Package 'sperich'

July 18, 2023

Type Package

Title Auxiliary Functions to Estimate Centers of Biodiversity

Version 1.5-9

Date 2023-07-13

Author Maximilian Lange, Sven Lautenbach, Claudia Raedig

Maintainer Maximilian Lange <maximilian-lange@t-online.de>

Description Provides some easy-to-use functions to interpolate species range based on species occurrences and to estimate centers of biodiversity.

Depends R (>= 3.6.0)

Imports sp, foreach, methods, grDevices, graphics, raster, stats,
utils

Suggests lattice, RColorBrewer

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2023-07-18 07:10:02 UTC

R topics documented:

`add.Edges` *Edge Creation*

Description

This function adds an edge between two points to a grid.

Usage

```
add.Edges(grid,point.a, point.b)
```

Arguments

`grid`	A grid to which the edge should be added.
`point.a`	A point in a grid given by its row and column.
`point.b`	A point in a grid given by its row and column.

Details

This function adds an edge between two points to a grid.

Value

The grid with new edge included.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##initialize data
m <- matrix(0,10,10)
point.a <- c(5,8)
point.b <- c(10,3)

##calculate the distance of the two points
m <- add.Edges(m,point.a,point.b)
image(m)
```

---

adjustment *Adjustment of interpolated species richness*

Description

This routine adjusts the weighted species richness for sampling effort.

Usage

```
adjustment(species.richness, noninterpolatedgrid, clusterlist)
```

Arguments

`species.richness`
The grid that should be adjusted. It contains the result of species richness estimation or cross-validation.

`noninterpolatedgrid`
A grid containing the species occurrences (could be created via function ’createNonInterpolatedGrid’).

`clusterlist` A list of vectors of pixelpositions (created by ’searchClusters’) which are spatial related.

Details

This routine adjusts the interpolated species richness for sampling effort by incorporating an additional weighting factor. Therefor, it needs clusters of pixels of the result grid to find centers of species richness and to create these weighting factor.

Value

This function returns a grid containing the adjusted species richness information.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.weighted <- species.richness(dataset.all.species,
landwatermask.nocoast, distances=1:5, weight=0.5, dimension, origin, resolution=1,
upperbound=5, all.species=1:2)

##prepare adjustment
noninterpolatedgrid <- createNonInterpolatedGrid(dataset.all.species,
dimension, origin, resolution=1, all.species=-1)
clusterlimit <- 100

clusterlist <- searchClusters(species.richness.weighted,
dimension, origin, resolution=1, clusterlimit)

##adjust species richness
species.richness.weighted.adjusted <- adjustment(species.richness.weighted,
noninterpolatedgrid, clusterlist)
```

---

cclFun *Connected Component Labeling*

Description

This routine identifies disjunct patches within a distribution matrix.

Usage

```
cclFun(mat)
```

Arguments

`mat`	Matrix with 0 as background value and positive values as occurrences to check patches for.

Details

This routine identifies disjunct patches of positive values within a matrix containing zeros as background values.

Value

This function returns a matrix with zero as background value and positive values numbering and representing the disjunct patches.

Author(s)

Maximilian Lange

Examples

```
## example data
tmat = { matrix(c( 0,0,0,1,0,0,1,1,0,1,
                   0,0,1,0,1,0,0,0,0,0,
                   0,1,NA,1,0,1,0,0,0,1,
                   1,0,1,1,1,0,1,0,0,1,
                   0,1,0,1,0,1,0,0,0,1,
                   0,0,1,0,1,0,0,1,1,0,
                   1,0,0,1,0,0,1,0,0,1,
                   0,1,0,0,0,1,0,0,0,1,
                   0,0,1,1,1,0,0,0,0,1,
                   1,1,1,0,0,0,0,0,0,1),
nr=10,byrow=TRUE) }

ccl.mat <- cclFun(tmat)
```

`createHeightmask` *Height-Mask Creation*

Description

This function creates a height-mask as a grid based on given dimension and coordinates.

