Package 'seraphim'

December 3, 2025

Type Package

Title Studying environmental rasters and phylogenetic informed
movements

Version 2.0

Date 2025-08-31

Author Simon Dellicour, Rebecca Rose, Nuno Faria, Philippe Lemey, Oliver G. Pybus

Maintainer Simon Dellicour <simon.dellicour@ulb.be>

Depends ape, doMC, fields, gdistance, HDInterval, ks, phytools,
raster, RColorBrewer, R.utils, vegan

Description
The source code for SERAPHIM is distributed under a GNU Lesser General Public License

License GPL-3

NeedsCompilation no

Contents

seraphim-package	*A package to analyse continuous phylogeographic reconstructions and conduct hypotheses testing based on landscape phylogeographic analyses.*

Description

The package "seraphim" (for "studying environmental rasters and phylogenetic informed movements"; Dellicour *et al*. 2016a) can, for instance, be used to investigate the association between environmental factors and the diffusion velocity (Dellicour *et al*. 2025) of lineages for which the dispersal history has been inferred by a continuous phylogeographic analysis (Lemey *et al*. 2010). In such analytical procedure initially described in Dellicour (*et al*. 2016b), functions of the R package "seraphim" can be used to (i) extract the spatio-temporal information embedded in spatiotemporally-referenced trees inferred, for instance, by the continuous phylogeographic method implemented in BEAST (Lemey *et al*. 2010), (ii) compute the environmental distances associated with each phylogenetic branch for given environmental rasters, (iii) estimate the correlation between the dispersal duration and these environmental distances associated with each branch, and (iv) estimate statistical supports for these correlations using null models generated by randomisation procedures (Dellicour *et al*. 2017). In addition, the package "seraphim" can be used to display the outcome of a continuous phylogeographic reconstruction, estimate a series of lineage dispersal statistics (Dellicour *et al*. 2024), investigate the impact of environmental factors on the *dispersal frequency* (Dellicour *et al*. 2018) and *dispersal location/direction* (Dellicour *et al*. 2019) of lineages. Finally, the package also includes different different functions to conduct phylogeographic simulations either along phylogenetic trees (Dellicour *et al*. 2018) or with a birth-death process (Dellicour *et al*. 2024, 2025).

Author(s)

Simon Dellicour, Rebecca Rose, Nuno Faria, Philippe Lemey, Oliver G. Pybus

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Dellicour S, Rose R, Faria NR, Vieira LFP, Bourhy H, Gilbert M, Lemey P, Pybus OG (2017). Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus epidemics. *Molecular Biology \& Evolution* 34: 2563-2571.

Dellicour S, Baele G, Dudas G, Faria NR, Pybus OG, Suchard M, Rambaut A, Lemey P (2018). Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak. *Nature Communications* 9: 2222.

Dellicour S, Bastide P, Rocu P, Fargette D, Hardy OJ, Suchard MA, Guindon S, Lemey P (2024). How fast are viruses spreading in the wild? *PLoS Biology* 22: e3002914.

Dellicour S, Gámbaro F, Jacquot M, Lequime S, Baele G, Gilbert M, Pybus OG, Suchard MA, Lemey P (2025). *Comparative performance of viral landscape phylogeography approaches. Proceedings of the National Academy of Sciences of the USA* 122: e2506743122.

Lemey P, Rambaut A, Welch JJ, Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology & Evolution* 27: 1877-1885.

See Also

https://beast.community/
http://www.circuitscape.org
https://github.com/Circuitscape/Circuitscape.jl

---

cartogramTransformation

A function to modify sampling coordinates according to the cartogram transformation of environmental rasters

---

Description

This function first performs cartogram transformation (Dougenik *et al*. 1985) of environmental rasters based on their cell values. Cartogram transformations are here performed while considering each raster cell as a distinct polygon for which an environmental value is assigned (i.e. the original raster cell value). In a second step and for each cartogram transformation that has been performed, the function then modifies sampling geographic coordinates accordingly. These sampling coordinates have to be contained in sequence names available in a fasta file or in a data frame containing these coordinates (see below for further detail).

Usage

```
cartogramTransformation(input,
                        envVariables,
                        resistances = list(),
                        outputName = "")
```

Arguments

input   list of sequences in a fasta format or data frame only containing the sampling coordinates to transform. In the first case, the fasta object has to be obtained by reading a fasta file withe the "read.dna" function of the "ape" package: read.dna ("file_name.fasta", format="fasta"). Important requirement: in the fasta file, the sequence names have to contain the following details in a specific order and separated by underscores: (i) sequence ID (e.g., GenBank accession number), (ii) sampling date in a decimal format (e.g., 2015.47), (iii) latitude coordinate and (iv) longitude coordinate in a decimal format. Here is an example: ">hOH10_1997.2_41.053_-80.706". See the example files for further details.

In the second case, the data frame has to contain only four columns: (i) the genomic sequence IDs, (ii) their collection dates, as well as their sampling (iii) latitude and (iv) longitude coordinates to transform.

envVariables   list of environmental rasters to that will be used to perform distinct cartogram transformations.

resistances   list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the rasters of the "envVariables" list have to be treated as resistance or conductance factors. If treated as a conductance factor, raster cell values will be inverted (1/x) prior to the cartogram transformation.

outputName   name (prefix) to give to the different output files.

Value

The function performs one cartogram transformation per provided environmental raster and creates a new fasta/text file per cartogram transformation. In the case of new fasta files, only the sequence names are modified: longitude and latitude coordinates are updated according to the cartogram transformation and thus indicate the new position of each sampled sequence in the transformed space. If text files are generated, these files will then only contain the transformed coordinates.

Note

The use of this function requires the preliminary installation of two additional R packages: the "sf" and "cartogram" packages.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Gámbaro F, Jacquot M, Lequime S, Baele G, Gilbert M, Pybus OG, Suchard MA, Lemey P (2025). *Comparative performance of viral landscape phylogeography approaches. Proceedings of the National Academy of Sciences of the USA* 122: e2506743122.

Dougenik JA, Chrisman NR, Niemeyer DR (1985). An algorithm to construct continuous area cartograms. *The Professional Geographer* 37: 75-81.

---

circuitScape1   *A function to launch Circuitscape analyses with the related Python package*

Description

The function calls the external Python library "circuitscape" to compute the environmental distance associated with each phylogenetic branch of a spatially annotated tree, e.g., inferred by a continuous phylogeographic analysis.

Usage

```
circuitScape1(envVariable,
              envVariableName,
              resistance = TRUE,
              avgResistance = TRUE,
              fourCells = FALSE,
              fromCoor,
              toCoor,
              OS = "Unix",
              prefix = "",
              ID = "",
              nberOfCores_CS=1)
```

Arguments

envVariable	raster file on which environmental distances have to be computed.
envVariableName	name (string) of the raster file. This name can be different from the name saved in the raster object and will be used to name the temporary file generated for the Circuitscape analysis.
resistance	boolean variable specifying if the raster will be treated as a resistance or conductance factor.
avgResistance	boolean variable specifying if the raster cells have to be connected by average resistance (TRUE) or average conductance (FALSE).
fourCells	boolean variable specifying if a given raster cell should be connected to either its four first-order (TRUE) or eight first- and second-order (FALSE). This option is only relevant for the least-cost and Circuitscape path models.
fromCoor	matrix containing the coordinates of the starting nodes of each phylogenetic branch coming from one tree (one line per node, 1st column with the longitude and the 2nd column with the latitude).
toCoor	matrix containing the coordinates of the ending nodes of each phylogenetic branch coming from one tree (one line per node, 1st column with the longitude and the 2nd column with the latitude).
OS	operating system on which the function is used ("Unix" or "Windows").
prefix	a character string that will be used as a prefix to name the temporary folders generated for Circuitscape.
ID	a number to identify the Circuitscape operation.
nberOfCores_CS	number of available cores to parallelise the computations in Circuitscape (we advise to set this parameter value to one and rather use the "numberOfCores" variable to set a parallelisation process).

Value

The function returns a vector containing the environmental distance computed for each phylogenetic branch.

Note

This function requires the preliminary installation of the program Circuitscape as a Pyhton package. Instructions on how to install this Python package are available on the Circuitscape website:

https://circuitscape.org/.

