA       NA
## Acacia implexa           NA       NA       NA
## Acacia longifolia         3 1.562500       20
## Acacia mearnsii           3 1.769231       15
## Acacia melanoxylon       NA       NA       NA
## Acacia paradoxa          NA       NA       NA
## Acacia piligera          NA       NA       NA
## Acacia podalyriifolia    NA       NA       NA
## Acacia provincialis      NA       NA       NA
## Acacia pycnantha          3 3.000000        3
## Acacia saligna           NA       NA       NA
## Acacia schinoides        NA       NA       NA
## Acacia stricta           NA       NA       NA
## Acacia ulicifolia        NA       NA       NA
## Acacia viscidula         NA       NA       NA
```

1.4. Compute impact risk map

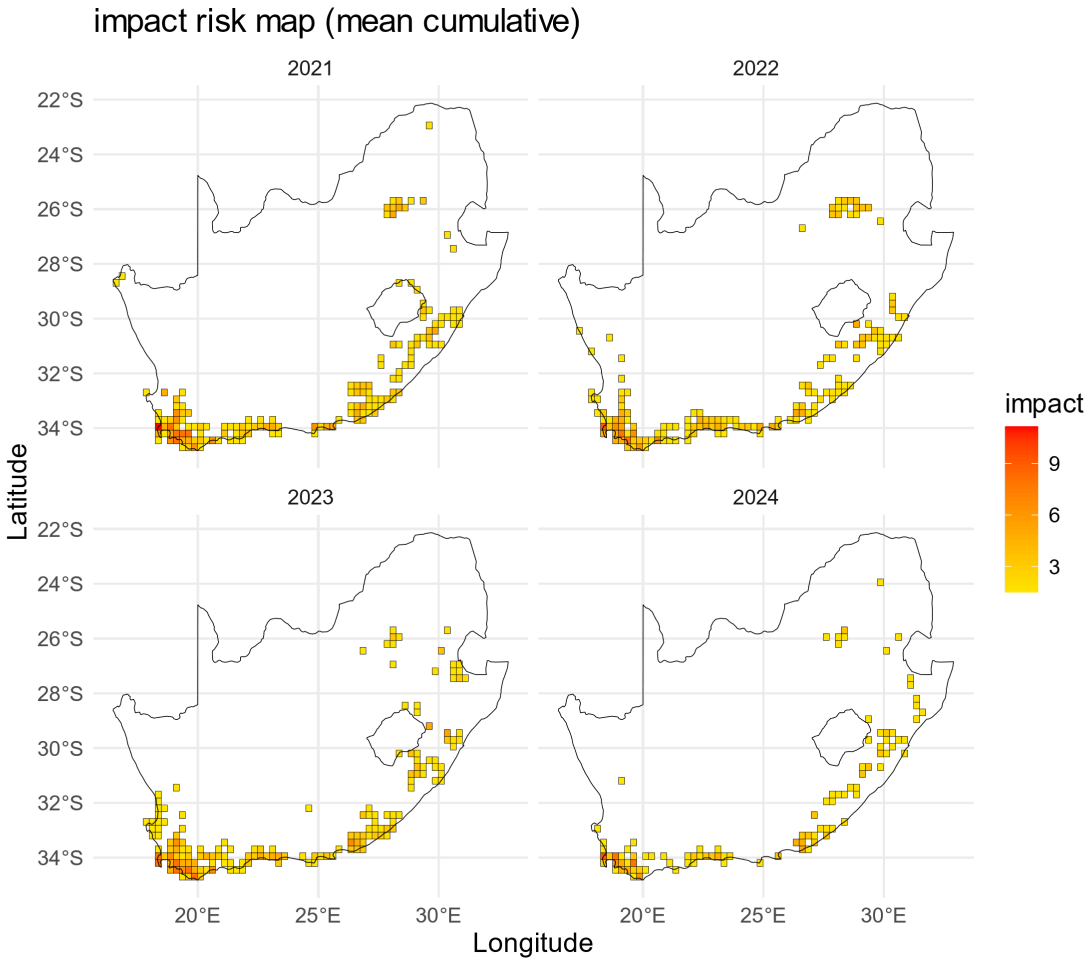
The impact risk map shows the impact score for each site, where multiple species can be present. To compute the risk per site, aggregated scores across species at each site are needed. The `impact_indicator()` uses ***max***, ***sum*** and ***mean*** metrics to aggregate impact scores as proposed by Boulesnane-Guengant et al., (in preparation). The combinations of aggregation metrics for each species and site lead to five types of indicators, namely, ***precautionary***, ***precautionary cumulative***, ***mean***, ***mean cumulative*** and ***cumulative***.

- ***precautionary***: maximum score across species' max in each site
- ***precautionary cumulative***: cumulative score across species' max in each site
- ***mean***: mean score across species' mean in each site
- ***mean cumulative***: cumulative score across species' mean in each site
- ***cumulative***: cumulative score across species' sum of maximum score per mechanism