Usage

```
createHeightmask(dataset.height, dimension, origin, resolution=1)
```

Arguments

`dataset.height`	A dataframe containing the longitude and lattitude of the grid cell location as 'long' and 'lat' and the height of the cell as 'height'. If the value of 'dataset.height' is 'NULL', the returned grid contains 0 as height-information of every cell.
`dimension`	The dimension of the processed grid.
`origin`	The geographic coordinates of the origin of the grid.
`resolution`	The resolution of the grid in (geographical) degree.

Details

This routine creates a height-mask as a grid. As Input a dataset with height-information is needed. The grid will be created with a given dimension.

Value

This function returns a grid which contains the height information of the observed area.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)
data(dataset.height)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create height-matrix
height.matrix <- createHeightmask(dataset.height, dimension, origin, resolution=1)
```

`createImage` *Image Creation*

Description

This function creates a levelplot of the result-grid of species range estimation and saves it as a PNG-File.

Usage

```
createImage(grid, landwatermask, image.title,
directory, filename, origin,
parts=10, resolution=1)
```

Arguments

`grid` The grid that should be plotted. It contains the weighted species richness information.

`landwatermask` A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0.

`image.title` The heading of the created image.

`directory` The directory in which the image should be stored.

`filename` The filename of the created PNG-File.

`origin` The geographic coordinates of the origin of the grid.

`parts` An Integer value determining the number of divisions in the colorkey of the created image.

`resolution` The resolution of the grid in (geographical) degree.

Details

This routine creates a levelplot of the weighted species richness information and stores it as a PNG-File. The package lattice is required.

Value

The return type is boolean. The value is TRUE if the image-creation was successful and FALSE if the image-creation has failed.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.weighted <- species.richness(dataset.all.species,
landwatermask.nocoast, distances=1:5, weight=0.5, dimension,
origin, resolution=1, upperbound=3000, all.species=1:2)

##create image
```

```
## Not run:
image.success <- createImage(species.richness.weighted, landwatermask.nocoast,
    image.title="Test", directory=getwd(), filename="species.richness.test01.png",
    origin, parts=10, resolution=1)
print(image.success)

## End(Not run)
```

createLandwatermask *Land-Water-Mask Creation*

Description

This function creates a land-water-mask as a grid based on given dimension and coordinates.

Usage

```
createLandwatermask(dataset.landwater, dimension, origin, resolution=1)
```

Arguments

`dataset.landwater`
A dataset containing the percentage of land on a cell of a grid. The coordinates of the cell should be given as longitude (named: long) and lattude (named: lat) and the percentage of land shoud be named landsum. If the value of 'dataset.landwater' is 'NULL', a grid containing only land-cells will be created.

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

Details

This routine creates a land-water-mask as a grid. As Input a dataset with land-water-information is needed. The grid will be created with a given dimension.

Value

This function returns a grid which contains the land-water information of the observed area. The value of cells containing land is 0, the value of water-cells is -1.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)
```

`createNonInterpolatedGrid`

Create a species occurrences grid with no interpolation

Description

This routine creates a grid containing the species occurrences.

Usage

```
createNonInterpolatedGrid(dataset.all.species, dimension,
origin, resolution=1, all.species=-1)
```

Arguments

`dataset.all.species`
A dataset containing all observed species with their ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

`all.species` The vector with the numbers of the species which should be mentioned. If the first value is -1, all species in the database will be added to the grid.

Details

This routine creates a grid and adds all species mentioned in 'all.species' to it.

Value

A grid containing the species occurrences without interpolation.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)
resolution <- 1
all.species <- -1

##create grid containing species occurrences
species.occurrences <- createNonInterpolatedGrid(dataset.all.species,
dimension, origin, resolution, all.species)
```

`data.into.Grid` *Species occurrences to grid*

Description

This function adds the species occurrence data to a new grid with given dimension, shift and resolution.

Usage

```
data.into.Grid(dataset.one.species, dimension, origin, resolution=1)
```

Arguments

`dataset.one.species`
A dataset containing one species with its ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of the occurrence locations of that species.

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

Details

This routine adds the species occurrence data to the grid through adding their location as a point into it.