Also note that "fromCoor" and "toCoor" matrices have to display the same dimensions and lines in the same order: coordinates in line *i* in "fromCoor" have to be the coordinates of the starting node connected to the ending node for which geographic coordinates are on line *i* in "toCoor"".

This function can be used to compute (resistance or conductance) environmental distance associated with any movement vector for which start and end locations are respectively specified in the "fromCoor" and "toCoor" arguments.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

McRae BH (2006). Isolation by resistance. *Evolution* 60: 1551-1561.

---

circuitScape2	*A function to launch Circuitscape analyses with the related Julia package*

Description

The function calls the external Julia library "circuitscape" to compute the environmental distance associated with each phylogenetic branch of a spatially annotated tree, e.g., inferred by a continuous phylogeographic analysis.

Usage

```
circuitScape2(envVariable,
             envVariableName,
             resistance = TRUE,
             avgResistance = TRUE,
             fourCells = FALSE,
             fromCoor,
             toCoor,
             OS = "Unix",
             prefix = "",
             ID = "",
             nberOfCores_CS=1)
```

Arguments

envVariable	raster file on which environmental distances have to be computed.
envVariableName	
	name (string) of the raster file. This name can be different from the name saved in the raster object and will be used to name the temporary file generated for the Circuitscape analysis.
resistance	boolean variable specifying if the raster will be treated as a resistance or conductance factor.
avgResistance	boolean variable specifying if the raster cells have to be connected by average resistance (TRUE) or average conductance (FALSE).
fourCells	boolean variable specifying if a given raster cell should be connected to either its four first-order (TRUE) or eight first- and second-order (FALSE). This option is only relevant for the least-cost and Circuitscape path models.
fromCoor	matrix containing the coordinates of the starting nodes of each phylogenetic branch coming from one tree (one line per node, 1st column with the longitude and the 2nd column with the latitude).
toCoor	matrix containing the coordinates of the ending nodes of each phylogenetic branch coming from one tree (one line per node, 1st column with the longitude and the 2nd column with the latitude).
OS	operating system on which the function is used ("Unix" or "Windows").
prefix	a character string that will be used as a prefix to name the temporary folders generated for Circuitscape.
ID	a number to identify the Circuitscape operation.
nberOfCores_CS	number of available cores to parallelise the computations in Circuitscape (we advise to set this parameter value to one and rather use the "numberOfCores" variable to set a parallelisation process).

Value

The function returns a vector containing the environmental distance computed for each phylogenetic branch.

Note

This function requires the preliminary installation of the program Circuitscape as a Julia package. Instructions on how to install this Julia package are available here: https://docs.circuitscape.org/Circuitscape.jl/latest/

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

McRae BH (2006). Isolation by resistance. *Evolution* 60: 1551-1561.

---

isolationByResistance	*A function to conduct "isolation-by-resistance" analyses, i.e. to investigate the impact of environmental factors on the isolation-by-distance pattern*

Description

This function estimates and tests the correlation between the log-transformed environmental distance(s) and patristic distance associated with each pair of tip nodes in posterior trees inferred through Bayesian continuous phylogeographic inference. While the continuous phylogeographic reconstruction is not used to compute the correlation metric, it is however use to generate a null dispersal model exploited to estimate its statistical support. The function performs three analytical steps: (i) for each environmental factor, the computation of the environmental distances associated with each pair of tip nodes (using the "straight-line", "least-cost" or "Circuitscape" path model, see the details below), (ii) the estimation of the correlation between the patristic distances and the corresponding log-transformed environmental distances using a linear regression approach, and (iii) an optional randomisation step to estimate the statistical support of the correlation metric (Q). To generate null models, several randomisation procedures of the branches position are implemented (see below for further detail).

Usage

```
isolationByResistance(localTreesDirectory = "",
                      nberOfExtractionFiles = 1,
                      envVariables=list(),
                      pathModel = 2,
                      resistances = list(),
                      avgResistances = list(),
                      fourCells = FALSE,
                      nberOfRandomisations = 0,
                      randomProcedure = 3,
                      outputName = "",
                      showingPlots = FALSE,
                      nberOfCores = 1,
                      OS = "Unix",
                      juliaCSImplementation = FALSE)
```

Arguments

localTreesDirectory
:   directory where the tree extraction matrices are saved (see the "treeExtractions" or "postTreeExtractions" function).

nberOfExtractionFiles
:   number of annotated phylogenetic trees from which spatio-temporal information has been extracted.

envVariables
:   list of environmental rasters to test.

pathModel	path model to use for computing environmental distances: "1" (straight-line path model), "2" (least-cost path model, Dijkstra 1959) or "3" (Circuitscape path model, McRae 2006).
resistances	list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the rasters of the "envVariables" list have to be treated as resistance or conductance factors (only relevant for the least-cost and Circuitscape path models).
avgResistances	list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the raster cells have to be connected by average resistance (TRUE) or average conductance (FALSE). This option is only relevant for the Circuitscape path model.
fourCells	boolean variable specifying if a given raster cell should be connected to either its four first-order (TRUE) or eight first- and second-order (FALSE). This option is only relevant for the least-cost and Circuitscape path models.
nberOfRandomisations	
	number of randomisation steps to perform in order to test the correlation between the patristic and log-transformed environmental distances.
randomProcedure	
	randomisation procedure to use: "1" (using external tree randomisations previously performed with the "treesRandomisation" function), "2" (using external RRW simulations previously performed with the "simulatorRRW1" function), "3" (randomisations of nodes position while maintaining the branches length, the tree topology and the location of the most ancestral node), "4" (randomisations of nodes position while maintaining the branches length and the location of the starting node of each branch), "5" (randomisations of nodes position while maintaining the branches length and the location of the ending node of each branch), or "6" (randomisations of nodes position while maintaining only the branches length).
outputName	name (prefix) to give to the different output files.
showingPlots	boolean variable specifying if a plot has to be displayed during the randomisation procedure.
nberOfCores	number of available cores to parallelise the computations in R (parallelisation will only work on Unix operating systems).
nberOfCores_CS	number of available cores to parallelise the computations in Circuitscape (we advise to set this parameter value to one and rather use the "numberOfCores" variable to set a parallelisation process).
OS	operating system on which the function is used ("Unix" or "Windows").
juliaCSImplementation	
	boolean variable specifying if the Julia implementation of Circuitscape should be called by the function. By default (FALSE), the Python implementation is called.

Details

Three different path models are available to compute the environmental distances associated with each pair of tip nodes: (i) the straight-line path model only considering the raster cells below the straight-line segment connecting each pair of tip nodes. In this model, the environmental distance associated with a given pair of tip nodes for a given environmental variable will be the sum of the values of the raster cells below the straight-line segment connecting both tip nodes. (ii) The least-cost path model (Dijkstra 1959). In this second model, environmental distance associated with a

given pair of tip nodes for a given environmental variable will be the sum of the values of the raster cells selected by a shortest-path algorithm while considering the environmental variable as a resistance or a conductance factor. (iii) The Circuitscape model based on circuit theory (McRae 2006). This latter model integrates the contribution of multiple possible pathways to compute the environmental distance (McRae 2006). The environmental distances based on the least-cost path model are here computed with the "costDistance" function available in the "gdistance" R package (van Etten 2012) and the environmental distances based on the Circuitscape path model are computed using the Python or Julia package "circuitscape" (McRae 2006).

Regarding the test to estimate the statistical support of the correlation metric Q, several randomisation procedures have been implemented to create null models: (i) using already performed tree randomisations (see the "treesRandomisation" function for further detail), (ii) using external RRW simulations corresponding to each tree (see the "simulatorRRW1" function for further detail), (iii) a randomisation of the branches position while maintaining their length, the tree topology and the location of the most ancestral node, (iv) a randomisation of the branches position while maintaining their length and the location of the starting node of each branch, and (v) a randomisation of the branches position while only maintaining their length. The latter procedure can be useful, e.g., when analysing non phylogenetically informed movement data like capture-mark-recapture movement data.

