

Package ‘MigClim’

February 19, 2015

Version 1.6

Date 2013-12-11

Title Implementing dispersal into species distribution models

Author Robin Engler <robin.engler@gmail.com> and
Wim Hordijk <wim@WorldWideWanderings.net> and
Loic Pellissier <loic.pellissier@unil.ch>

Maintainer Robin Engler <robin.engler@gmail.com>

Depends SDMTools, raster

Description Functions for implementing species dispersal into projections
of species distribution models (e.g. under climate change scenarios).

License GPL

NeedsCompilation yes

Repository CRAN

Date/Publication 2013-12-23 13:10:10

R topics documented:

MigClim.genClust	2
MigClim.migrate	4
MigClim.plot	8
MigClim.testData	9
MigClim.userGuide	10
MigClim.validate	10
Index	12

MigClim.genClust *Run a simulation of migration of genetic clusters.*

Description

Simulate the migration of genetic clusters. Centers of origin of the genetic clusters are picked randomly or defined by the user as the initial distribution. The simulation makes the genetic clusters migrate until the last time step for which data files are provided. Unlike the 'migrate' function in this package, this genetic clusters migration simulation assumes large time scales (e.g., 1000 years per step).

Usage

```
MigClim.genClust (hsMap="hsMap", barrier="barrier", nrClusters=4,
  nrIterations=1, threshold=445, outFile="out", initFile="")
```

Arguments

hsMap	The 'base' name of the raster files that contain the habitat suitability maps for each time step in ASCII grid format. Iteration numbers (1,2,3,...) and the file extension '.asc' are automatically added to this 'base' name to get the file name for the habitat suitability map for each next iteration. For example, if the habitat suitability raster files are named 'hsMap1.asc', 'hsMap2.asc', etc., the value of this argument should be 'hsMap'. Habitat suitability maps indicate the suitability of each cell to be colonized as a value between 0 (fully unsuitable) and 1000 (fully suitable).
barrier	The 'base' name of the raster files that contain the barriers for each time step in ASCII grid format. Iteration numbers (1,2,3,...) and the file extension '.asc' are automatically added to this 'base' name to get the file name for the barriers for each next iteration. For example, if the barrier raster files are named 'barrier1.asc', 'barrier2.asc', etc., the value of this argument should be 'barrier'. Barrier files indicate whether there is a barrier to migration present (1) or absent (0 or nodata_value) in each cell.
nrClusters	The number of genetic clusters to use.
nrIterations	The number of iterations for which the simulation will be run. This number must match the number of habitat suitability maps and barrier files.
threshold	The threshold value (in [0:1000]) above which a cell is considered suitable.
outFile	The 'base' name of the raster files that will contain the output for each time step in ASCII grid format. Iteration numbers (1,2,3,...) and the file extension '.asc' are automatically added to this 'base' name. For example, is the value of this argument is 'out', the output raster files will be named 'out1.asc', 'out2.asc', etc.
initFile	If an empty string (default value), initial starting points for the genetic clusters are generated at random, and then saved as a raster file with iteration number 0 (e.g., 'out0.asc'). Otherwise, the initial distribution is read from a file with the

name as given for this argument. The file name is assumed to be the full name (including the file extension), and to be a raster file in ASCII grid format.

Details

'nrClusters' origins of the genetic clusters represented by suitable pixels are randomly picked as the initial state. The remaining suitable pixels are assigned to one of these clusters using a nearest neighbor rule. Then, for each following time-step (e.g. every thousand years) up to the present, any suitable pixel in any timeframe t is colonized by the genetic cluster from the closest suitable pixel from timeframe $t-1$. Alternatively, a user defined distribution of the genetic clusters may be provided in ASCII grid format, with 0 (or `nodata_value`) as unsuitable, and a value of 1 to 'nrClusters' attributed to each suitable pixel. The habitat suitability maps should be provided in ASCII grid format with value from 0 (totally unsuitable) to 1000 (fully suitable), as typical output from BIOMOD. The function output will be written in ASCII grid format.

It is assumed that the input files are in the current working directory. If not, either set the working directory using 'setwd', or include the full pathname in the base name of the input files. The output files are also written to the current working directory, unless a full pathname is included.