Value

This function returns a grid which contains the species occurrence information.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

#extract datasets of one species out of database
dataset.one.species <- extract.species(dataset.all.species, 3)

##add species occurrence information to grid
grid <- data.into.Grid(dataset.one.species, dimension, origin, resolution=1)
```

`dataset.all.species` *Angiosperm Dataset*

Description

This data set gives an example of angiosperm species and their occurrence in southern america.

Format

A dataframe containing the longitude and lattitude of the occurrence location as 'long' and 'lat' and the species ID as 'speciesID'.

Source

Angiosperm Database

References

Morawetz, W. and Raedig, C. (2007). Angiosperm biodiversity, endemism and conservation in the Neotropics. Taxon 56:1245-1254.

`dataset.height` *Height-Information Dataset*

Description

This data set containes the height of 1 degree grid cells of southern america.

Format

A dataframe containing the longitude and lattitude of the grid cell location as ’long’ and ’lat’ and the height of the cell as ’height’.

`dataset.landwater` *Land-Water-Information Dataset*

Description

This data set containes the percentage of land/water of 1 degree grid cells of southern america.

Format

A dataframe containing the longitude and lattitude of the grid cell location as ’long’ and ’lat’ and the percentage of land/water as landsum.

`edgeNotValid` *Edge-Water-Test*

Description

This function checks if an edge would cross a water or mountain surface.

Usage

```
edgeNotValid(grid,point.a, point.b, landwatermask, upperbound)
```

Arguments

`grid` A grid to which the edge should be added if it hits no water or mountain surfaces.

`point.a` A point in a grid given by its row and column.

`point.b` A point in a grid given by its row and column.

`landwatermask` A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0. Additional, height-informations could be added for land surfaces. In this case, take care of the 'upperbound' value.

`upperbound` This value determines the height (based on values in 'landwatermask') which is considered to be a barrier for species distribution.

Details

This function checks if an edge would cross a water or mountain surface.

Value

A boolean value which determines if the edge would cross a water or mountain surface.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##initialize variables
distance <- 5

##create grid parameters
dimension <- getDimension(dataset.all.species)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater, dimension, origin)

##extract datasets of one species out of database
dataset.one.species <- extract.species(dataset.all.species, 3)

##create grid
grid <- matrix(0,dimension[1],dimension[2])

##add points
grid <- data.into.Grid(dataset.one.species, dimension, origin)

##points to list
points <- which(grid > 0)
```

```
points.xy <- list()
for (i in 1:length(points)){
points.xy[[i]] <- c(ifelse((points[i] %% dimension[1]) == 0, dimension[1],
points[i] %% dimension[1]), ceiling(points[i]/dimension[1]))
}

##Edge-Water-Test
check <- edgeNotValid(grid, points.xy[[1]],points.xy[[2]],
landwatermask.nocoast, 1000)
print(check)
```

evaluate *Result grid evaluation*

Description

This function evaluates the result grids of species richness estimation.

Usage

```
evaluate(result.grid.one, result.grid.two,
title.one="Histogramm of species richness",
title.two="Histogramm of species richness", xmax=400, ymax=1000,
directory=getwd(), filename="histogramm.png")
```

Arguments

`result.grid.one`
A result grid of species richness estimation. If the value is 'NULL', the grid is ignored and only the other grid will be included in the resulting png-file.

`result.grid.two`
A result grid of species richness estimation. If the value is 'NULL', the grid is ignored and only the other grid will be included in the resulting png-file.

`title.one` The title for the histogramm of the first grid ('result.grid.one').

`title.two` The title for the histogramm of the second grid ('result.grid.two').

`xmax` The maximum value of the abscissa respectively the maximum number of species in the grid.

`ymax` The maximum value of the ordinate respectively the maximum frequency shown in the histogramm.

`directory` The directory the png-file is created in.

`filename` The name of the created png-file.

