

Package ‘mopa’

June 6, 2017

Depends R (>= 3.4.0), raster, sp

Imports splancs, spam, spatstat, sampling, dismo, rpart, tree, e1071,
earth, randomForest, PresenceAbsence, lattice, gtools, ranger,
abind,

Suggests rJava (>= 0.5-0), caret (>= 6.0-76)

Type Package

Title Species Distribution MOdeling with Pseudo-Absences

Version 1.0.0

Date 2017-06-05

Description Tools for transferable species distribution modeling and pseudo-absence data generation allowing the straightforward design of relatively complex experiments with multiple factors affecting the uncertainty (variability) of SDM outputs (pseudo-absence sample, climate projection, modeling algorithm, etc.), and the quantification of the contribution of different factors to the final variability following the method described in Deque et al. (2010) <doi:10.1007/s00382-011-1053-x>. Multiple methods for pseudo-absence data generation can be applied, including the novel Three-step method as described in Iturbide et al. (2015) <doi:10.1016/j.ecolmodel.2015.05.018>. Additionally, a function for niche overlap calculation is provided, considering the metrics described in Warren et al. (2008) <10.1111/j.1558-5646.2008.00482.x> and in Pianka (1973) <10.1146/annurev.es.04.110173.000413>.

BugReports <https://github.com/SantanderMetGroup/mopa/issues>

URL <https://github.com/SantanderMetGroup/mopa/wiki>

License GPL-3

LazyLoad true

RoxygenNote 6.0.1

Maintainer Maialen Iturbide <miturbide@ifca.unican.es>

NeedsCompilation no

Author Maialen Iturbide [aut, cre],
Joaquin Bedia [ctb],
SantanderMetGroup <http://meteo.unican.es> [ctb]

Repository CRAN

Date/Publication 2017-06-06 15:40:42 UTC

R topics documented:

backgroundGrid	2
backgroundRadius	3
extractFromModel	5
extractFromPrediction	7
mods	8
mopaPredict	8
mopaTrain	10
nicheOver	12
Oak_phylo2	14
OCSVMprofiling	14
pseudoAbsences	16
Q_pubescens	18
varianceAnalysis	19
wrld	20

Index	22
--------------	-----------

backgroundGrid	<i>Create background coordinates from raster object</i>
----------------	---

Description

Creates the background coordinates used for defining the study area and for generating pseudo-absences.

Usage

```
backgroundGrid(raster, spatial.subset = NULL)
```

Arguments

raster	Raster object with projection (<code>?raster::projection</code> and <code>?crs</code>) from which to extract the the point grid (used as background coordinates for generating pseudo-absences)
spatial.subset	Object of class <code>extent</code> (see extent) or a two column data.frame (or matrix) of coordinates (xy, each row is a point).

Details

If a data.frame, matrix or a list of the previous is passed to `spatial.subset` the bounding coordinates are extracted to delimit the background. For example, to bound the study area to the spatial distribution of a species.

Value

A list with a `SpatialPolygons` object and a matrix of the background coordinates

Author(s)

M. Iturbide

References

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutierrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*. DOI:10.1016/j.ecolmodel.2015.05.018.

Examples

```
## Load presences
data(Oak_phylo2)

## Load rasters
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

projection(biostack$baseline) <- CRS("+proj=longlat +init=epsg:4326")
r <- biostack$baseline[[1]]

## Background around a set of coordinates
bg.species <- backgroundGrid(r, Oak_phylo2)
## Background of a subdomain of the study area
bg.subdomain <- backgroundGrid(r, extent(c(-10, 30, 35, 65)))
## Background of the whole study area
bg <- backgroundGrid(r)

# plot(bg$xy)
# plot(bg.subdomain$xy)
# plot(bg.species$xy$H11)
plot(bg.species$xy$H01)
```

backgroundRadius

*Background extent restriction for a sequence of distances***Description**

Creation of point-grid backgrounds through the establishment of extent limitations for a sequence of distances, from near presence locations to the length of the half diagonal of the bounding that encloses the background (study area).