Value

No value is returned, but the distribution of the genetic clusters for each iteration is written to a file in raster ASCII grid format.

References

Espindola A., Pellissier L., Maiorano L., Hordijk W., Guisan A., Alvarez N. Niche hindcasting across 24 millennia predicts the current intra-specific species genetic structure. *Ecology Letters*, in revision, 2012.

See Also

Migclim.validate ()

Examples

```
## Not run:
### Some example data files can be downloaded from the following web page:
### http://www.unil.ch/ecospat/page89413.html
###
### Run the example as follows (set the current working directory to the
### folder where the example data files are located):
###
MigClim.genClust (hsMap="hsMap", barrier="barrier", nrClusters=4,
  nrIterations=5, threshold=445, outFile="out", initFile="")
MigClim.validate (validateFile="Validation.txt", nrPoints=79,
  simFile="out5.asc", nrClusters=4)

## End(Not run)
```

MigClim.migrate *Initialize and run the MigClim simulation*

Description

Initialize the main MigClim simulation by passing the relevant parameter values, which will be written to a parameter file. The simulation is then run with these parameter values.

Usage

```
MigClim.migrate (iniDist="InitialDist", hsMap="HSmap", rcThreshold=0,
  envChgSteps=1, dispSteps=1, dispKernel=c(1.0,1.0),
  barrier="", barrierType="strong",
  iniMatAge=1, propaguleProd=c(1.0),
  lddFreq=0.0, lddMinDist=NULL, lddMaxDist=NULL,
  simulName="MigClimTest", replicateNb=1, overWrite=FALSE,
  testMode=FALSE, fullOutput=FALSE, keepTempFiles=FALSE)
```

Arguments

iniDist	The initial distribution of the species. This can be given either a string indicating the name of a raster file (see 'Details' for supported formats) or as a data frame object (see 'Details' for how to structure your data frame). Please note that the inputs for 'iniDist', 'hsMap' and 'barrier' (optional) must always be given in the same format. Note that the values of the species' initial distribution layer must be binary and integer numbers: 1 (species is present) or 0 (species is absent).
hsMap	The habitat suitability values. This can be given as a string indicating the 'base name' of the raster files that contain the habitat suitability maps. Iteration numbers (1,2,3,...) are automatically added to this 'base name' to get the file name for the habitat suitability map for each successive environmental change iteration (see the 'Details' section for supported formats). Alternatively, the habitat suitability information can also be given as a data frame object, where each column indicates a successive habitat suitability map (see the 'Details' section for further information on how this data frame must be structured). Note that the values of the habitat suitability layers must be integer numbers in the range 0 to 1000.
rcThreshold	The reclassification threshold: an integer value between 0 and 1000; default=0). If 'rcThreshold > 0', then the continuous values of the habitat suitability maps (in the range 0:1000) will be reclassified according to 'rcThreshold'. Values of habitat suitability < 'rcThreshold' are reclassified to '0' (unsuitable habitat) and values >= 'rcThreshold' are reclassified to '1000' (fully suitable habitat). In the case where 'rcThreshold=0', the habitat suitability values are not reclassified, and are instead considered as habitat 'invasibility', modulating the probability of an unoccupied cell to become colonized (probabilities are computed as 'habitat suitability / 1000').