Details

This routine evaluates the result grids of a species richness estimation (details in Raedig et al. 2010). It creates one or two histogramms (depending on input) of the species occurrences with additional information about the total number of species of the grid and other statistic scores. It is possible to create on png-file for every result grid or to compare two result-grids in one png-file, depending on how many grids are defined as input.

Value

This function creates a png-file with one or two histogramms (depending on input) and additional statistic informations.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
#load data
data(dataset.all.species)
data(dataset.landwater)

#create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

#create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

#estimate species richness
species.richness.weighted <- species.richness(dataset.all.species,
landwatermask.nocoast, distances=1:5, weight=0.5, dimension,
origin, resolution=1, upperbound=3000, all.species=1:2)

#evaluation
## Not run:
evaluate(species.richness.weighted, NULL, title.one="Histogramm 1")

## End(Not run)
```

export_tif *Export result grid as GeoTiff*

Description

This function exports the result grid as a GeoTiff.

Usage

```
export_tif(grid, origin, resolution, epsg_code, directory=getwd(),
filename="grid.tif")
```

Arguments

grid The grid that should be exported.

origin The geographic coordinates of the origin of the grid.

resolution The resolution of the grid in (geographical) degree.

epsg_code The epsg_code of the origin coordines. Defauls to 4326.

directory The directory in which the file will be created.

filename The name of the created GeoTiff-file.

Details

This routine exports a grid as a GeoTiff-file.

Author(s)

Maximilian Lange

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.weighted <- species.richness(dataset.all.species,
landwatermask.nocoast, distances=1:5, weight=0.5, dimension,
origin, resolution=1, upperbound=3000, all.species=1:2)

##export
## Not run:
export_tif(species.richness.weighted, origin, resolution=1, epsg_code="4326",
    directory=getwd(), filename="species.richness.tif")

## End(Not run)
```

`extract.species` *Species Extraktion*

Description

This function extracts the datasets of one species out of a database filled with species and their occurrence locations.

Usage

```
extract.species(dataset.all.species, species.number)
```

Arguments

`dataset.all.species`
A dataset containing all observed species with their ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`species.number` The number of the species which should be extracted.

Details

This function extracts the datasets of one species out of a database filled with species (named by their speciesID) and their occurrence locations mentioned with lattitude (named lat) and longitude (named long).

Value

This function returns a dataset which contains the species occurrence locations.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)

#extract datasets of one species (with speciesID=3) out of database
dataset.one.species <- extract.species(dataset.all.species, 3)
```

fill.Areas *Fill Algorithm*

Description

This function fills all bounded areas in a grid.

Usage

```
fill.Areas(grid, landwatermask)
```

Arguments

`grid` A grid in which the areas should be filled. Cells with values greater than zero represent the bounds, cells with values lower or equal zero will be filled with value 1 if they are bounded.

`landwatermask` A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0. Cells with negative values will not be filled.

Details

This function fills all bounded areas in a grid.

Value

The grid with filled areas.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##initialize data
m <- matrix(0,10,10)
m[2:8,2] <- 1
m[2:8,8] <- 1
m[2,2:8] <- 1
m[8,2:8] <- 1

##initialize mask (in this case: empty mask)
mask <- matrix(0,10,10)

##calculate the distance of the two points
m <- fill.Areas(m, mask)
image(m)
```

`generate.subsamples` *Subsample Creation*

Description

A function that creates an array of subsamples of a dataset.

Usage

```
generate.subsamples(number.of.occurrences, fold, loocv.limit)
```

Arguments

`number.of.occurrences`
The number of occurrences of the species.

`fold` The number of groups which should be created if the number of occurrences is greater than loocv.limit.

`loocv.limit` The limit below which the subsamples are created for a leave-one-out-cross-validation instead of a k-fold-cross-validation.

Details

This routine creates subsamples of species occurrences based on the number of occurrences of the chosen species and the number of groups which should be created. If the number of occurrences of the chosen species is below the loocv.limit, the groups contain only one member and the fold value will be ignored. The subsamples can be used for a k-fold or leave-one-out-cross-validation.