Usage

```
backgroundRadius(xy, background, start = 0.166, by = 0.083,
  unit = c("decimal degrees", "utm"))
```

Arguments

<code>xy</code>	Data frame or list of data frames with coordinates (each row is a point) — typically species presence data— to be considered as starting points from which different background extents are created.
<code>background</code>	Matrix or list of matrices of background coordinates. Object derived from function <code>OCSVMprofiling</code> (component <code>\$absence</code>). Alternatively, object derived from function <code>backgroundGrid</code> (component <code>\$xy</code>) if the environmental profiling step is going to be avoided in the pseudo-absence generation process).
<code>start</code>	Value for the minimum distance to consider for extent limitations. Default is 0.166
<code>by</code>	Value of the distance to consider from one extent to the following. Default is 0.083
<code>unit</code>	Character indicating the coordinate system of the objects. Default is "decimal degrees", alternatively "utm" can be used

Details

Argument `unit` is only used to set extent distances in km. This function is aimed at creating backgrounds of different extent for pseudo-absence sampling based on an initial point grid (derived from function `OCSVMprofiling` or function `backgroundGrid`). If this function is used for a subsequent application of functions `pseudoAbsences` and `mopaTrain`, the last will perform species distribution modeling for each of the extents here established, and will return the fitted model that belongs to the optimum background extent (see references).

Value

List/s of matrixes with `xy` coordinates, each matrix correspond to a different background extent.

Author(s)

M. Iturbide

References

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutierrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*. DOI:10.1016/j.ecolmodel.2015.05.018.

See Also

[mopaTrain](#), [pseudoAbsences](#), [backgroundGrid](#), [OCSVMprofiling](#)

Examples

```
## Considering a single group of presence points
data(Q_pubescens)
presences <- Q_pubescens[sample(1:300, size = 100),]
```

```

# Define the spatial characteristics of the study area
r <- raster(nrows=50, ncols=50, xmin=-10, xmax=20, ymn=35, ymx=65, vals = rep(1, 50*50))

# Background of the whole study area
bg <- backgroundGrid(r)

# Partition of the study area
bg.extents <- backgroundRadius(xy = presences, background = bg$xy,
                              start = 0.166, by = 0.083*50, unit = "decimal degrees")

## Considering more than one groups of presence points
data(Oak_phylo2)

# Obtaining the raster that defines the spatial characteristics of the study area
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

projection(biostack$baseline) <- CRS("+proj=longlat +init=epsg:4326")
r <- biostack$baseline[[1]]
# Background of the whole study area
bg <- backgroundGrid(r)

# Partition of the study area
bg.extents <- backgroundRadius(xy = Oak_phylo2, background = bg$xy,
                              start = 0.166, by = 0.083*10, unit = "decimal degrees")

```

extractFromModel *Extract objects from lists returned by function [mopaTrain](#)*

Description

Extract values returned by function [mopaTrain](#)

Usage

```
extractFromModel(models, value = c("model", "auc", "kappa", "tss",
                                   "fold.models", "ObsPred"))
```

Arguments

models	Object returned by mopaTrain .
value	Any character of the following: "model", "auc", "kappa", "tss", "fold.models", "ObsPred"

Details

ObsPred allows to calculate further accuracy measures.

Value

Depending on the specified value:

"model"	fitted model using all data for training
"auc"	AUC statistic in the cross validation
"kappa"	kappa statistic in the cross validation
"tss"	true skill statistic in the cross validation
"fold.models"	fitted model with partitioned data
"ObsPred"	Observed and predicted (cross model prediction) values

Author(s)

M. Iturbide

Examples

```
## Load presence data
data(Oak_phylo2)

## Load Climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Spatial reference
r <- biostack$baseline[[1]]

## Create background grid
bg <- backgroundGrid(r)
## Generate pseudo-absences
RS_random <- pseudoAbsences(xy = Oak_phylo2, background = bg$xy,
                             exclusion.buffer = 0.083*5, prevalence = -0.5, kmeans = FALSE)

## Model training
fittedRS <- mopaTrain(y = RS_random, x = biostack$baseline,
                     k = 10, algorithm = "glm", weighting = TRUE)

## Extract fitted models
mods <- extractFromModel(models = fittedRS, value = "model")
## Extract observed and predicted values
ObsPred <- extractFromModel(models = fittedRS, value = "ObsPred")
```

extractFromPrediction *Extract values from objects or list of objects*

Description

Extract objects from lists returned by function `mopaPredict`.