Value

The function generates several output files. If the number of randomisations to perform is set to zero, the function will only generate a text file listing the linear regression results for each phylogenetic tree: the β regression coefficient and coefficient of determination R^2_{env} obtained from the univariate linear regression between the patristic distances and the log-transformed environmental distances computed for a given environmental factor, as well as the difference Q between R^2_{env} and R^2_{null} (the coefficient of determination obtained from the linear regression between the patristic distances and the log-transformed geographical distances associated with each pair of tip nodes). If the number of randomisations to perform is higher than zero, the function will return a second text file reporting the Bayes factor support approximated for each environmental factor and each randomisation step (similar to the approach described in Dellicour *et al*. 2017; see also the related tutorial for detailed information about the Bayes factor estimations). Note that the geographical distance is computed using the selected environmental distance computation method on a "null" raster with uniform cell values equal to "1".

Note

The computation of Circuitscape environmental distances requires the preliminary installation of the program Circuitscape as a Python or Julia package. Instructions on how to install this Python package are available on the Circuitscape website: https://circuitscape.org/

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Faria NR, Vieira LFP, Bourhy H, Gilbert M, Lemey P, Pybus OG (2017). Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus

epidemics. *Molecular Biology \& Evolution* 34: 2563-2571.

Dellicour S, Gámbaro F, Jacquot M, Lequime S, Baele G, Gilbert M, Pybus OG, Suchard MA, Lemey P (2025). *Comparative performance of viral landscape phylogeography approaches. Proceedings of the National Academy of Sciences of the USA* 122: e2506743122.

Dijkstra EW (1959). A note on two problems in connexion with graphs. *Numerische Mathematik* 1: 269-271.

McRae BH (2006). Isolation by resistance. *Evolution* 60: 1551-1561.

Van Etten J (2012). R package gdistance: distances and routes on geographical grids. R package version 1.1.

See Also

http://cran.r-project.org/web/packages/gdistance/index.html
http://www.circuitscape.org
https://github.com/Circuitscape/Circuitscape.jl

---

mccTreeExtractions	*A function to extract the spatio-tremporal information embedded in the maximum clade credibility (MCC) tree obtained from a continuous phylogeographic inference conducted with the software package BEAST*

Description

This function extracts the spatio-temporal information contained in the maximum clade credibility (MCC) phylogenetic trees inferred by the continuous phylogeographic method implemented in BEAST (Lemey *et al*. 2010). To use this function the "latitute-longitude" continuous traits partition should have been named "location" while preparing the input XML file for the continuous phylogeographic analysis conducted with BEAST.

Usage

```
mccTreeExtractions(mcc_tre, mostRecentSamplingDatum)
```

Arguments

mcc_tre   MCC tree read/imported by the "readAnnotatedNexus" function of the same package.

mostRecentSamplingDatum
                   most recent sampling date in a decimal format (e.g., 2007.4).

Value

The function returns an extraction table containing the spatio-tremporal information extracted from the MCC tree.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Lemey P, Rambaut A, Welch JJ, Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology & Evolution* 27: 1877-1885.

---

mdsTransformation	*A function to transform sampling coordinates according to multidimensional scaling (MDS) transformations based on pairwise environmental distances*

Description

This function transforms sampling coordinates according to multidimensional scaling (MDS) analyses based on pairwise environmental distances computed among sampling points. A MDS analysis is performed for each specified environmental raster, and original sampling coordinates are replaced by new coordinates found on two main MDS dimensions. Original sampling coordinates have to be contained in sequence names available in a fasta file or in a data frame containing these coordinates (see below for further details).

Usage

```
mdsTransformation(input,
                envVariables,
                pathModel = 2,
                resistances = list(),
                avgResistances = list(),
                fourCells = FALSE,
                outputName = "",
                OS = "Unix")
```

Arguments

input
: list of sequences in a fasta format or data frame containing the sampling coordinates to transform. In the first case, the fasta object has to be obtained by reading a fasta file withe the "read.dna" function of the "ape" package: read.dna ("file_name.fasta", format="fasta"). Important requirement: in the fasta file, the sequence names have to contain the following details in a specific order and separated by underscores: (i) sequence ID (e.g., GenBank accession number), (ii) sampling date in a decimal format (e.g., 2015.47), (iii) latitude coordinate and (iv) longitude coordinate in a decimal format. Here is an example: ">hOH10_1997.2_41.053_-80.706". See the example files for further details. In the second case, the data frame has to contain only four columns: (i) the genomic sequence IDs, (ii) their collection dates, and their sampling (iii) latitude and (iv) longitude coordinates to transform.

envVariables
: list of environmental rasters that will be used to compute the pairwise environmental distances on which MDS analyses will be based.

pathModel
: path model to use for computing environmental distances: "1" (straight-line path model), "2" (least-cost path model, Dijkstra 1959) or "3" (Circuitscape path model, McRae 2006).

resistances	list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the rasters of the "envVariables" list have to be treated as resistance or conductance factors (only relevant for the least-cost and Circuitscape path models).
avgResistances	list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the raster cells have to be connected by average resistance (TRUE) or average conductance (FALSE). This option is only relevant for the Circuitscape path model.
fourCells	boolean variable specifying if a given raster cell should be connected to either its four first-order (TRUE) or eight first- and second-order (FALSE). This option is only relevant for the least-cost and Circuitscape path models.
outputName	name (prefix) to give to the different output files.
OS	operating system on which the function is used ("Unix" or "Windows").

Value

For each environmental raster that is provided, the function first computes pairwise environmental distances among all sampling coordinates. These pairwise environmental distances are then used to perform MDS analyses and tranform the original geographic coordinates by taking new values on the two main MDS dimensions. In practice, new fasta or text files are generated, one per MDS analysis that is performed. In the case of new fasta files, only the sequence names are modified: longitude and latitude coordinates are updated according to the MDS transformation and thus indicate the new position of each sampled sequence in the transformed space. If text files are generated, these files will then contain the transformed coordinates.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Gámbaro F, Jacquot M, Lequime S, Baele G, Gilbert M, Pybus OG, Suchard MA, Lemey P (2025). *Comparative performance of viral landscape phylogeography approaches. Proceedings of the National Academy of Sciences of the USA* 122: e2506743122.

Dijkstra EW (1959). A note on two problems in connexion with graphs. *Numerische Mathematik* 1: 269-271.

McRae BH (2006). Isolation by resistance. *Evolution* 60: 1551-1561.

Van Etten J (2012). R package gdistance: distances and routes on geographical grids. R package version 1.1.

See Also

http://cran.r-project.org/web/packages/gdistance/index.html
http://www.circuitscape.org
https://github.com/Circuitscape/Circuitscape.jl

plotRaster	*A function to plot a raster with a simplified layout*

Description

This function generates a raster plot using a simplified layout.

Usage

```
plotRaster(rast,
           cols = NULL,
           colNA = "gray90",
           add = FALSE,
           new = TRUE,
           addBox = TRUE,
           addAxes = FALSE,
           addLegend = FALSE,
           legendOnly = FALSE)
```

Arguments

rast	raster object to plot.
cols	colour scale to use (by default, it will be the "RdBu" colour scale from the "RColorBrewer" package.
colNA	colour to use for the "NA" raster cell values (default is "gray90").
add	boolean variable specifying if the plot has to be added to the current plot.
new	boolean variable specifying if the plot has to be displayed in a new window (only used if "add" is set to "FALSE").
addBox	boolean variable specifying if a box has to be added around the raster plot (only used if "add" is set to "FALSE").
addAxes	boolean variable specifying if axes displaying minimum and maximum geographic coordinates have to added to the plot (only used if "add" is set to "FALSE").
addLegend	boolean variable specifying if a legend has to be added to the plot.
legendOnly	boolean variable specifying if only the legend has to be added to a plot.

Author(s)

Simon Dellicour

postTreeExtractions	*A function to extract the spatio-tremporal information embedded in annotated phylogenetic trees sampled from the posterior distribution of a Bayesian continuous phylogeographic inference*

Description

This function extracts the spatio-temporal information contained in the annotated phylogenetic trees inferred, e.g., by the continuous phylogeographic method implemented in BEAST (Lemey *et al.* 2010). To use this function the "latitute-longitude" continuous traits partition should have been named "location" while preparing the input XML file for the continuous phylogeographic analysis conducted with BEAST.

