Package 'occUncertain'

January 20, 2023

Type Package

Title Addressing Occurrence Point Uncertainty When Calculating Spatial Metrics

Version 0.1.0

Date 2023-01-14

Description Repeatable processing of species occurrence datasets that makes it easier to propagate georeferencing imprecisions and data input mistakes to downstream analyses, allowing analysts to assess
the impacts of these imprecisions in quantifying range of occurrence (EOO) or area of occupancy (AOO) . Users can use the software to:
(a) change each coordinate record's uncertainty from meters to decimal degrees,
The formula for converting from meters to decimal degrees is in part based on information from the ESRI ArcUser magazine ``Measuring in Arc-
Seconds'' at this site<https://www.esri.com/news/arcuser/0400/wdside.html>
(b) deal with records that don't have uncertainty values in multiple ways,
(c) create a new random location for each occurrence using a uniform distribution with a defined interval
within the occurrence location uncertainty, and
(d) use repetitions to quantify EOO and AOO with attribute uncertainty.

BugReports https://github.com/mlammens/occUncertain/issues

Depends R (>= 3.5.0), knitr, ConR, rgdal

License GPL-3

Encoding UTF-8

LazyData true

VignetteBuilder knitr

RoxygenNote 7.2.3

Suggests rmarkdown, testthat, ggplot2, dplyr

NeedsCompilation no

Author Samuel V. Chang Triguero [aut, cre],
Matthew E. Aiello-Lammens [aut]

Maintainer Samuel V. Chang Triguero <changalfsam@gmail.com>

Repository CRAN

Date/Publication 2023-01-20 10:10:06 UTC

R topics documented:

`generate_occ_uncertain` *Random latitude and longitude coordinates accounting uncertainty values*

Description

`generate_occ_uncertain` Given a data frame of georeferenced occurrences this function generates a new set of coordinates with added uncertainty.

Usage

```
generate_occ_uncertain(
  occs_df,
  lat_col = "latitude",
  lon_col = "longitude",
  lat_uncertainty = "latitude_uncertainty",
  lon_uncertainty = "longitude_uncertainty",
  taxa_col = "species"
)
```

Arguments

`occs_df`	A `data.frame` of occurrence locations that includes *at least these four columns* - latitude, longitude, latitude uncertainty and longitude uncertainty in degrees.
`lat_col`	Name of column of latitude dbl values. Caps sensitive.
`lon_col`	Name of column of longitude dbl values. Caps sensitive.
`lat_uncertainty`	Name of column of latitude uncertainty in degree values. Caps sensitive.
`lon_uncertainty`	Name of column of longitude uncertainty in degree values. Caps sensitive.
`taxa_col`	Name of column of taxa (species) values. Caps sensitive.

Details

Input as a `dataframe` should have the following structure:

[,1]	ddlat	numeric, latitude (in decimal degrees)
[,2]	ddlon	numeric, longitude (in decimal degrees)
[,3]	ddlat unc	numeric, longitude uncertainty (in decimal degrees)
[,4]	ddlon unc	numeric, longitude uncertainty (in decimal degrees)
[,5]	tax	character or factor, taxa names

It is mandatory to respect field positions, but field names do not matter

Value

random_dd A `data.frame` of a random latitude, random longitude and taxa name for each occurence record.

`Leopardus_wiedii_gbif` *Occurence data for* Leopardus wiedii *acquired from GBIF*

Description

The margay (Leopardus wiedii, Schinz 1821) is a small Neotropical cat that is native with distribution from South America into Northeastern Mexico. Citation - GBIF.org (January 2023) GBIF Occurrence Download http://api.gbif.org/v1/occurrence/download/request/0066556-200221144449610.zip

Usage

```
Leopardus_wiedii_gbif
```

Format

An object of class `data.frame` with 229 rows and 4 columns.

Details

DOI 10.15468/dl.ljhrm9

Source

https://www.gbif.org

`meters_to_decdeg` *Convert from meters to degrees correcting for global position*

Description

`meters_to_decdeg` converts from meters to degrees at a specified position on the globe. The use case this function was developed for was to calculate occurence point uncertainty values, which are usually reported in meters, as degrees.

The formula for converting from meters to decimal degrees is in part based on information from the ESRI ArcUser magazine at this site https://www.esri.com/news/arcuser/0400/wdside.html

Usage

```
meters_to_decdeg(
  occs_df,
  lat_col = "latitude",
  lon_col = "longitude",
  distance,
  na_action = "NA as 0"
)
```

Arguments

`occs_df`	A `data.frame` of occurrence locations that incudes *at least these three columns* - latitude, longitude, and a distance in meters to be converted to decimal degrees.
`lat_col`	Name of column of latitude values. Caps sensitive.
`lon_col`	Name of column of longitude values. Caps sensitive.
`distance`	Name of column of distance values, in meters. Caps sensitive.
`na_action`	Enact distance options for NA values. Caps sensitive

Value

dist_dd A `data.frame` of latitude and longitude distances in units of degree decimal.

`random_geo_range` *Random geographic occurrences and preliminary conservation status assessment following IUCN Criterion B. Species area of occupancy (AOO) and extent of occurrence (EOO), from latitude and longitude coordinates accounting uncertainty values*

Description

`random_geo_range` Given georeferenced coordinates and associated uncertainty. This function generates random statistics values (Extent of Occurrence, Area of Occupancy, number of locations, number of subpopulations) and provide a preliminary conservation status following Criterion B of IUCN. A graphical map output is also available.