Usage

```
extractFromPrediction(predictions, value)
```

Arguments

predictions	Listed lists of objects (e.g. as returned by <code>mopaPredict</code>)
value	Character pointing to the name of component/s in the list

Author(s)

M. Iturbide

Examples

```
## Load and prepare presence data
data(Q_pubescens)
presences <- Q_pubescens[sample(1:300, size = 100),]

## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Create the background of the whole study area
bg <- backgroundGrid(biostack$baseline$bio1)

## Generate pseudo-absences
RS_random <- pseudoAbsences(xy = presences, background = bg$xy,
                            exclusion.buffer = 0.083*5, prevalence = -0.5, kmeans = FALSE)

## Model training
fittedRS <- mopaTrain(y = RS_random, x = biostack$baseline,
                    k = 10, algorithm = "glm", weighting = TRUE)

## Extract fitted models
mods <- extractFromModel(models = fittedRS, value = "model")

## Model prediction
preds <- mopaPredict(models = mods, newClim = biostack$future)
predsMPI <- extractFromPrediction(predictions = preds, value = "MPI")
splot(predsMPI)
```

 mods

Fitted models

Description

List of fitted models as returned by functions [mopaTrain](#) and [extractFromModel](#).

Source

```
# RS_random is the result of running the following code:
data(Oak_phylo2)
presences <- Oak_phylo2$H11
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)
r <- biostack$baseline[[1]]
## Background of the whole study area bg <- backgroundGrid(r)
## Considering an unique background extent
#' RS_random <- pseudoAbsences(xy = presences, background = bg$xy, realizations = 5, exclu-
sion.buffer = 0.083*5, prevalence = -0.5, kmeans = FALSE)
fittedModels <- mopaTrain(y = RS_random, x = biostack$baseline, k = 10, algorithm = "mars")
mods <- extractFromModel(models = fittedModels, value = "model")
```

 mopaPredict

Model prediction

Description

Model projection into a RasterStack

Usage

```
mopaPredict(models, newClim)
```

Arguments

models	Model class object (e.g. "glm") or list of model class objects, e.g. as returned by function extractFromModel .
newClim	RasterStack or list of RasterStack objects with variables for projecting

Value

RasterStack of the projected probabilities

Author(s)

M. Iturbide

References

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutierrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*. DOI:10.1016/j.ecolmodel.2015.05.018.

See Also

[mopaTrain](#), [extractFromPrediction](#)

Examples

```
# SHORT EXAMPLE
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Fitted models
data(mods)
?mods

## Model prediction
newClim <- lapply(1:4, function(x){
  crop(biostack$future[[x]], extent(-10, 10, 35, 65))
})

prdRS.fut <- mopaPredict(models = mods, newClim = newClim)

# FULL WORKED EXAMPLE
## Load presence data
data(Oak_phylo2)

## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Spatial reference
r <- biostack$baseline[[1]]
## Create background grid
bg <- backgroundGrid(r)
```

```

## Generate pseudo-absences
RS_random <-pseudoAbsences(xy = Oak_phylo2, background = bg$xy,
                           exclusion.buffer = 0.083*5, prevalence = -0.5, kmeans = FALSE)

## Model training
fittedRS <- mopaTrain(y = RS_random, x = biostack$baseline,
                    k = 10, algorithm = "glm", weighting = TRUE)

## Extract fitted models
mods <- extractFromModel(models = fittedRS, value = "model")

## Model prediction
preds <- mopaPredict(models = mods, newClim = biostack$future)

```

mopaTrain

Easy species distribution modeling and cross validation

Description

Species distribution modeling and k-fold cross validation for a set of presence/absence data per species, also considering different background extents (optional). Algorithms supported are "glm", "svm", "maxent", "mars", "rf", "cart.rpart" and "cart.tree"

Usage

```

mopaTrain(y, x, k = 10, algorithm = c("glm", "svm", "maxent", "mars", "rf",
  "cart.rpart", "cart.tree"), algorithm.args = NULL, weighting = FALSE,
  threshold = NULL, diagrams = FALSE, tuneRF.args = NULL)