envChgSteps	The number of environmental change steps to perform. At each environmental change step the habitat suitability values are updated with the values of the corresponding habitat suitability map (and therefore the number of environmental change steps must match the number of habitat suitability maps available).
dispSteps	The number of dispersal steps to perform within each environmental change step. For instance, if one wants to simulate dispersal to occur once a year, and the habitat suitability maps represent 5 years intervals, then 'dispSteps' should be set to 5.
dispKernel	The dispersal kernel. A vector of dispersal probabilities (values in the range 0.0 to 1.0) giving the conditional probability for a source cell to colonize an empty cell given the distance between both cells. The distance unit is the 'pixel', with the first value in the vector representing the probability for a source cell to colonize a directly adjacent cell. See also the MigClim user guide (available by typing 'MigClim.userGuide' in R) for more details on this parameter.
barrier	The name of the raster file that contains barrier information or a single column data frame (or vector) containing this information. If an empty string is given (default value), no barrier information is used. The values of the barrier layer must integer numbers and binary: either 1 (indicating that the cell is a barrier) or 0 (indicating that the cell is not a barrier).
barrierType	The barrier type to use. Values can be either 'strong' (default value) or 'weak'. Not relevant if barrier information is not used. 'weak' barriers will allow dispersal to proceed through two diagonally adjacent barrier pixels, 'strong' barriers won't. See the MigClim user guide (type 'MigClim.userGuide()' in R) for detailed explanations of the difference between these two barrier types.
iniMatAge	The initial maturity age of newly colonized cells. Newly colonized cells younger than this age cannot produce propagules and hence cannot colonize other cells. When newly colonized cells reach an age equal to 'iniMatAge', then their probability to produce propagules is set to the first value indicated in the 'propaguleProd' vector. The time unit that measures cell 'age' is a dispersal step, which usually should be equal to a year.
propaguleProd	The propagule production probability as a function of cell 'age'. A vector where each successive value indicates the propagule production probability of a cell that has reached its 'iniMatAge' age. The first value of the vector corresponds to the cells having an age equal to 'iniMatAge' and successive values correspond to an increase in 1 unit of age from the 'iniMatAge'. When the probability of propagule production reaches 1 (full maturity age), then it is no longer needed to indicate this value in the 'propaguleProd' vector as it will be considered to be 1 from then on. The length of the 'propaguleProd' vector is thus equal to 'full maturity age'-'iniMatAge' (or a length of 1 if both ages are equal). Propagule production probabilities must be given in the range 0.0 to 1.0.
lddFreq	The long-distance dispersal frequency, i.e., the probability for an occupied cell with full propagule production potential to generate a long distance dispersal event. If set to 0.0 (default), no long-distance dispersal is performed. Value should be given in the range 0.0 and 1.0.
lddMinDist	The minimum distance for long-distance dispersal (not used if 'lddFreq=0'). This value must be larger than the regular dispersal distance given by the length of 'dispKernel'.

<code>lddMaxDist</code>	The maximum distance for long-distance dispersal (not used if <code>'lddFreq=0'</code>). This value must be \geq <code>'lddMinDist'</code> .
<code>simulName</code>	The 'base name' to be used for the different outputs produced by the MigClim simulation. Three different types of outputs are produced by the <code>'MigClim.migrate()'</code> function: <code>ascii</code> grid files named <code>'simulName'+'_raster.asc'</code> that contains the final state of the simulation, <code>'simulName'+'_stats.txt'</code> files that contain the simulation's outputs after each dispersal step, and <code>'simulName'+'_summary.txt'</code> files that contain a single-line summary of the entire simulation.
<code>replicateNb</code>	Number of times a simulation should be replicated. The final outputs include all the outputs from individual runs as well as the average of all runs.
<code>overWrite</code>	If <code>'TRUE'</code> then any existing file with the same name as an output of the MigClim.migrate function will be mercilessly overwritten. If <code>'FALSE'</code> then the function will stop if any output file does already exist.
<code>testMode</code>	If <code>'TRUE'</code> then the MigClim.migrate function will check all the provided input data but will not run the actual simulation. Useful for testing your data before running several successive simulations or simulations that might take a long time.
<code>fullOutput</code>	If <code>'TRUE'</code> , the current state of the simulation is written to an ASCII raster file after each dispersal step (allowing to reconstruct the dispersal process at each step). If <code>'FALSE'</code> (default), only the final state of the simulation is written to an ASCII grid file.
<code>keepTempFiles</code>	If <code>'FALSE'</code> (default), then any <code>'.asc'</code> file created from a conversion process in the function will be deleted when the simulation completes. If you wish to keep these files then set the value of this parameter to <code>'TRUE'</code> .

Details

The input data for initial distribution (`'iniDist'`), habitat suitability (`'hsMap'`), and (optionally) barriers (`'barrier'`) can be provided as either a string giving the name of a raster file (the name should be given relative to the working directory) or as a data frame object. For a given simulation, all these inputs must be given in the same format. Option 1: Input data provided as raster files. In this case, the input must be a string that contains the name of the raster files relative to the working directory. The following raster formats are supported: (i) `ascii` grid (files must have a `'.asc'` extension), (ii) R rasterLayer (see `'raster'` package), (iii) ESRI GRID, (iv) GeoTIFF (files must have a `'.tif'` extension). Note that all input grids need to have exactly the same pixel size and the same extent (i.e. the same number of rows and columns).