Usage

```
postTreeExtractions(post_tre, mostRecentSamplingDatum)
```

Arguments

post_tre   annotated phylogenetic tree read/imported by the "readAnnotatedNexus" function of the same package.

mostRecentSamplingDatum
                   most recent sampling date in a decimal format (e.g., 2007.4).

Value

The function returns an extraction table containing the spatio-tremporal information extracted from an annotated phylogenetic tree.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Lemey P, Rambaut A, Welch JJ, Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology & Evolution* 27: 1877-1885.

rasterSimulation	*A function to simulate an environmental raster*

Description

This function simulates an environmental raster based on a variogram model that can be preliminary estimated with the "variogramModel" function. After the simulation step, a n-score transformation is automatically performed on the simulated raster cell values in order to obtain a distribution of values matching the one observed on the original raster file.

Usage

```
rasterSimulation(rast, variogramModel)
```

Arguments

rast   raster template for the raster simulation.

variogramModel   estimated variogram model for the environmental variable.

Details

This function was called by a previous version of the "spreadFactor" function.

Value

The function returns a raster simulated under a specific variogram model.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Pybus OG (2016a). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016b). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

---

readAnnotatedNexus *Read annotated tree files in a Nexus format*

Description

This function reads annotated trees from files in a Nexus format and was previously available in the R package "OutbreakTools" (Jombart *et al.* 2014). Except for the annotations, this functions mimic the R package "ape" function "read.nexus". Annotations are ordered by edges, i.e. matching the `edge.length` slot of a `phylo` object.

Usage

```
readAnnotatedNexus(file, tree.names=NULL)
```

Arguments

`file`	a file name specified by either a variable of mode character, or a double-quoted string; if 'file = ""' (the default) then the tree is input on the keyboard, the entry being terminated with a blank line
`tree.names`	if there are several trees to be read, a vector of mode character that gives names to the individual trees; if 'NULL' (the default), the trees are named '"tree1"', '"tree2"', ...

Details

See the function "read.nexus" in the R package "ape" for a specification of Nexus formatted tree files. This function additionally extracts BEAST annotations for all branches/nodes in the trees and returns these annotations as lists of lists in the resulting "phylo" objects

Value

An object of class "phylo" with an additional slot called annotations. This slot is a list indexed by the nodes.

Author(s)

Marc Suchard

References

Jombart T, Aanensen DM, Baguelin M, Birrell P, Cauchemez S, Camacho A, Colijn C, Collins C, Cori A, Didelot X, Fraser C, Frost S, Hens N, Hugues J, Höhle M, Opatowski L, Rambaut A, Ratmann O, Soubeyrand S, Suchard MA, Wallinga J, Ypma R, Ferguson N (2014). OutbreakTools: A new platform for disease outbreak analysis using the R software. *Epidemics* 7: 28-34.

| simulatorRRW1 | *A function to simulate a relaxed random walk (RRW) diffusion process along a time-scaled phylogenetic tree* |

Description

This function simulates a relaxed random walk (RRW) diffusion process along the branches of a time-scaled phylogenetic tree using RRW parameter values previously estimated with the continuous phylogeographic model implemented in BEAST (Lemey *et al.* 2010).

Usage

```
simulatorRRW1(tree,
              rates,
              sigmas = c(0.1,0.1),
              cor = 0,
              envVariables = list(),
              mostRecentSamplingDatum,
              ancestPosition = c(0,0),
              reciprocalRates = TRUE,
              n1 = 100,
              n2 = 0,
              showingPlots = FALSE,
              newPlot = TRUE)
```

Arguments

tree	phylogenetic tree (object of class "phylo").
rates	vector of trait evolutionary rates estimated by BEAST and associated with each branch (vector elements have to be in the same order as their associated branch edges in the "tree" object).
sigmas	vector of scale parameters for the latitude and longitude (can be retrieved from the BEAST output; see the related tutorial for further details).
cor	observed correlation between latitude and longitude (can be retrieved from the BEAST output; see the related tutorial for further details).
envVariables	list of environmental rasters that will determine the accessible area (simulated node positions will not fall into raster cells with "NA" values).
mostRecentSamplingDatum	
	most recent sampling date in a decimal format (e.g., 2007.4).
ancestPosition	vector of geographic coordinates (longitude, latitude) of the most ancestral node position, i.e. the starting position of the RRW simulation.
reciprocalRates	
	BEAST parameter of the continuous diffusion model: boolean variable specifying if branch lengths have to be divided (TRUE) or not (FALSE) by the trait evolutionary rates.
n1	maximum number of branch rotations allowed per simulated node position. If a simulated node position falls into an inaccessible area (i.e. a raster cell with a "NA" value), the simulator can rotate the branch around its oldest node position and while maintaining the geographic distance travelled by the branch until

its youngest node position falls into an accessible area. In that situation, "n1" defines the maximum number of times that this rotation trial can be attempted. After "n1" rotation trials, the branch is not rotated anymore and the youngest node position (and hence its actual distance from the oldest node of the branch) is re-simulated if "n2" is higher than "0" (see below) or the entire RRW simulation is restarted from the root of the tree if "n2" is set to "0".

n2   maximum number of re-simulations allowed per node position. If a simulated node position falls into an inaccessible area (i.e. a raster cell with a "NA" value), the simulator will first attempt to rotate the branch around its oldest node position until its youngest node position falls into an accessible area. But after "n1" rotation trials, and if "n2" is higher than "0", it will re-simulate the youngest node position and hence its actual distance from the oldest node of the branch. In that situation, "n2" defines the maximum number of times that a new position is re-simulated (and then also rotated up to n1 times) before restarting the entire RRW simulation from the root of the tree.

showingPlots   boolean variable specifying if the different plots have to be displayed or not.

newPlot   boolean variable specifying if a new plot window has to be opened or if the simulated branches/nodes have to be plotted on a previously opened plot window, e.g., with a specific raster and/or polygons already mapped on it (graphical parameter that is only useful when "showingPlots=TRUE").

Value

The function simulates RRW diffusion process from a specified ancestral location and along a given phylogenetic tree. In practice, the function generates a table containing the spatio-tremporal information of the simulated RRW diffusion process. This extraction table is similar to those generated by the "treeExtractions", "postTreeExtractions" and "mccTreeExtractions" functions.

Author(s)

Simon Dellicour, Paul Bastide, Philippe Lemey

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Baele G, Dudas G, Faria NR, Pybus OG, Suchard M, Rambaut A, Lemey P (2018). Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak. *Nature Communications* 9: 2222.

Lemey P, Rambaut A, Welch JJ, Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology & Evolution* 27: 1877-1885.

---

simulatorRRW2   *A function to conduct phylogeographic simulations based on a birth-death process and a Brownian random walk (BRW) or a relaxed random walk (RRW) diffusion process*

Description

This function can be used to conduct forward-in-time joint simulations of both time-scaled phylo-
genies and the dispersal history of their branches on an underlying geo-referenced grid. At each
time step of the BRW simulation, both the longitudinal and latitudinal displacements of evolving
lineages are randomly drawn from a Gaussian distribution, while in the RRW simulations such
longitudinal and latitudinal displacements are randomly drawn from a Cauchy distribution before
randomly rotating the resulting displacement around its origin. In both cases, lineage dispersal is
prevented in inaccessible raster cells.