```

Arguments

y	Object returned by function pseudoAbsences or data frame or list/s of data frames with coordinates in the first two columns and presence/absence (1=presence, 0=absence) in the third column.
x	RasterStack or list of RasterStacks of variables for modeling, a.k.a baseline environment/climatology
k	Integer. Number of folds for cross validation. Default is 10
algorithm	Character string of the algorithms for modeling. Options are the following: "glm", "svm", "maxent", "mars", "rf", "cart.rpart" and "cart.tree" (see details)
algorithm.args	Further arguments to be passed to the selected algorithm for modeling (functions involved are described in details).
weighting	Logical for model fitting with weighted presence/absences. Applicable for algorithms "glm", "mars", "rf", "cart.tree" and "cart.rpart". Default is FALSE. The processing time is considerably increased if weighting option is selected when the "mars" algorithm (see earth) is applied.
threshold	Cut value between 0 and 1 to calculate the confusion matrix. Default is NULL (see Details).

diagrams	Logical. Only applied if <code>x</code> contains data for different background extents (see backgroundRadius and pseudoAbsences). Should diagrams of AUC extent fitting be printed? default is FALSE.
tuneRF.args	list of arguments from function tuneRF . Only used when <code>algorithm = "rf"</code>

Details

This function calculates the AUC with the function [auc](#) from package **PresenceAbsence**. **Note:** Package **SDMTools** must be detached.

If `threshold` is not specified the value that maximizes the TSS (true skill statistic) is used to calculate the confusion matrix.

If `y` contains data for different background extents (see [backgroundRadius](#) and [pseudoAbsences](#)), [mopaTrain](#) performs the species distribution modeling for each different background extent, and fits obtained AUCs (corresponding to different background extents) to three non linear models (Michaelis-Menten, exponential2 and exponential3). The model that scores the lowest error is automatically selected to extract the V_m coefficient (equation 1 in Iturbide et al., 2015). Then, the minimum extent at which the AUC surpasses the V_m value is selected as the threshold extent (see Figure 3 in Iturbide et al., 2015), being the corresponding fitted SDM the one returned by [mopaFitting](#). If argument `diagrams` is set to TRUE, A fitted model plot (as in Fig. 3 in Iturbide et al., 2015) is printed in the plotting environment.

[mopaTrain](#) uses the algorithm implementations of the following functions and R packages:

- "mars" function [earth](#) from package **earth**
- "rf" function [ranger](#) from package **ranger**
- "maxent" function [maxent](#) from package **dismo**
- "cart.rpart" function [rpart](#) from package **rpart**
- "svm" function [best.svm](#) from package **e1071**
- "cart.tree" function [tree](#) from package **tree**
- "glm" function [glm](#) from package **stats**

For example, when applying "glm", further arguments from function [glm](#) can be passed to [mopaTrain](#) by using `algorithm.args`.

Value

A list of six components is returned for each species in `x`:

- `$model` fitted model using all data for training
- `$auc` AUC statistic in the cross validation
- `$kappa` kappa statistic in the cross validation
- `$tss` true skill statistic in the cross validation
- `$fold.models` fitted models of each data partition for cross validation
- `$ObsPred` cross model prediction (e.g. for further assessment of model accuracy)

Author(s)

M. Iturbide

References

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutierrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*. DOI:10.1016/j.ecolmodel.2015.05.018.

See Also

[mopaPredict](#), [pseudoAbsences](#), [backgroundGrid](#), [OCSVMprofiling](#), [backgroundRadius](#), [extractFromModel](#)

Examples

```
## Load presence data
data(Oak_phylo2)

## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Spatial reference
r <- biostack$baseline[[1]]
## Create background grid
bg <- backgroundGrid(r)

## Generate pseudo-absences
RS_random <- pseudoAbsences(xy = Oak_phylo2, background = bg$xy,
                             exclusion.buffer = 0.083*5, prevalence = -0.5, kmeans = FALSE)

## Model training
fittedRS <- mopaTrain(y = RS_random, x = biostack$baseline,
                     k = 10, algorithm = "glm", weighting = TRUE)

## Extract fitted models
mods <- extractFromModel(models = fittedRS, value = "model")
```

nicheOver

Niche overlap

Description

Compute niche overlap among rasters in a RasterStack

Usage

```
nicheOver(stack, metric = c("D", "O"))
```

Arguments

stack RasterStack
 metric Metric for niche overlap. Options are D and O (see details).

Details

Niche overlap measures the similarity of the environmental ranges occupied by each constructed model via operating the difference between two vectors of probability distributions.

D: Schoeners statistic for niche overlap (Warren et al., 2008). O: Pianka index (Pianka, 1973).

Value

Matrix of overlap values for metric D or O.