Value

This function returns a grid which contains the subsamples.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
subsamples <- generate.subsamples(15,3,5)
```

getDimension *getDimension*

Description

This function calculates the dimension of the grid which should be used for a species richness estimation.

Usage

```
getDimension(dataset.all.species, resolution=1)
```

Arguments

`dataset.all.species`
A dataset containing all observed species with their ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`resolution` The resolution of the grid in (geographical) degree.

Details

This function calculates the dimension of the grid which should be used for a species richness estimation on a given database of species occurrence locations.

Value

The dimension which is needed to create a grid used for a species richness estimation on a given database of species occurrence locations.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)

#extract dimension out of the database
dimension <- getDimension(dataset.all.species, resolution=1)
print(dimension)
```

getDistance *Distance between two points in a grid*

Description

This function calculates the distance of two points in a grid.

Usage

```
getDistance(point.a, point.b, resolution=1)
```

Arguments

`point.a`	A point in a grid given by its row and column.
`point.b`	A point in a grid given by its row and column.
`resolution`	The resolution of the grid in (geographical) degree.

Details

This function calculates the distance of two points in a grid by the Pythagorean theorem.

Value

The distance of the two points as an numeric value based on geographical degrees.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##initialize data
point.a <- c(5,8)
point.b <- c(10,3)

##calculate the distance of the two points
distance <- getDistance(point.a,point.b, resolution=1)
print(distance)
```

getNarrowEndemics *Narrow endemic species extraction*

Description

This function extracts the row number of narrow endemic species in the given dataset.

Usage

```
getNarrowEndemics(dataset.all.species, all.species,
narrow.endemic.limit, dimension, origin,
resolution)
```

Arguments

`dataset.all.species`
: A dataset containing all observed species with their ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`all.species`
: A vector with the numbers of the species which should be mentioned. If the first value is -1, all species in the database will be used.

`narrow.endemic.limit`
: This value determines the limit of points and the maximum distance of these points up to which a species is considered as narrow endemic species.

`dimension`
: The dimension of the grid which should be processed.

`origin`
: The geographic coordinates of the origin of the grid which should be processed.

`resolution`
: The resolution of the grid which should be processed in (geographical) degree.

Details

This function extracts the row numbers of narrow endemic species in the given datasets.

Value

This function returns a vector containing the row numbers of narrow endemic species in the given dataset.

Author(s)

Maximilian Lange, Claudia Raedig

Examples

```
## load data
data(dataset.all.species)

## create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

## search narrow endemic species in row 100 to 200 in the dataset
narrow.endemic.species <- getNarrowEndemics(dataset.all.species,
all.species=100:200, narrow.endemic.limit=8,
dimension, origin, resolution=1)
```

getOrigin *Origin Koordinates*

Description

This function calculates the origin of the grid used for a species richness estimation.

Usage

```
getOrigin(dataset.all.species)
```

Arguments

`dataset.all.species`
A dataset containing all observed species with their ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of their occurrence location.

Details

This function calculates the coordinates of the origin of the grid which should be used for a species richness estimation.

Value

The coordinates of the origin of the grid used for species richness estimation.

Author(s)

Maximilian Lange, Sven Lautenbach

Examples

```
##load data
data(dataset.all.species)

#extract coordinates of origin out of database
origin <- getOrigin(dataset.all.species)
print(origin)
```

searchClusters *Search clusters in the result grid of species richness estimation.*

Description

This routine searches clusters in the result grid of species richness estimation.

Usage

```
searchClusters(species.richness, dimension, origin, resolution, clusterlimit)
```

Arguments

`species.richness`
: The grid that should be adjusted. It contains the result of species richness estimation or cross-validation.

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

`clusterlimit` The limit under which values of 'species.richness' should not be used to build clusters.

Details

This routine searches clusters in the result grid of species richness estimation to find centers of species richness.

Value

This function returns a list containing vectors of the position of pixels which belong together.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.weighted <- species.richness(dataset.all.species,
landwatermask.nocoast, distances=1:5, weight=0.5, dimension,
origin, resolution=1, upperbound=3000, all.species=1:2)

##prepare clustersearch
clusterlimit <- 100

##adjust species richness
clusterlist <- searchClusters(species.richness.weighted,
dimension, origin, resolution=1, clusterlimit)
```

species.range *Species range estimation*

Description

This function estimates the species range based on given occurrences of one species.