The standard ASCII grid Raster format looks as follows (actual values depend on file content):

```
ncols      100
nrows      100
xllcorner  0
yllcorner  0
cellsize   10
nodata_value -9999
0 0 1 0 0 1 1 1 0 ...
1 0 1 0 1 0 1 0 1 ...
:
```

:

Option 2: Input data provided as data frames. It is important to note that, unlike when inputting data in a raster format, inputting it as data frames requires to provide the actual data frame object to the function, not a string indicating its name. 'iniDist' should have 3 columns (in this order): X coordinate, Y coordinate and initial distribution of the species. Only values of 0 or 1 are accepted: 1 = species present, 0 = species absent.

Xcoordinate	Ycoordinate	IniDist
520000	200100	1
520050	200100	1
520100	200100	0
...		

'hsMap' should have a number of columns equal to 'envChgSteps'. Each column giving the suitability of the habitat for a given environmental change step (in chronological order). Values of habitat suitability must be in the range [0:1000].

HSmap1	HSmap2	HSmap3	...
789	800	230	...
560	720	150	...
120	540	86	...
...			

'barrier' is optional and must have only one column (it can also be a vector). It must contain only values of either 1 (pixel is a barrier feature) or 0 (pixel is not a barrier feature).

The function output(s) will be written in ascii GRID format (with .asc extension).

Value

The number of environmental change steps performed. The function also writes the following outputs into the current working directory: an ASCII grid raster file named 'simulName'+ '_raster.asc' that contains the final state of the simulation, a 'simulName'+ '_stats.txt' file that contains the simulation's outputs after each dispersal event, and a 'simulName'+ '_summary.txt' file that contains a single-line summary of the entire simulation. If fullOutput=TRUE then an ASCII raster file containing the state of the simulation at the end of each dispersal step is also saved as output with the following name structure: 'simulName' + '_step_' + dispersal step code + '.asc'. The output file 'simulName' + '_stats.txt' contains summary statistics for each individual dispersal step. The output file 'simulName' + '_summary.txt' contains summary statistics over the entire simulation (including the running time in seconds).

References

Engler R., Hordijk W. and Guisan A. The MigClim R package - seamless integration of dispersal constraints into projections of species distribution models. *Ecography*, in review.

See Also

MigClim.plot(), MigClim.userGuide()

Examples

```

data(MigClim.testData)

### Run MigClim with a data frame type input.
n<-MigClim.migrate (iniDist=MigClim.testData[,1:3],
  hsMap=MigClim.testData[,4:8], rcThreshold=500,
  envChgSteps=5, dispSteps=5, dispKernel=c(1.0,0.4,0.16,0.06,0.03),
  barrier=MigClim.testData[,9], barrierType="strong",
  iniMatAge=1, propaguleProd=c(0.01,0.08,0.5,0.92),
  lddFreq=0.1, lddMinDist=6, lddMaxDist=15,
  simulName="MigClimTest", replicateNb=1, overWrite=TRUE,
  testMode=FALSE, fullOutput=FALSE, keepTempFiles=FALSE)

## Not run:   ### Convert data from a dataframe to ASCII grid format
library(SDMTools)
dataframe2asc(MigClim.testData[,c(2,1,3:9)])
### and then run MigClim with "raster" type input.
n<-MigClim.migrate (iniDist="InitialDist",
  hsMap="HSmep", rcThreshold=500,
  envChgSteps=5, dispSteps=5, dispKernel=c(1.0,0.4,0.16,0.06,0.03),
  barrier="Barrier", barrierType="strong",
  iniMatAge=1, propaguleProd=c(0.01,0.08,0.5,0.92),
  lddFreq=0.1, lddMinDist=6, lddMaxDist=15,
  simulName="MigClimTest", replicateNb=1, overWrite=TRUE,
  testMode=FALSE, fullOutput=FALSE, keepTempFiles=FALSE)
## End(Not run)

```

MigClim.plot

Plots ascii raster outputs of the MigClim.migrate function

Description

Plots ascii raster outputs of the MigClim.migrate function with an adequate color scale and saves the result as a JPEG or PNG file (or does simply display it in the R console).