Author(s)

M. Iturbide

References

Warren DL, Glor RE, Turelli M, Funk D (2008) Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, 62, 2868-2883. doi:10.1111/j.1558-5646.2008.00482.x.

Pianka ER (1973) The Structure of Lizard Communities. *Annual Review of Ecology and Systematics*, 4, 53-74. doi:10.1146/annurev.es.04.110173.000413.

Examples

```
## Load presence data
data(Oak_phylo2)

## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Spatial reference
r <- biostack$baseline[[1]]
## Create background grid
bg <- backgroundGrid(r)

## Generate pseudo-absences
RS_random <- pseudoAbsences(xy = Oak_phylo2, background = bg$xy,
                             exclusion.buffer = 0.083*5, prevalence = -0.5, kmeans = FALSE)

## Model training
fittedRS <- mopaTrain(y = RS_random, x = biostack$baseline,
                     k = 10, algorithm = "glm", weighting = TRUE)

## Extract fitted models
```

```

mods <- extractFromModel(models = fittedRS, value = "model")

## Model prediction
preds <- mopaPredict(models = mods, newClim = biostack$future)

## Extract predictions for species phylogeny H11
predsH11 <- extractFromPrediction(predictions = preds, value = "H11")

## Compute niche overlap
no <- nicheOver(predsH11, metric = "D")
library(lattice)
levelplot(no, col.regions = rev(terrain.colors(16)))

```

Oak_phylo2

Oak distribution

Description

A dataset consisting of a list with two data frames with xy coordinates corresponding to the distribution of two phylogenetic groups of oaks (H11 and H1)

Source

Oak_phylo2 is a modified subset of the **Quercus sp Europe Petit 2002** database (Petit et al., 2002b), which is available in the **Georeferenced Database of Genetic Diversity** or (GD)².

References

Petit, R. J. *et al* 2002. Chloroplast DNA variation in european white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management* 156 (1-3), 5-26.

Evolution of Trees and Forest Communities

Ten years of the EVOLTREE network

EvoTREE E-Lab - An information system for forest genetics

ISBN: 978-2-9519296-3-9

OCSVMprofiling

Environmental profiling with One-Classification Support Vector Machine

Description

Presence-only modeling and classification of coordinates predicted as presence and absence

Usage

```
OCSVMprofiling(xy, varstack, background = NULL, nu = 0.5)
```

Arguments

<code>xy</code>	Data frame or list of data frames with coordinates (each row is a point)
<code>varstack</code>	RasterStack of variables for modeling
<code>background</code>	Object derived from function backgroundGrid . If NULL (default), the background is extracted from varstack. Matrix or list of matrixes of the background xy coordinates in columns.
<code>nu</code>	Parameter needed for one-classification svm . Default is 0.5

Details

This function constitutes the first step from a three-step process to generate pseudo-absences, and is aimed at excluding the suitable areas for the species (xy records) from the background from which pseudo-absences are sampled.

Value

A list with two components:

<code>absence</code>	Matrix or list of matrixes with xy coordinates predicted as absence
<code>presence</code>	Matrix or list of matrixes with xy coordinates predicted as presence

Author(s)

M. Iturbide

References

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutierrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*. DOI:10.1016/j.ecolmodel.2015.05.018.

See Also

[svm](#), [backgroundGrid](#)

Examples

```
## Load presence data
data(Oak_phylo2)

## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)
```

```

## Spatial reference
projection(biostack$baseline) <- CRS("+proj=longlat +init=epsg:4326")
r <- biostack$baseline[[1]]
## Background of the whole study area
bg <- backgroundGrid(r)

## Environmental profiling
bg.profiled <- OCSVMprofiling(xy = Oak_phylo2, varstack = biostack$baseline,
                             background = bg$xy)

## Plot
plot(bg.profiled$absence$H11, pch="*")
points(bg.profiled$presence$H11, pch="*", col= "pink")

```

pseudoAbsences

Pseudo-absence data generation

Description

Pseudo-absence data generation at random or by k-means clustering inside a single background or a group of backgrounds (e.g. of different extent, [backgroundRadius](#))

Usage

```

pseudoAbsences(xy, background, realizations = 1, exclusion.buffer = 0.0166,
               prevalence = 0.5, kmeans = FALSE, varstack = NULL)