```
impact_value<-impact_indicator(cube=acacia_cube$cube,
                               impact_data = eicat_data,
                               col_impact=NULL,
                               col_name=NULL,
                               type = "mean cumulative",
                               coords=acacia_cube$coords)
```

impact risk map

```
#impact risk map
#visualise last four years for readability
impact_value$sitedf%>%
  gather(-c(siteID,X,Y),key="year",value="impact") %>%
  na.omit() %>%
  filter(year>=2021) %>%
  ggplot() +
  geom_tile(
    aes(x=X,y=Y,fill=impact),color="black")+
  geom_sf(data = SA.sf, fill = NA, color = "black", alpha = 0.5)+
  scale_fill_gradient2(low = "forestgreen",
                       mid = "yellow",
                       high = "red")+
  theme_minimal() +
   labs(
    title = "impact risk map (mean cumulative)",
    y = "Latitude", x="Longitude"
  )+
  theme(text=element_text(size=14))+
  facet_wrap(~year)
```



1.5. Compute impact indicators

1.5.1.Impact indicators of alien taxa

To compute the impact indicator of alien taxa, we sum all the yearly impact scores of each site of the study region. To correct for sampling effort we divide the yearly impact scores by the number of sites in the study region with at least a single occurrence throughout the whole year.

$$I_i = \frac{\Sigma S_i}{N}$$

I_i is impact score at year i.

S_i is the sum of risk map value, where $S = \{s_1, s_2, ..., s_n\}$ and s_n is the site score for the site n

N is the number of sites occupied throughout the study years of the region.

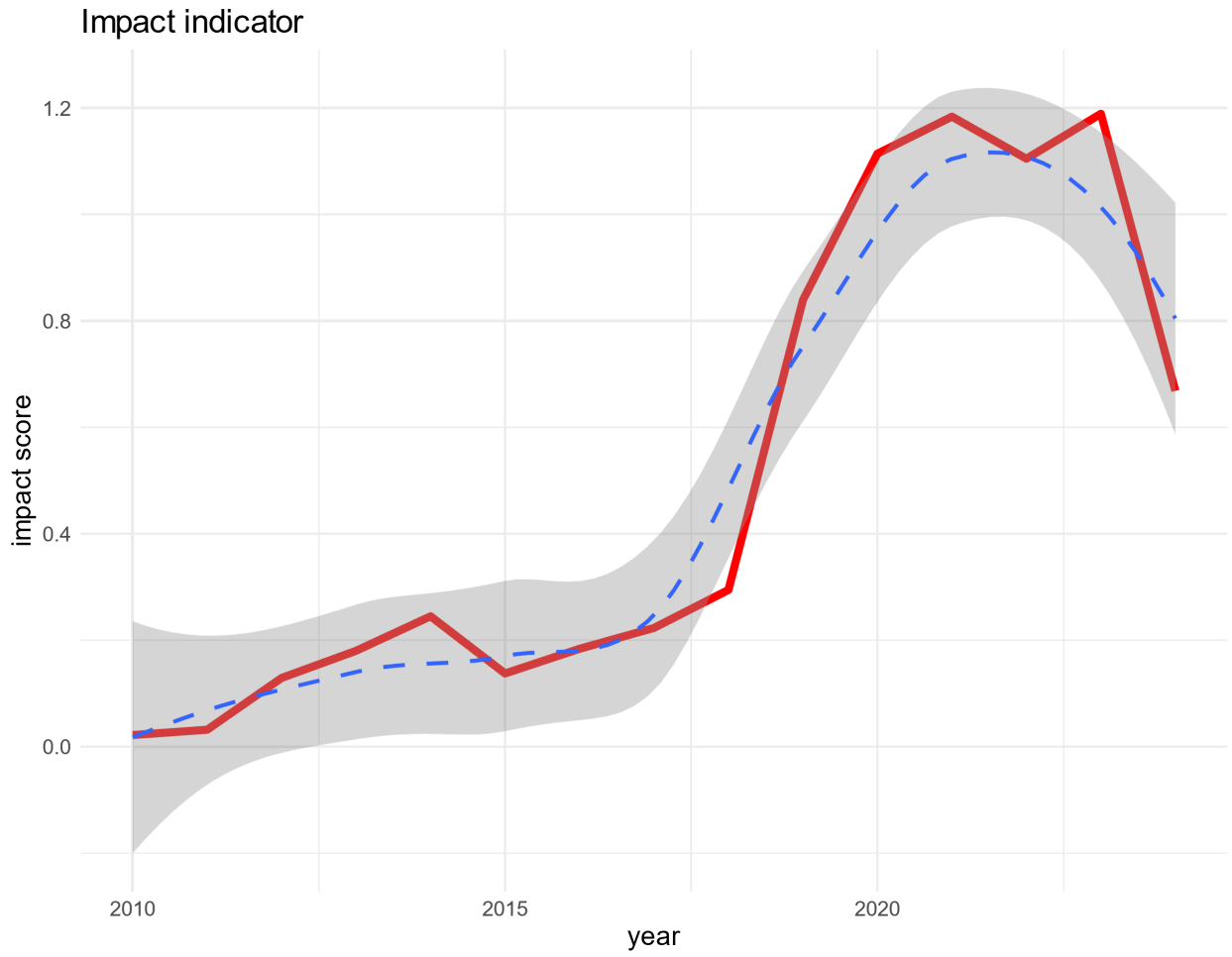
Note: This is the only method incorporated as of now. Other methods will be considered later.

Note: A function `impact_uncertainty()` is being developed to use the bootstrap method to compute the confidence interval of the indicator instead of `geom_smooth()` used below.