Usage

```
simulatorRRW2(RRW = TRUE,
              envVariable =
raster(matrix(c(runif(600,5,10),runif(1000,0,5)),nrow=40,ncol=40)),
              sigma = 1,
              ancestPosition = c(0.4,0.5),
              birthRate = 0.2,
              samplingRate = 0.2,
              startingYear = 0,
              samplingWindow = c(10,50),
              timeSlice = 0.1,
              timeIntervale = 1,
              showingPlots = FALSE)
```

Arguments

RRW   boolean variable specifying if a Bownian (BRW) or relaxed (RRW) random walk
                    diffusion process should be simulated.

envVariable   raster that will determine the environmental heterogeneity (cells with a "NA"
                    value will be considered as inaccessible).

sigma   standard deviation of the normal distribution with mean parameter set to zero in
                    which longitudinal and latitudinal displacement values are randomly picked (if
                    "RRW" is set to "FALSE", i.e. in the case of a BRW simulation), or the value
                    of the scale parameter of the Cauchy distribution with location parameter set
                    to zero in which longitudinal and latitudinal displacement values are randomly
                    picked (if "RRW" is set to "TRUE", i.e. in the case of a RRW simulation (see
                    also the note below).

ancestPosition   vector of geographic coordinates (longitude, latitude) of the most ancestral node
                    position, i.e. the starting position of the RRW diffusion simulation.

birthRate   value defining the rate at which lineages are splited in two new lineages (ex-
                    pressed in events per lineage per time unit).

samplingRate   value defining the rate at which lineage are sampled and thus stopped moving
                    and splitting on the raster (expressed in events per lineage per time unit).

startingYear   year (or moment in another time unit) at which the simulation begins from the
                    ancestral node position.

samplingWindow   vector of time values (in the same time unit as the "startingYear") defining the
                    starting and ending times of the sampling period, i.e. the period during which
                    the sampling rate will be effective.

timeSlice   time interval (in the same time unit as the "startingYear") at which unsampled
                    lineages perform a new movement on the raster.

timeIntervale   time interval (in the same time unit as the "startingYear") at which the plot displaying the simulation is updated (only taken into account when "showing-Plots=TRUE").

showingPlots   boolean variable specifying if the different plots have to be displayed or not.

Value

The function jointly simulates time-scaled phylogenies and the dispersal history of their branches either based on a BRW or RRW diffusion process and while preventing dispersal in inaccessible raster cells. In practice, the function generates a list of two objects: (i) a table containing the spatio-tremporal information of the simulated BRW/RRW diffusion, with one line per lineage/phylogenetic branch (such a table similar to those generated by the "treeExtractions" function), and (ii) a tree object of class "phylo" that thus summarizes the phylogenetic relationships among simulated lineages.

Note

At each time step, once the longitudinal and latitudinal displacements have been independently drawn in a normal (BRW simulation) or Cauchy distribution (RRW simulation), the resulting movement vector is then randomly rotated around the position of the evolving lineage at the previous time step to avoid that sporadic long-distance dispersal events solely occure along a longitudinal or latitudinal gradient (a rotation step that is thus mostly relevant for RRW simulations in which such long-distance dispersal events are more frequent due to the nature of the Cauchy distribution).

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Bastide P, Rocu P, Fargette D, Hardy OJ, Suchard MA, Guindon S, Lemey P (2024). How fast are viruses spreading in the wild? *PLoS Biology* 22: e3002914.

---

simulatorRRW3   *A function to conduct phylogeographic simulations based on a birth-death process and a relaxed random walk (RRW) diffusion process during which the dispersal velocity of lineages is impacted by an environmental raster*

Description

This function can be used to conduct forward-in-time joint simulations of both time-scaled phylogenies and the dispersal history of their branches on an underlying geo-referenced grid (raster) whose environmental values impact the dispersal velocity of lineages. At each time step of this relaxed random walk (RRW) simulation, both the longitudinal and latitudinal displacements of evolving lineages are randomly drawn from a Gaussian distribution whose standard deviation is proportional to the underneath raster cell value while preventing lineage dispersal in inaccessible raster cells.

Usage

```
simulatorRRW3(envVariable =
raster(matrix(c(runif(600,5,10),runif(1000,0,5)),nrow=40,ncol=40)),
                resistance = TRUE,
                scalingValue = 1,
                ancestPosition = c(0.4,0.5),
                birthRate = 0.2,
                samplingRate = 0.2,
                startingYear = 0,
                samplingWindow = c(10,50),
                timeSlice = 0.1,
                timeIntervale = 1,
                showingPlots = FALSE,
                extractionOfValuesOnMatrix = FALSE)
```

Arguments

envVariable	raster that will determine the environmental heterogeneity impacting the dispersal velocity of evolving lineages (cells with a "NA" value will be considered as inaccessible).
resistance	boolean variable specifying if the environmental raster has to be treated as a resistance or conductance factor. If treated as a resistance factor, raster cell values will be inverted (1/x) prior to the simulation.
scalingValue	value used to rescale the standard deviation of the two normal distributions with a mean set to zero in which longitudinal and latitudinal displacement values are randomly picked (see details below).
ancestPosition	vector of geographic coordinates (longitude, latitude) of the most ancestral node position, i.e. the starting position of the RRW diffusion simulation.
birthRate	value defining the rate at which lineages are splited in two new lineages (expressed in events per lineage per time unit).
samplingRate	value defining the rate at which lineage are sampled and thus stopped moving and splitting on the raster (expressed in events per lineage per time unit).
startingYear	year (or moment in another time unit) at which the simulation begins from the ancestral node position.
samplingWindow	vector of time values (in the same time unit as the "startingYear") defining the starting and ending times of the sampling period, i.e. the period during which the sampling rate will be effective.
timeSlice	time interval (in the same time unit as the "startingYear") at which unsampled lineages perform a new movement on the raster.
timeIntervale	time interval (in the same time unit as the "startingYear") at which the plot displaying the simulation is updated (only taken into account when "showingPlots=TRUE").
showingPlots	boolean variable specifying if the different plots have to be displayed or not.
extractionOfValuesOnMatrix	
	boolean variable specifying if the raster object as to be preliminary transformed into a matrix object. This operation allows a significantly faster simulation but has to be avoided if possible. Indeed, with this transformation, the raster and the ancestral position are projected on a flat grid and thus lose their initial projection during the whole simulation. This aspect can, in some cases, lead to

over-simplistic situations. Note that if this option is selected, coordinates of simulated nodes are re-projected on the initial raster at the end of the simulation.

Value

The function jointly simulates time-scaled phylogenies and the dispersal history of their branches based on a RRW diffusion process whose dispersal velocity is proportional (in the case of a conductance factor) or inversely proportional (in the case of a resistance raster) to the underneath environmental raster cell values. In practice, the function generates a list of two objects: (i) a table containing the spatio-tremporal information of the simulated RRW diffusion process, with one line per lineage/phylogenetic branch (such a table similar to those generated by the "treeExtractions" function), and (ii) a tree object of class "phylo" that thus summarizes the phylogenetic relationships among simulated lineages.

Note

In order to avoid null standard deviation values, all the raster cell values are preliminary increased by "1" (see the formulas below). Furthermore, if the environmental raster is treated as a resistance factor, all its cell values will be inverted (after having increased all the cell values by one). This operation allows using the same standard deviation formulas for both the cases where the environmental raster as to be treated as a conductance or a resistance factor.

For a given raster cell i with environmental v_i, standard deviations (SDx_i and SDy_i) used to define the two normal distributions with mean = 0, in which longitudinal d_x and latitudinal d_y displacement values are randomly picked, are computed as follows:

$$SDx_i = (v_i/v_{max}) * cell_{width} * scalingValue$$
$$SDy_i = (v_i/v_{max}) * cell_{height} * scalingValue$$

with v_{max} being the highest value recorded in the environmental raster. Note that $cell_{width}$ and $cell_{height}$ are taken from the raster properties/parameters. Final note: in order to obtain to simulate trees with a realistic sampling pattern through time, tip branches that have not been sampled at the end of simulation are removed from the simulated tree topology.

Author(s)

Simon Dellicour, Philippe Lemey

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, *et al.* (*in prep.*). Peaks and valleys of landscape phylogeography.

---

spreadFactors	*A function to investigate the impact of environmental factors on the dispersal dynamic of (viral) lineages*

Description

This function estimates and tests the correlation between environmental distance(s) and dispersal duration associated with each phylogenetic branch. The function performs three analytical steps: (i) for each environmental factor, the computation of the environmental distances associated with each phylogenetic branch (using the "straight-line path", "least-cost path" or "Circuitscape path" model, see the details below), (ii) the estimation of the correlation between the dispersal durations and the corresponding environmental distances using a linear regression approach, and (iii) an optional randomisation step to estimate the statistical support of the correlation metrics (Q). To generate null models, several randomisation procedures of the branches position are implemented (see below for further detail).