Usage

```
species.range(dataset.one.species, distance, dimension, origin,
resolution=1, landwatermask, upperbound, cross.validation=FALSE)
```

Arguments

`dataset.one.species`
A dataset containing one species with its ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of the occurrence locations of that species.

`landwatermask` A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0. Additional, height-informations could be added for land surfaces. In this case, take care of the 'upperbound' value.

`distance` The maximum distance in which two occurrences are considered as related occurrences. These two occurrences will be connected with an edge.

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

`upperbound` This value determines the height (based on values in 'landwatermask') which is considered to be a barrier for species distribution.

`cross.validation`
A logical value determining wether a cross-validation is performed.

Details

This routine estimates the species range based on given occurrences of one species through a geometric interpolation model (details in Raedig et al. 2010).

Value

This function returns a grid which contains the species range information.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##extract datasets of one species out of database
dataset.one.species <- extract.species(dataset.all.species, 3)

##estimate species range
species.range.tmp <- species.range(dataset.one.species, distance=5,
dimension, origin, resolution=1, landwatermask.nocoast)
```

species.richness *Species richness estimation*

Description

This function estimates the species richness based on given species occurrences.

Usage

```
species.richness(dataset.all.species, landwatermask,
distances=2:10, weight=0.5, dimension, origin, resolution=1,
upperbound, all.species=-1, silent=TRUE, do.parallel=FALSE)
```

Arguments

`dataset.all.species`	A dataset containing the species with their ID (named: speciesID)and the longitude (named: long) and latitude (named: lat) of their occurrence location.
`landwatermask`	A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0. Additional, height-informations could be added for land surfaces. In this case, take care of the 'upperbound' value.
`distances`	The distances which will be used for species range estimation.
`weight`	The tuning parameter of the weighting procedure (details in Raedig et al. 2010).
`dimension`	The dimension of the processed grid.
`origin`	The geographic coordinates of the origin of the grid.
`resolution`	The resolution of the grid in (geographical) degree.
`upperbound`	This value determines the height (based on values in 'landwatermask') which is considered to be a barrier for species distribution.
`all.species`	The vector with the numbers of the species which should be mentioned. If the first value is -1, all species in the database will be used for species richness estimation.
`silent`	A boolean flag that determines wether the report of status messages should be suppressed or not.
`do.parallel`	A boolean flag determining wether the function uses the 'foreach'-package to estimate species richness via parallel processing.

Details

This routine estimates the species richness based on given species occurrences through a geometric interpolation model (details in Raedig et al. 2010).

Value

This function returns a grid which contains the weighted species richness information.

Author(s)

Maximilian Lange, Sven Lautenbach, Claudia Raedig

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.weighted <- species.richness(dataset.all.species,
landwatermask.nocoast, distances=1:5, weight=0.5, dimension,
origin, resolution=1, upperbound=5, all.species=1:2, silent=TRUE)
```

species.richness.cv *Species richness cross-validation*

Description

This function cross-validates the species richness based on given species occurrences.

Usage

```
species.richness.cv(dataset.all.species, landwatermask, fold=5,
loocv.limit=10, distances=3:10, weight=0.5, dimension, origin,
resolution=1, upperbound, all.species=-1,
silent=TRUE, do.parallel=FALSE)
```

Arguments

`dataset.all.species`
: A dataset containing the species with their ID (named: speciesID)and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`landwatermask` A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0. Additional, height-informations could be added for land surfaces. In this case, take care of the 'upperbound' value.

`fold` The number of groups which should be created if the number of occurrences is greater than loocv.limit.

`loocv.limit` The limit below which the subsamples are created for a leave-one-out-cross-validation instead of a k-fold-cross-validation.

`distances` The distances which will be used for species range estimation.

`weight` The tuning parameter of the weighting procedure (details in Raedig et al. 2010).