Usage

```
MigClim.plot(asciiFile, outDir="", fileFormat="jpeg", fullOutput=FALSE)
```

Arguments

asciiFile	Name of the main ascii raster output "_raster.asc" obtained when running a simulation using MigClim.migrate. If the ascii file is not located in the working directory, then the file name should include its path as well. File names should end in "_raster.asc" or "_raster".
outDir	Optional output directory. When this value is set to its default ("") then the output files are saved in the same directory as the ascii grids that are being plotted.

fileFormat	Format in which the output graphics are saved. Can be either "jpeg", "png" or "inR". If "inR" is passed as parameter value, the plotted image is not saved to disk but simply displayed in the R console.
fullOutput	If "TRUE" then all the intermediate outputs for the given simulation are also plotted and saved to disk.

Value

The function does not return any value but saves "jpeg" or "png" image files to disk, or displays a graphic in the R console (fileFormat="inR").

See Also

MigClim.migrate ()

MigClim.testData	<i>Test data for the MigClim dispersal simulation tool</i>
------------------	--

Description

Data frame that contains data for simulating the dispersal of a species under environmental change conditions. The data frame contains information on the species' initial distribution in the landscape, how the suitability of its habitat evolves over time, and some information on which areas in the landscape represent barriers to dispersal for the species. Each row of the data frame contains the information for a given 50 x 50m pixel in the landscape

Usage

```
data(MigClim.testData)
```

Format

MigClim.testData is a data frame with the following columns Xcoord: The X coordinate of the given pixel. Ycoord: The Y coordinate of the given pixel. InitialDist: The species's initial distribution. 1 = species presence, 0 = species absence. HSmap1: Habitat suitability in the range 0:1000 for the first time step. HSmap2: Habitat suitability in the range 0:1000 for the second time step. HSmap3: Barrier: Indicate whether a given pixel is a barrier feature or not (barrier features impede dispersal across them). 1 = pixel is a barrier, 0 = pixel is not a barrier.

MigClim.userGuide *Displays the MigClim user guide in PDF format*

Description

Displays the MigClim user guide in PDF format. The function takes no argument.

Usage

```
MigClim.userGuide()
```

Value

The function does not return any value but opens the "MigClim_userGuide.pdf" PDF file containing the MigClim user guide.

See Also

MigClim.migrate ()

MigClim.validate *Validation of genetic clusters migration simulation results.*

Description

Compare the output of a genetic clusters migration simulation to an observed genetic clusters distribution.

Usage

```
MigClim.validate (validateFile="Validation.txt", nrPoints=0, simFile="out1.asc",
nrClusters=4)
```

Arguments

validateFile	The name of the file containing the observed distribution. See details below for the required file format. A full file name (including file extension) is expected.
nrPoints	The number of points in the observed distribution in the 'validateFile' file.
simFile	The name of an output file from the genetic cluster migration simulation. Here, a full file name is expected, i.e., including an iteration number and the '.asc' extension, for example 'out5.asc'.
nrClusters	The number of genetic clusters.

Details

This function calculates a matching score between the genetic clusters distribution predicted by a simulated scenario and an observed population genetic structure. The observed distribution (i.e., the validation file) should be in the following format:

```
N X      Y      C
3 24.62192 45.60239 1
18 19.23068 48.88922 1
23 24.63611 46.59972 2
25 23.59028 46.70118 3
30 23.32314 42.19968 4
:   :         :     :
```

The first line in the file is a header line, and each next line should contain four values, being an identification number (not used here), an X-coordinate, a Y-coordinate, and finally the number of the genetic cluster the point belongs to. There should be 'nrPoints' lines in this file.

Value

A score (between 0 and 1) indicating the fit of the distribution in the output file. This is given as an array with two elements. The first element gives the "total" score, i.e., measured over all points regardless of which cluster each point belongs to. The second element give the "average" score, where the fit is calculated over each cluster separately, and then averaged over all cluster scores (so small clusters get the same weight as large clusters).

See Also

MigClim.genClust ()

Index

MigClim.genClust, [2](#)
MigClim.migrate, [4](#)
MigClim.plot, [8](#)
MigClim.testData, [9](#)
MigClim.userGuide, [10](#)
MigClim.validate, [10](#)