Usage

```
spreadFactors(localTreesDirectory = "",
              nberOfExtractionFiles = 1,
              envVariables=list(),
              pathModel = 2,
              resistances = list(),
              avgResistances = list(),
              fourCells = FALSE,
              nberOfRandomisations = 0,
              randomProcedure = 3,
              outputName = "",
              showingPlots = FALSE,
              nberOfCores = 1,
              OS = "Unix",
              juliaCSImplementation = FALSE)
```

Arguments

localTreesDirectory

> directory where the tree extraction matrices are saved (see the "treeExtractions" or "postTreeExtractions" function).

nberOfExtractionFiles

> number of annotated phylogenetic trees from which spatio-temporal information has been extracted.

envVariables   list of environmental rasters to test.

pathModel   path model to use for computing environmental distances: "1" (straight-line path model), "2" (least-cost path model, Dijkstra 1959) or "3" (Circuitscape path model, McRae 2006). When set to "0", the "spreadFactors" function will rather conduct an analysis of the association between environmental factors and the dispersal location (instead of the diffusion velocity) of lineages - this analysis is not described here; see the related tutorial ("impact on dispersal location") for further detail.

resistances   list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the rasters of the "envVariables" list have to be treated as resistance or conductance factors (only relevant for the least-cost and Circuitscape path models).

avgResistances   list of boolean variables corresponding to the "envVariables" list. These boolean variables specify if the raster cells have to be connected by average resistance (TRUE) or average conductance (FALSE). This option is only relevant for the Circuitscape path model.

fourCells boolean variable specifying if a given raster cell should be connected to either its four first-order (TRUE) or eight first- and second-order (FALSE). This option is only relevant for the least-cost and Circuitscape path models.

nberOfRandomisations

number of randomisation steps to perform in order to test the correlation between environmental distances and dispersal durations.

randomProcedure

randomisation procedure to use: "1" (using external tree randomisations previously performed with the "treesRandomisation" function), "2" (using external RRW simulations previously performed with the "simulatorRRW1" function), "3" (randomisations of nodes position while maintaining the branch lengths, the tree topology, and the location of the most ancestral node), "4" (randomisations of nodes position while maintaining the branch lengths and the location of the starting node of each branch), "5" (randomisations of node positions while maintaining the branch lengths and the location of the ending node of each branch), or "6" (randomisations of node positions while maintaining only the branch lengths).

outputName name (prefix) to give to the different output files.

showingPlots boolean variable specifying if a plot has to be displayed during the randomisation procedure.

nberOfCores number of available cores to parallelise the computations in R (parallelisation will only work on Unix operating systems).

nberOfCores_CS number of available cores to parallelise the computations in Circuitscape (we advise to set this parameter value to one and rather use the "numberOfCores" variable to set a parallelisation process).

OS operating system on which the function is used ("Unix" or "Windows").

juliaCSImplementation

boolean variable specifying if the Julia implementation of Circuitscape should be called by the function. By default (FALSE), the Python implementation is called.

Details

Three different path models are available to compute the environmental distances associated with each phylogenetic branch: (i) the straight-line path model only considering the raster cells below each phylogenetic branch. In this model, the environmental distance associated with a given branch for a given environmental variable will be the sum of the values of the raster cells below the straight-line segment connecting the starting and ending nodes of the branch. (ii) The least-cost path model (Dijkstra 1959). In this second model, environmental distance associated with a given branch for a given environmental variable will be the sum of the values of the raster cells selected by a shortest-path algorithm while considering the environmental variable as a resistance or a conductance factor. (iii) The Circuitscape model based on circuit theory (McRae 2006). This latter model integrates the contribution of multiple possible pathways to compute the environmental distance (McRae 2006). The environmental distances based on the least-cost path model are here computed with the "costDistance" function available in the "gdistance" R package (van Etten 2012) and the environmental distances based on the Circuitscape path model are computed using the Python or Julia package "circuitscape" (McRae 2006).

Regarding the test to estimate the statistical support of the correlation metrics *Q*, several randomisation procedures have been implemented to create null models: (i) using already performed tree

randomisations corresponding to each tree (see the "treesRandomisation" function for further detail), (ii) using already performed RRW simulations corresponding to each tree (see the "simulatorRRW1" function for further detail), (iii) a randomisation of the branches position while maintaining their length, the tree topology and the location of the most ancestral node, (iv) a randomisation of the branches position while maintaining their length and the location of the starting node of each branch, and (v) a randomisation of the branches position while maintaining only their length. The latter procedure can be useful, e.g., when analysing non phylogenetically informed movement data like capture-mark-recapture movement data.

Value

The function generates several output files. If the number of randomisations to perform is set to zero, the function will only generate a text file listing the results of the two linear regression (LR) analyses conducted for each phylogenetic tree: (i) the univariate linear regression between the dispersal durations (t) and the environmental distances (d) computed for a given environmental factor (for which the output files are labelled "LR1"), and (ii) the univariate linear regression between four times the dispersal durations ($4t$) and the squared environmental distances (d^2) computed for a given environmental factor (for which the output files are labelled "LR2"). While the first linear regression (LR1) aims to assess the impact of a given environmental factor on the heterogeneity of the lineage dispersal velocity ($t \sim d$), the second linear regression (LR2) aims to assess the impact of a given environmental factor on the heterogeneity of the diffusion velocity ($4t \sim d^2$), which is related to the diffusion coefficient identified as a metric more robust to the sampling intensity (i.e. the sampling size) than the lineage dispersal velocity (Dellicour *et al*. 2024). For both linear regressions, the function will thus generate a first text file gathering the following elements: the β regression coefficients and coefficient of determination R^2_{env} obtained from each linear regression, as well as the difference Q between R^2_{env} and R^2_{null} (the coefficient of determination obtained from the linear regression based on environmental distances computed on a "null" raster with uniform cell values equal to "1", the later distances corresponding to a proxy of the geographic distances). If the number of randomisations to perform is higher than zero, the function will return a additional text files reporting the Bayes factor support approximated for each environmental factor and each randomisation step (Dellicour *et al*. 2017; see also the related tutorial for detailed information about the Bayes factor estimations). Note that if the "pathModel" parameter is set to "0", the "spreadFactors" function will rather conduct association analyses between the lineage dispersal locations and local environmental distances (Dellicour *et al*. 2019) — those analyses and their outcomes are not detailed here but in a dedicated tutorial ("impact on lineage location").

Note

The computation of Circuitscape environmental distances requires the preliminary installation of the program Circuitscape as a Python or Julia package. Instructions on how to install this Python package are available on the Circuitscape website: https://circuitscape.org/

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics:

a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Dellicour S, Rose R, Faria NR, Vieira LFP, Bourhy H, Gilbert M, Lemey P, Pybus OG (2017). Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus epidemics. *Molecular Biology \& Evolution* 34: 2563-2571.

Dellicour S, Troupin C, Jahanbakhsh F, Salama A, Massoudi S, Moghaddam MK, Baele G, Lemey P, Gholami A, Bourhy H (2019). Using phylogeographic approaches to analyse the dispersal history, velocity, and direction of viral lineages – application to rabies virus spread in Iran. *Molecular Ecology* 28: 4335-4350.

Dellicour S, Bastide P, Rocu P, Fargette D, Hardy OJ, Suchard MA, Guindon S, Lemey P (2024). How fast are viruses spreading in the wild? *PLoS Biology* 22: e3002914.

Dellicour S, Gámbaro F, Jacquot M, Lequime S, Baele G, Gilbert M, Pybus OG, Suchard MA, Lemey P (2025). *Comparative performance of viral landscape phylogeography approaches. Proceedings of the National Academy of Sciences of the USA* 122: e2506743122.

Dijkstra EW (1959). A note on two problems in connexion with graphs. *Numerische Mathematik* 1: 269-271.

Faria NR, Suchard MA, Rambaut A, Streicker DG, Lemey P (2013). Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. *Philosophical Transactions of the Royal Society Series B* 368: 20120196.

McRae BH (2006). Isolation by resistance. *Evolution* 60: 1551-1561.

Van Etten J (2012). R package gdistance: distances and routes on geographical grids. R package version 1.1.