```

Arguments

xy	Data frame or list of data frames with coordinates (each row is a point), this is, presence data
background	Matrix or list/s of matrixes with background coordinates in columns. Object derived from function backgroundGrid , OCSVMprofiling or backgroundRadius .
realizations	Integer. Number of realizations (default = 1).
exclusion.buffer	value of the minimum distance to be kept between presence data and pseudo-absence data. Default is 0.0166
prevalence	Proportion of presences against absences. Default is 0.5 (equal number of presences and absences)
kmeans	Logical. If FALSE (default) pseudo-absences are generated at random. If TRUE k-means clustering of the background is done and centroids are extracted as pseudo-absences.
varstack	RasterStack of variables for to compute the k-means clustering. Used if kmeans = TRUE.

Details

Details. The application of this function could be preceded by the application of functions [OCSVMprofiling](#) and/or [backgroundRadius](#) in order to consider alternative methods for pseudo-absence data generation (see references).

Value

data frame or list/s of data frames

Author(s)

M. Iturbide

References

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutierrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*. DOI:10.1016/j.ecolmodel.2015.05.018.

See Also

[mopaTrain](#)

Examples

```
# SHORT EXAMPLE
## Load and prepare presence data
data(Q_pubescens)
presences <- Q_pubescens[sample(1:300, size = 100),]

## Define the spatial characteristics of the study area
r <- raster(nrows=50, ncols=50, xmn=-10, xmx=20, ymn=35, ymx=65, vals = rep(1, 50*50))

## Background of the whole study area
bg <- backgroundGrid(r)

## Generate pseudo-absences considering an unique background extent
RS_random <-pseudoAbsences(xy = presences, background = bg$xy,
                           exclusion.buffer = 0.083*5,
                           prevalence = -0.5, kmeans = FALSE)

# FULL WORKED EXAMPLE
## Load presence data
data(Oak_phylo2)

## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
```

```

download.file(data.url, destfile)
load(destfile, verbose = TRUE)

projection(biostack$baseline) <- CRS("+proj=longlat +init=epsg:4326")
r <- biostack$baseline[[1]]
## Background of the whole study area
bg <- backgroundGrid(r)

## Environmental profiling of the background
bg.profiled <- OCSVMprofiling(xy = Oak_phylo2, varstack = biostack$baseline,
                             background = bg$xy)

## Generate pseudo-absences considering an unique background extent
RS_random <-pseudoAbsences(xy = Oak_phylo2, background = bg$xy,
                           exclusion.buffer = 0.083*5,
                           prevalence = -0.5, kmeans = FALSE)
RSEP_random <-pseudoAbsences(xy = Oak_phylo2, background = bg.profiled$absence,
                             exclusion.buffer = 0.083*5,
                             prevalence = -0.5, kmeans = FALSE)

## Background partition into different extents
bg.extents <- backgroundRadius(xy = Oak_phylo2, background = bg$xy,
                              start = 0.166, by = 0.083*20,
                              unit = "decimal degrees")

## Generate pseudo-absences considering different background extents
TS_random <-pseudoAbsences(xy = Oak_phylo2, background = bg.extents,
                           exclusion.buffer = 0.083*10,
                           prevalence = -0.5, kmeans = FALSE)

## with k-means clustering
TS_kmeans <-pseudoAbsences(xy = Oak_phylo2, background = bg.extents,
                           exclusion.buffer = 0.083*5,
                           prevalence = -0.5, kmeans = TRUE,
                           varstack = biostack$baseline)

```

Q_pubescens

Quercus pubescens distribution

Description

A data frame with xy coordinates of *Quercus pubescens* distribution

Source

Q_pubescens is a modified subset of occurrences obtained from **GBIF.org**

References

GBIF.org (14th March 2017) GBIF Occurrence Download <http://doi.org/10.15468/dl.4ss6vr>

varianceAnalysis *Variance analysis of RasterStack objects*

Description

Extract components of lists of objects (as returned by function `mopaPredict`) and perform variance analysis to obtain raster objects of the contribution of each component to the observed variability.