```
#sum of impact risk map for each year

ggplot(data = impact_value$impact_values) +
  geom_line(aes(y = value, x = year),colour="red",
            stat="identity",
            linewidth=2)+
  geom_smooth(aes(y = value, x = year),linetype=2)+
  labs(
    title = "Impact indicator",
    y = "impact score"
  )+
  theme_minimal() +
  theme(text=element_text(size=14))

## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```

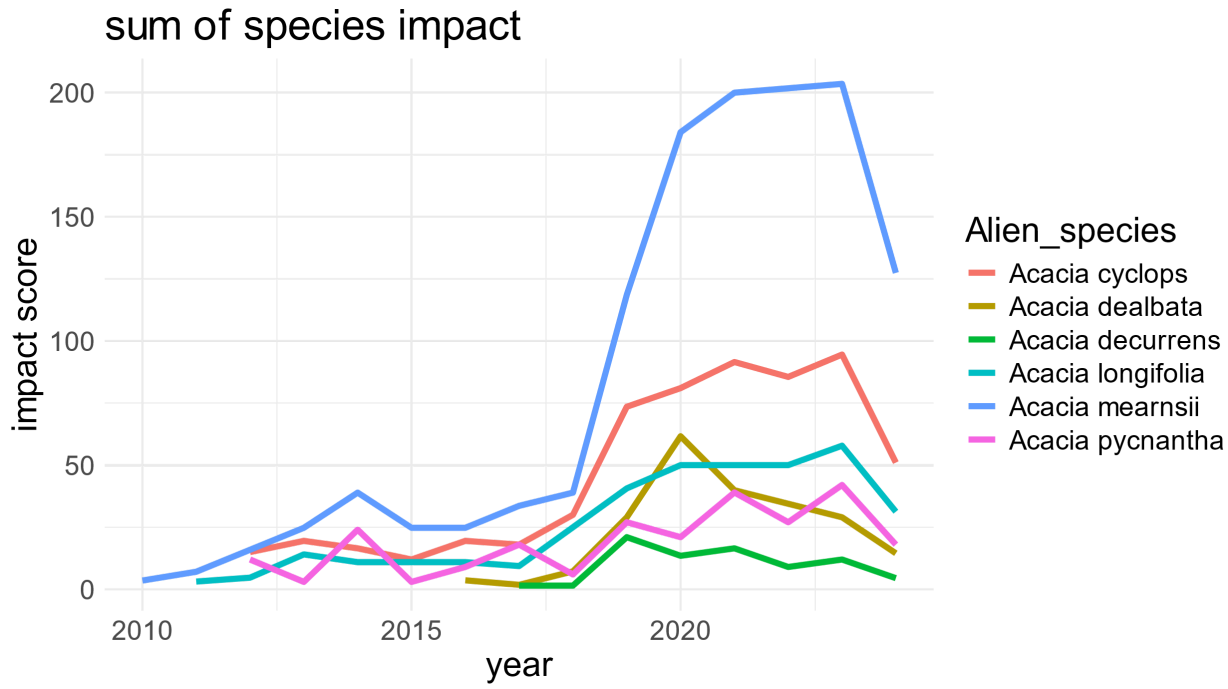


1.5.2.Impact indicator per species

We compute the impact indicator per species by summing the impact risk map per species and correct for sampling effort by dividing by N.

```
#  impact indicator per species
impact_value$species_values %>%
  rownames_to_column("year") %>%
  mutate(year=as.numeric(year)) %>%
  gather(-year,key = "Alien_species", value = "impact_score") %>%
  ggplot(aes(x = year, y = impact_score)) +
  geom_line(aes(color = Alien_species),linewidth=1.5)+
  theme_minimal() +
  labs(
    title = "sum of species impact",
    y = "impact score"
  )+
  theme(text=element_text(size=14))
```

```
## Warning: Removed 5 rows containing missing values or values outside the
scale range
## (`geom_line()`).
```



1.5.3.Comparing type of indicators

To compare the type of impact indicators for a case study, we provide a plot which can be adapted by a user to compare a set of methods.