Usage

```
random_geo_range(
  n_length,
  occs_df,
  lat_col = "latitude",
  lon_col = "longitude",
  lat_uncertainty = "lat_uncertainty",
  lon_uncertainty = "lon_uncertainty",
  taxa_col = "species",
  country_map = NULL,
  exclude.area = TRUE,
  method.range = "convex.hull",
  export_shp = FALSE,
  write_shp = FALSE,
  map_pdf = FALSE,
  draw.poly.EOO = TRUE,
  Cell_size_AOO = 2,
  Cell_size_locations = 10,
  DrawMap = TRUE,
  add.legend = TRUE,
  write_results = FALSE,
  write_file_option = "excel"
)
```

Arguments

n_length	Number of iterations
occs_df	a dataframe georeferenced occurrence
lat_col	latitude values decimal degrees column
lon_col	longitude decimal degrees column
lat_uncertainty	latitude uncertainty decimal degrees column
lon_uncertainty	longitude uncertainty decimal degrees column
taxa_col	character or factor, taxa names
country_map	a SpatialPolygonsDataFrame or SpatialPolygons showing for example countries or continent borders. This shapefile will be used for cropping the SpatialPolygons used for EOO computation if exclude.area is TRUE. By default, it is land
exclude.area	a logical, if TRUE, areas outside of country_map are cropped of SpatialPolygons used for EOO computation. By default, it is TRUE
method.range	a character string, if "convex.hull", EOO is based on a convex hull. if "alpha.hull", EOO is based on alpha hull of alpha value. By default, it is "convex.hull"
export_shp	a logical, if TRUE, shapefiles of SpatialPolygons used for EOO computation are exported. By default, it is FALSE

`write_shp` a logical, if TRUE, shapefiles of `SpatialPolygons` used for EOO computation are written as ESRI shapefiles in a sub-directory in the working directory. By default, it is FALSE

`map_pdf` a logical, if TRUE, maps are exported in one pdf file. Otherwise, each species map is exported in png. By default, it is FALSE

`draw.poly.EOO` a logical, if TRUE, the polygon used for estimating EOO is drawn. By default, it is TRUE

`Cell_size_AOO` a numeric, value indicating the grid size in kilometers used for estimating Area of Occupancy. By default, equal to 2

`Cell_size_locations`
a numeric, value indicating the grid size in kilometers used for estimating the number of location. By default, equal to 10

`DrawMap` a logical, if TRUE a map is produced for each species in png format, unless map_pdf is TRUE. By default, it is TRUE

`add.legend` a logical, if TRUE a legend and a submap showing distribution in 'country_map' are displayed for each map. By default, it is TRUE

`write_results` a logical, if TRUE, results are exported in a file which can csv or excel, see write_file_option. By default, it is FALSE

`write_file_option`
a character, if "excel", results are exported in excel file, if "csv", results are exported in csv. By default, it is "excel"

Details

Input as a `dataframe` should have the following structure:

[,1]	ddlat	numeric, latitude (in decimal degrees)
[,2]	ddlon	numeric, longitude (in decimal degrees)
[,3]	ddlat unc	numeric, longitude uncertainty (in decimal degrees)
[,4]	ddlon unc	numeric, longitude uncertainty (in decimal degrees)
[,5]	tax	character or factor, taxa names

It is mandatory to respect field positions, but field names do not matter Starting position of the raster used for estimating the Area Of Occupancy

Different starting position of the raster used for estimate the AOO may provide different number of occupied cells. Hence, by default, 4 different translations of the raster is done (fixed increment of 1/4 resolution north and east) and the minimum number of occupied cells is used for estimating AOO. It is also possible to define a given number of random starting position of the raster using the argument `nbe.rep.rast.AOO`

Estimating number of locations

Locations are estimated by overlaying a grid of a given resolution (see `Cell_size_locations` for

specifying the resolution). The number of locations is simply the number of occupied locations. Note that the grid position is overlaid in order to minimize the number of locations (several translation of the grid are performed and the one providing the minimum number of occupied cells is provided).

Taking into account protected area for estimating the number of locations
A location is defined by the IUCN as a "geographically or ecologically distinct area in which a single threatening event can affect all individuals of the taxon". A simple way to include threat level is to rely on a map of protected areas and assume that populations within and

outside protected areas are under different threat level.

If a map of protected area is provided, this one is used for estimating the number of locations by the following procedure:
- if `method_protected_area` is "no_more_than_one", all occurrences within a given protected area will be considered as one location. Occurrences outside protected area will be used for estimating the number of locations using overlaying grid as described above. See the vignette for illustration.
- if `method_protected_area` is NOT "no_more_than_one", number of locations will be estimated by the overlaying grid as described above, but by considering differently occurrences outside and inside protected area.

The protected areas layers should be given as as `SpatialPolygonsocc_randomFrame` in `protec.areas`. The `ID_shape_PA` should also be given and should represent the unique ID of each protected area in the provided shapefile. This can be checked by the following code:

`colnames(ProtectedAreas@data)` Where ProtectedAreas is the name of your shapefile.

Limitation in the estimations of EOO

For a species whose occurrences span more than 180 degrees, EOO is not computed. This is the case for example for species whose distribution span the 180th meridian.

Value

A `data.frame` of EOO, AOO, Category in Criterion B, Number of locations, Category of AOO, Category of EOO, Cateogry Code for each iteration.

Index