See Also

https://beast.community/
http://cran.r-project.org/web/packages/gdistance/index.html
http://www.circuitscape.org
https://github.com/Circuitscape/Circuitscape.jl

---

spreadGraphic1	*A function to generate a collection of higher posterior density (HPD) regions retrieved for each internal node from an annotated maximum clade credibility (MCC) tree and representing the uncertainty associated with the Bayesian continuous phylogeographic inference*

Description

This function returns a list of spatial polygon data frames (SPDFs), with one SPDF per internal node. Specifically, these SPDF correspond to the 80% HPD uncertainty regions associated with each internal node and are extracted from an annotated MCC tree, which was previously retrieved and annotated with the program TreeAnnotator (https://beast.community/treeannotator). In the resulting list of SPDFs, the name assigned to a SPDF is the time of occurrence of the associated

internal node. See also, for instance, the programs SPREAD 4 (Nahata *et al.* 2022) and spread.gl (Li *et al.* 2024) implementing similar approaches.

Usage

```
spreadGraphic1(mcc_tre,
               mcc_tab)
```

Arguments

mcc_tre   MCC tree read/imported by the "readAnnotatedNexus" function of the same package.

mcc_tab   extraction table containing the spatio-tremporal information extracted from the MCC tree using the "readAnnotatedNexus" function of the same package.

Value

The function returns a list of SPDFs, with one SPDF per internal node.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Li Y, Bollen N, Hong SL, Brusselmans M, Gambaro F, Suchard MA, Rambaut A, Lemey P, Dellicour S, Baele G (2024). Spread.gl: visualising pathogen dispersal in a high-performance browser application. *Bioinformatics* 40: btae721.

Nahata K, Bielejec F, Monetta J, Dellicour S, Rambaut A, Suchard MS, Lemey P (2022). SPREAD 4: online visualization of pathogen phylogeographic reconstructions. *Virus Evolution* 8: veac088.

---

spreadGraphic2	*A function to generate a collection of higher posterior density (HPD) regions computed by defined time slice and representing the uncertainty associated with the Bayesian continuous phylogeographic inference*

Description

This function generates a list of spatial polygon data frames (SPDFs), with one SPDF per considered time slice. Each SPDF is build by estimating the HPD (highest posterior density) region based on all the ending positions of phylogenetic branches whose ending time falls within the considered time slice. See also, for instance, the programs SPREAD 4 (Nahata *et al.* 2022) and spread.gl (Li *et al.* 2024) implementing similar approaches.

Usage

```
spreadGraphic2(localTreesDirectory,
               nberOfExtractionFiles,
               prob = 0.95,
               startDatum,
               precision = 1)
```

Arguments

localTreesDirectory

> directory where the tree extraction matrices are saved (see the "treeExtractions" or "postTreeExtractions" function).

nberOfExtractionFiles

> number of annotated phylogenetic trees from which spatio-temporal information has been extracted.

prob
> probability corresponding to the HPD (highest posterior density) regions.

startDatum
> number in a decimal format (e.g., 2007.4) defining the beginning of the dispersal history.

precision
> number, time interval that will be used to define the successive time slices.

Value

The function returns a list of SPDFs, with one SPDF per time slice.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Li Y, Bollen N, Hong SL, Brusselmans M, Gambaro F, Suchard MA, Rambaut A, Lemey P, Dellicour S, Baele G (2024). Spread.gl: visualising pathogen dispersal in a high-performance browser application. *Bioinformatics* 40: btae721.

Nahata K, Bielejec F, Monetta J, Dellicour S, Rambaut A, Suchard MS, Lemey P (2022). SPREAD 4: online visualization of pathogen phylogeographic reconstructions. *Virus Evolution* 8: veac088.

---

spreadStatistics   *A function to compute several lineage dispersal statistics*

Description

This function estimates several lineage dispersal statistics based on a continuous phylogeographic reconstruction

Usage

```
spreadStatistics(localTreesDirectory = ".",
                 nberOfExtractionFiles = 1,
                 timeSlices = 200,
                 onlyTipBranches = FALSE,
                 showingPlots = TRUE,
                 outputName = gsub(" ","_",date()),
                 nberOfCores = 1,
                 slidingWindow = 1)
```

Arguments

localTreesDirectory
:   directory where the tree extraction matrices are saved (see the "treeExtractions" or "postTreeExtractions" function).

nberOfExtractionFiles
:   number of annotated phylogenetic trees from which spatio-temporal information has been extracted.

timeSlices
:   number of distinct time slices that will be used to generate the maximal wavefront distance evolution plots.

onlyTipBranches
:   boolean variable defining if the statistic estimations have to be performed only while considering the tip branches of the trees.

showingPlots
:   boolean variable specifying if the different plots have to be displayed or not.

outputName
:   name (prefix) to give to the different output files.

nberOfCores
:   number of available cores to parallelise the computations (parallelisation will only work on Unix operating systems).

slidingWindow
:   sliding window, in units of time, that will be used to generate the dispersal velocity evolution plot (optional).

Value

The function generates several outputs. The first output is a text file gathering the estimates of several dispersal statistics, with one row corresponding to the estimates obtained for a distinct posterior tree: the mean and weighted branch dispersal velocities, the original and weighted diffusion coefficients, the isolation-by-distance (IBD) signal estimated by the correlation between the patristic and great-circle geographic distances computed for each pair of tip nodes, as well as branch velocity and diffusion coefficients variation among branches, estimated for each phylogenetic tree. See the related tutorial for the different formulas, as well as Pybus *et al.* (2012) for the formula used to estimate the original diffusion coefficient and Trovão *et al.* (2015) for the formula used to estimate the weighted diffusion coefficient. As for the IBD signal, the function will estimate three distinct metrics: the Spearman correlation (r_S) between the patristic and great-circle geographic distances computed for each pair of tip nodes, the Pearson correlation (r_P1) between the patristic and great-circle geographic distances computed for each pair of tip nodes, and the Pearson correlation (r_P2) between the patristic and the log-transformed great-circle geographic distances computed for each pair of tip nodes. The function also generates and automatically saves figures displaying the kernel density estimates of mean branch velocity, weighted dispersal velocity and diffusion coefficients parameters (coefficients of variation against mean values). Finally, if the parameter onlyTipBranches is set to "FALSE", the function will also generate and save graphs displaying the evolution of the maximal wavefront distance, as well as the evolution of the diffusion coefficient metrics through

time. Two types of wavefront distance are estimated: the maximal *spatial* wavefront distance corresponding to the straight-line distance from to the estimated location of the root, and the maximal *patristic* wavefront distance corresponds to the distance computed as the sum of geographical distances associated with each branch connecting a given node to the root. The maximal *patristic* wavefront distance is here defined as the *patristic* distance from the root to the node associated with the highest *spatial* distance from the root location at a given point in time. On these graphs, grey area corresponds to the 95% highest posterior density (HPD) interval of the estimated wavefront position. Note that geographic distances used to estimate the different statistics are great circle distances computed with the function "rdist.earth" from the R package "fields" (Nychka *et al*. 2014).

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Bastide P, Rocu P, Fargette D, Hardy OJ, Suchard MA, Guindon S, Lemey P (2024). How fast are viruses spreading in the wild? *PLoS Biology* 22: e3002914.

Nychka D, Furrer R, Sain S (2014). fields: Tools for spatial data. R package version 7.1.

Pybus OG, Suchard MA, Lemey P, Bernardin FJ, Rambaut A, Crawford FW, Gray RR, Arinaminpathy N, Stramer SL, Busch MP, Delwart EL (2012). Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *PNAS* 109: 15066-15071.

Trovão NS, Suchard MA, Baele G, Gilbert M, Lemey P (2015). Bayesian inference reveals host-specific contributions to the epidemic expansion of Influenza A H5N1. *Molecular Biology and Evolution* 32 (12): 3264-3275.

---

spreadValues	*A function to generate skyline-like plot reporting averaged environmental conditions/values crossed by phylogenetic branches*

Description

This function will extract, for each considered time slice, environmental values crossed by the branch segments occurring during this time slice. All the extracted values are then averaged and reported for each sampled tree and per time slice.