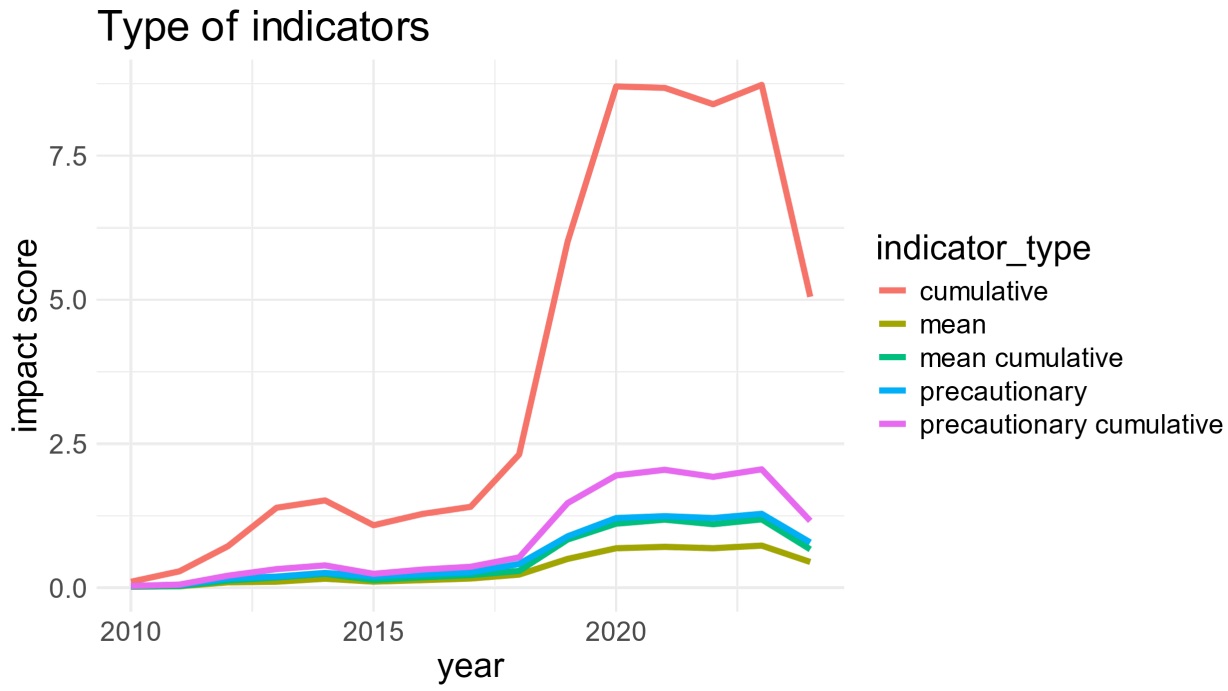
```
# plot all type of impact indicators
types<-c("precautionary",
         "precautionary cumulative",
```

```
         "mean",
         "mean cumulative",
         "cumulative")

all_impact<-data.frame("year"=unique(acacia_cube$cube$data$year))
for(type in types){
  impact_value<-impact_indicator(cube=acacia_cube$cube,
                                 impact_data = eicat_data,
                                 col_impact=NULL,
                                 col_name=NULL,
                                 type = type,
                                 coords=acacia_cube$coords)

  all_impact[type]<-impact_value$impact_values$value
}

all_impact %>%
  gather(-year,key = "indicator_type", value = "impact_score") %>%
  ggplot(aes(x = year, y = impact_score)) +
  geom_line(aes(color = indicator_type),linewidth=1.5)+
  theme_minimal() +
   labs(
    title = "Type of indicators",
    y = "impact score"
  )+
  theme(text=element_text(size=14))
```



2. Link to code repository

https://github.com/mmyahaya/impact.indicator

3. Acknowledgements

We appreciate the inputs received from John Wilson, Maarten Trekels, Tsungai Zengeya, Dave Richardson, Tammy Robins, Christiaan Gildenhuys, Lorenzo Ruaro at the Mathematical Biosciences Hub, Stellenbosch University workshop titled "Quantifying potential impact of biological invasions using occurrence cubes and consensus-based impact categories" held on 21 October 2024.