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

`upperbound` This value determines the height (based on values in 'landwatermask') which is considered to be a barrier for species distribution.

`all.species` The vector with the numbers of the species which should be mentioned. If the first value is -1, all species in the database will be used for species richness estimation.

`silent` A boolean flag that determines wether the report of status messages should be suppressed or not.

`do.parallel` A boolean flag determining wether the function uses the 'foreach'-package to cross-validate via parallel processing.

Details

This routine cross-validates the species richness based on given species occurrences through a geometric interpolation model (details in Raedig et al. 2010).

Value

This function returns a grid which contains the cross-validated species richness information.

Author(s)

Maximilian Lange, Sven Lautenbach, Claudia Raedig

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.weighted <- species.richness.cv(dataset.all.species,
landwatermask.nocoast, fold=5, loocv.limit=10, distances=2:5,
weight=0.5, dimension, origin, resolution=1, upperbound=5,
all.species=1:2)
```

species.richness.main *Main Function for species richness estimation*

Description

The main function to estimate or cross-validate the species richness based on given species occurrences.

Usage

```
species.richness.main(dataset.all.species, dataset.landwater,
dataset.height, distances=1:10, weight=0.5, resolution=1,
narrow.endemic=FALSE, narrow.endemic.limit=5, upperbound=5,
cross.validation=FALSE, fold=5, loocv.limit=10,
create.image=FALSE, image.title="Interpolated Species Richness",
directory=getwd(), filename="species.richness.png",
evaluation=FALSE, eval.title="Histogramm", adjust=FALSE,
clusterlimit=100, predefinedClusterlist=NULL, all.species=-1,
export=FALSE, exportname="species.richness.tif",
noninterpolatedgrid=NULL, silent=TRUE, do.parallel=FALSE)
```

Arguments

`dataset.all.species`
: A dataset containing all observed species with their ID (named: speciesID) and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`dataset.landwater` A dataset containing the percentage of land on a cell of a grid. The coordinates of the cell should be given as longitude (named: long) and lattude (named: lat) and the percentage of land shoud be named landsum. If the value of 'dataset.landwater' is 'NULL', the land-water-information has no effect on the species richness estimation.

`dataset.height` A dataset containing the longitude and lattitude of the grid cell location as 'long' and 'lat' and the height of the cell as 'height'. If the value of 'dataset.height' is 'NULL', the height-information has no effect on the species richness estimation.

`distances` The distances used for species range estimation or cross validation.

`weight` The tuning parameter of the weighting procedure (details in Raedig et al. 2010).

`resolution` The resolution of the grid in (geographical) degree.

`narrow.endemic` A boolean flag that determines if only narrow endemic species should be considered in species richness estimation.

`narrow.endemic.limit` This value determines the limit of points and the maximum distance of these points up to which a species is considered as narrow endemic species.

`upperbound` This value determines the height which is considered to be a barrier for species distribution.

`cross.validation` A logical value determining wether a cross-validation is performed. If the value is true, the parameters narrow.endemic and narrow.endemic.limit will be ignored.

`fold` The number of groups which should be created if the number of occurrences is greater than loocv.limit.

`loocv.limit` The limit below which the subsamples are created for a leave-one-out-cross-validation instead of a k-fold-cross-validation.

`create.image` A boolean flag that determines if an image (PNG-File) is created.

`image.title` The heading of the created image.

`directory` The directory in which the created files should be stored.

`filename` The filename of the created PNG-Files.

`evaluation` A boolean value determining wether the routine 'evaluate' is used or not. If the value is true, a PNG-File with a histogramm of the result grid will be created.

`eval.title` The heading of the created histogramm.

`adjust` A boolean value determining wether an adjustment of the result grid should be done or not.

`clusterlimit` The limit below values of 'species.richness' should not be used to build clusters. Each cluster will be adjusted with an own centre of species richness.

`predefinedClusterlist` A list of vectors of pixelpositions (created by 'searchClusters') which are spatial related. The default value is NULL because the list will be created if 'adjust' is 'TRUE', but it may be useful for robustness estimation, because the cross-validation result should be adjusted with the same clusterlist as the related species richness estimation (and without a pre-defined clusterlist, other clusters will be build).

`all.species` A vector with the identification numbers of the considered species. If the first value is -1, all species in the database will be used.

`export` A boolean value that determines if the routine should export the results as GeoTiff.

`exportname` The name of the created GeoTiff-file.

`noninterpolatedgrid`
A grid containing the species occurrences (could be created via function 'createNonInterpolatedGrid'). If the value of this parameter is 'NULL', it will be created if an adjustment should be done. The parameter may be usefull to save time while processing more than one 'species.richness.main'.

`silent` A boolean flag that determines wether the report of status messages should be suppressed or not.

`do.parallel` A boolean flag determining wether the function uses the 'foreach'-package to process in parallel.

Details

This routine is the main function of this package. It either estimates or cross-validates the species richness based on given species occurrences using a geometric interpolation model (details in Raedig et al. 2010).

Value

This function returns a grid which contains either the weighted or the cross-validated species richness information. Additionally, an image (PNG-File) of that grid can be created.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)
data(dataset.height)

##estimate species richness
species.richness.weighted <- species.richness.main(dataset.all.species,
    dataset.landwater, dataset.height, distances=1:5, weight=0.5,
    resolution=1, narrow.endemic=FALSE, narrow.endemic.limit=5,
    upperbound=5, cross.validation=FALSE, fold=5, loocv.limit=10,
    create.image=FALSE, directory=getwd(), filename="species.richness.png",
    all.species=1:2, export=FALSE)
```