Usage

```
varianceAnalysis(predictions, component1, component2, fixed = NULL)
```

Arguments

<code>predictions</code>	listed lists of raster objects as returned by <code>mopaPredict</code>
<code>component1</code>	Character. Options are "SP", "PA", "SDM", "baseClim" and "newClim" (see Details). If exist, "foldModel" is another option. Selected option corresponds to the first component in the variance analysis.
<code>component2</code>	Character. Options are "SP", "PA", "SDM", "baseClim" and "newClim" (see Details). If exist, "foldModel" is another option. Selected option corresponds to the second component in the variance analysis.
<code>fixed</code>	Optional. Character of the component names corresponding to the components that are not being analyzed (<code>component3</code> , <code>component4</code> ...). One name for each component is provided, components that only have one choice (e.g. a single species, a single baseline climate, etc.) are internally fixed. If <code>fixed = NULL</code> the first element of each component is selected. If <code>fixed</code> is specified, the selected name must be provided for each of the components that have multi-choices.

Details

Rasters are extracted using function `grep`, by matching names in the lists and characters in `component1` and `component2`. The contribution of `component1` in front `component2` to the spread (uncertainty) of the projected probabilities in `predictions` is here assessed using a simple analysis of variance approach, where the total variance (V) can be decomposed as the summation of the variance explained by `component1` (V_{comp1}), `component2` (V_{comp2}) and the combination of the previous two (V_{comp12}):

$$V = V_{comp1} + V_{comp2} + V_{comp12}.$$

Description of the components:

- SP: presence data sets
- PA: pseudo-absence realizations
- SDM: modeling algorithms

- baseClim: baseline climate, i.e. sets of variables used for model calibration in function `mopaTrain`,
- newClim: new climate, i.e. sets of variables used to project models (e.g. future climate projections) in function `mopaPredict`.

Value

A list of two RasterStack objects, the first containing the global mean and standard deviation and the second containing the percentage of variance corresponding to each component in the analysis (component1, component2 and components 1 and 2).

Author(s)

M. Iturbide

References

San Martin, D., Manzanar, R., Brands, S., Herrera, S., & Gutierrez, J.M. (2016) Reassessing Model Uncertainty for Regional Projections of Precipitation with an Ensemble of Statistical Downscaling Methods. *Journal of Climate* 30, 203-223.

Examples

```
## Load climate data
destfile <- tempfile()
data.url <- "https://raw.githubusercontent.com/SantanderMetGroup/mopa/master/data/biostack.rda"
download.file(data.url, destfile)
load(destfile, verbose = TRUE)

## Fitted models
data(mods)
?mods

## Model prediction and analysis of the variability in projections
newClim <- lapply(1:4, function(x){
  crop(biostack$future[[x]], extent(-10, 5, 35, 60))
})

prdRS.fut <- mopaPredict(models = mods, newClim = newClim)
result <- varianceAnalysis(prdRS.fut, "PA", "newClim")
splot(result$variance, col.regions = rev(get_col_regions()))
```

wrld

World map

Description

A dataset of class SpatialPolygonsDataFrame of the World map

Author(s)

M. Iturbide

Index

- *Topic **map**
 - wrld, [20](#)
- *Topic **records**
 - mods, [8](#)
 - Oak_phylo2, [14](#)
 - Q_pubescens, [18](#)
- *Topic **xy**
 - mods, [8](#)
 - Oak_phylo2, [14](#)
 - Q_pubescens, [18](#)

- auc, [11](#)

- backgroundGrid, [2](#), [4](#), [12](#), [15](#), [16](#)
- backgroundRadius, [3](#), [11](#), [12](#), [16](#), [17](#)
- best.svm, [11](#)

- earth, [10](#), [11](#)
- extent, [2](#)
- extractFromModel, [5](#), [8](#), [12](#)
- extractFromPrediction, [7](#), [9](#)

- glm, [11](#)
- grep, [19](#)

- maxent, [11](#)
- mods, [8](#)
- mopaPredict, [7](#), [8](#), [12](#), [19](#), [20](#)
- mopaTrain, [4](#), [5](#), [8](#), [9](#), [10](#), [11](#), [17](#), [20](#)

- nicheOver, [12](#)

- Oak_phylo2, [14](#)
- OCSVMprofiling, [4](#), [12](#), [14](#), [16](#), [17](#)

- pseudoAbsences, [4](#), [10–12](#), [16](#)

- Q_pubescens, [18](#)

- ranger, [11](#)
- rpart, [11](#)

- svm, [15](#)
- tree, [11](#)
- tuneRF, [11](#)

- varianceAnalysis, [19](#)

- wrld, [20](#)