species.richness.nonweighted

Species richness estimation without weighting

Description

This function estimates the species richness for a given distance based on given species occurrences without using weighting factor.

Usage

```
species.richness.nonweighted(dataset.all.species, landwatermask,
distance=10, dimension, origin, resolution=1,
upperbound, all.species=-1, silent=TRUE,
do.parallel=FALSE)
```

Arguments

`dataset.all.species`
A dataset containing the species with their ID (named: speciesID)and the longitude (named: long) and latitude (named: lat) of their occurrence location.

`landwatermask` A grid containing the land-water-information of the observed area. If a grid cell containes no land, the value of the cell in the landwatermask is -1, otherwise it is 0. Additional, height-informations could be added for land surfaces. In this case, take care of the 'upperbound' value.

`distance` The distance which will be used for species range estimation.

`dimension` The dimension of the processed grid.

`origin` The geographic coordinates of the origin of the grid.

`resolution` The resolution of the grid in (geographical) degree.

`upperbound` This value determines the height (based on values in 'landwatermask') which is considered to be a barrier for species distribution.

`all.species` The vector with the numbers of the species which should be mentioned. If the first value is -1, all species in the database will be used for species richness estimation.

`silent` A boolean flag that determines wether the report of status messages should be suppressed or not.

`do.parallel` A boolean flag determining wether the function uses the 'foreach'-package to estimate species richness via parallel processing.

Details

This routine estimates the species richness for a given distance (without weighted summation) based on given species occurrences through a geometric interpolation model (details in Raedig et al. 2010).

Value

This function returns a grid which contains the species richness information for a given distance.

Author(s)

Maximilian Lange, Sven Lautenbach

References

Raedig, C., Dorman, C.F., Hildebrandt, A. and Lautenbach, S. (2010). Reassessing Neotropical angiosperm distribution patterns based on monographic data: a geometric interpolation approach. Biodivers Conserv, 19, 1523-1546.

Examples

```
##load data
data(dataset.all.species)
data(dataset.landwater)

##create grid parameters
dimension <- getDimension(dataset.all.species, resolution=1)
origin <- getOrigin(dataset.all.species)

##create landwatermask
landwatermask.nocoast <- createLandwatermask(dataset.landwater,
dimension, origin, resolution=1)

##estimate species richness
species.richness.noweight <- species.richness.nonweighted(dataset.all.species,
landwatermask.nocoast, distance=5, dimension,
origin, resolution=1, upperbound=5, all.species=1:2)
```

Index