Usage

```
spreadValues(localTreesDirectory = ".",
             nberOfExtractionFiles = 1,
             envVariables = list(),
             startTime,
             endTime,
             timeSlices = 100,
             slidingWindow = 1/12,
```

```
                     showingPlots = TRUE,
                     outputName = gsub(" ","_",date()),
                     nberOfCores = 1)
```

Arguments

localTreesDirectory
:   directory where the tree extraction matrices are saved (see the "treeExtractions" or "postTreeExtractions" function).

nberOfExtractionFiles
:   number of annotated phylogenetic trees from which spatio-temporal information has been extracted.

envVariables
:   list of environmental rasters to analyse.

startTime
:   starting time of the plot (x-axis).

endTime
:   ending time of the plot (x-axis).

timeSlices
:   number of distinct time slices that will be used to generate the maximal wave-front distance evolution plots.

slidingWindow
:   sliding window, in units of time, that will be used to generate the environmental skyline-like plot.

showingPlots
:   boolean variable specifying if the different plots have to be displayed or not.

outputName
:   name (prefix) to give to the different output files.

nberOfCores
:   number of available cores to parallelise the computations (parallelisation will only work on Unix operating systems).

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Lequime S, Vrancken B, Gill MS, Bastide P, Gangavarapu K, Matteson N, Tan Y, du Plessis L, Fisher AA, Nelson MI, Gilbert M, Suchard MS, Andersen KG, Grubaugh ND, Pybus OG, Lemey P (2020). Epidemiological hypothesis testing using a phylogeographic and phylodynamic framework. *Nature Communications* 11: 5620.

---

torusRandomisation   *A function to perform the torus randomisation of an original raster*

Description

This function generates a randomised raster by performing random torus translation and reflection of an original raster (vertical and horizontal reflections both having a probability of occurrence equals to 0.5). Torus translation and reflection randomisation can be used to break down the association between variables while keeping their respective spatial autocorrelation pattern intact, as far as possible (Harms *et al.* 2001, Hardy 2009).

Usage

```
torusRandomisation(rast)
```

Arguments

rast   raster to randomise with a torus translation and reflection.

Details

This function was called by a previous version of the "spreadFactor" function.

Value

The function returns a raster object with cell values randomised by a torus translation and reflection.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Harms KE, Condit R, Hubbell SP, Foster RB (2001). Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *Journal of Ecology* 89: 947-959.

Hardy OJ (2009). TOROCOR: a program to assess the association between spatially autocorrelated variables using a torus-translation test on multiple grids.

---

treeExtractions   *A function to extract the spatio-tremporal information embedded in spatiotemporally-referenced phylogenetic trees*

Description

This function extracts the spatio-temporal information contained in the phylogenetic trees inferred, e.g., by the continuous phylogeographic method implemented in BEAST (Lemey *et al.* 2010). Depending on the settings of the BEAST analysis, this function can sometimes fail to properly extract information contained in posterior trees. We thus recommend to rather use the latest "postTreeExtractions" function also available in the package.

Usage

```
treeExtractions(localTreesDirectory,
                allTrees, burnIn,
                randomSampling,
                nberOfTreesToSample,
                mostRecentSamplingDatum,
                coordinateAttributeName,
                nberOfCores = 1)
```

Arguments

`localTreesDirectory`
　　　　　　　name of the (new) directory where the tree extraction matrices will be saved.

`allTrees`　　　all the trees in a Nexus format (read with the "scan" function).

`burnIn`　　　number of posterior trees to discard as burn-in.

`randomSampling`　boolean variable specifying if the trees have to be randomly sampled in the list. If FALSE, the trees will be sampled at a regular interval (at the largest possible interval).

`nberOfTreesToSample`
　　　　　　　number of trees to sample.

`mostRecentSamplingDatum`
　　　　　　　most recent sampling date in a decimal format (e.g., 2007.4).

`coordinateAttributeName`
　　　　　　　attribute name used to indicate the geographic coordinates within the trees file.

`nberOfCores`　number of available cores to parallelise the extractions (parallelisation will only work on Unix operating systems).

Value

The function generates "csv" files (one per phylogenetic tree) containing the spatio-tremporal information extracted from annotated phylogenetic trees.

Author(s)

Rebecca Rose, Simon Dellicour

References

Dellicour S, Rose R, Pybus OG (2016a). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016b). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Lemey P, Rambaut A, Welch JJ, Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology & Evolution* 27: 1877-1885.

---

treesRandomisation | *A function to randomise the position of pmhylogenetic tree branches*

Description

This function randomises the position of phylogenetic branches on the study area and is based on the "extraction" tables previously generated by the "treeExtractions" or "postTreeExtractions" function, which both summarise each phylogenetic branch as a movement vector. Four distinct randomisations have been implemented (see below) and, in the case of two of them, can be conducted while considering or not the influence of an environmental raster on the dispersal locations of lineages.

Usage

```
treesRandomisation(localTreesDirectory = "",
                   nberOfExtractionFiles = 1,
                   envVariable,
                   randomProcedure = 3,
                   repulsion = NULL,
                   overwrite = FALSE,
                   showingPlots = FALSE)
```

Arguments

localTreesDirectory

directory where the tree extraction matrices are saved (see the "treeExtractions" or "postTreeExtractions" function).

nberOfExtractionFiles

number of annotated phylogenetic trees from which spatio-temporal information has been extracted.

envVariable

environmental raster to consider and on which the tree branches randomisation will be conducted. If the "repulsion" parameter is set to "NULL", this raster is only used to prevent that randomised node positions fall in non-accessible areas, i.e. raster cells associated with a "NA" value (e.g., sea areas). If the "repulsion" parameter is set to "TRUE" or "FALSE", the environmental raster will influence the tree branch randomisation process, higher environmental values tending to repulse ("TRUE") or attract ("FALSE") the dispersal locations of lineages (see below for further detail).

randomProcedure

randomisation procedure to use: "3" (randomisations of nodes position while maintaining the branch lengths, the tree topology, and the location of the most ancestral node), "4" (randomisations of node positions while maintaining the branches length and the location of the starting node of each branch), "5" (randomisations of node positions while maintaining the branch lengths and the location of the ending node of each branch), or "6" (randomisations of nodes position while maintaining only the branch lengths).

repulsion

boolean variable specifying if the environmental raster should be treated as an environmental variable repulsing ("TRUE") or attracting ("FALSE") the dispersal location of the branches (see below for further detail), or if its environmental values should not be considered in the tree branches randomisation process ("NULL"). In the latter case, the environmental raster is only used to prevent

that randomised node positions fall in non-accessible areas, i.e. raster cells associated with a "NA" value (e.g., sea areas).

overwrite   boolean variable specifying if existing simulation outputs should be overwritten or not (if set to "TRUE" and if output files already exists for the simulation n°*i*, this simulation will not be conducted and the function will directly conduct simulation *i*+1).

showingPlots   boolean variable specifying if visualisations the different randomisations have to be displayed (and here automatically saved) or not.

Details

Several randomisation procedures of tree branches have been implemented: (i) a randomisation of the branch positions while maintaining their length, the tree topology, and the location of the most ancestral node (procedure n°3), (ii) a randomisation of the branch positions while maintaining their length and the location of the starting node of each branch (procedure n°4), (iii) a randomisation of the branch positions while maintaining their length and the location of the ending node of each branch (procedure n°5), and (iv) a randomisation of the branches position while maintaining only their length (procedure n°6). The latter procedure can be useful, e.g., when analysing non phylogenetically informed movement data like capture-mark-recapture movement data. Note that in each case, branch positions are independently randomised for each tree. Note that randomisation procedures n°1 (using already performed tree randomisations) and n°2 (sing already performed RRW simulations) are not implemented in this function and correspond to procedures solely implemented in the "spreadFactors" function.

If the "repulsion" parameter is set to "TRUE" or "FALSE", the environmental raster will influence the tree branch randomisation process, higher environmental values tending to repulse ("TRUE") or attract ("FALSE") the dispersal locations of lineages. In practice, each branch position will be randomised 100 times and the environmental values at all the potential randomised child (or "ending") node locations are recorded. Those environmental values V are then used to specify the probability p_j to randomly pick a potential randomised location j given the environmental value v_j at that location: $p_j = (max(V) - v_j)/(max(V) - min(V))$ if "repulsion" is set to "TRUE" (the environmental raster tending to attract the dispersal location of lineages) or $p_j = (v_j - min(v))/(max(V) - min(V))$ if "repulsion" is set to "FALSE" (the environmental raster tending to repulse the dispersal location of lineages). Important note: because only relevant in those cases, these repulsion/attraction randomisation settings are not available for the randomisation procedures n°3 and 4 (if the randomisation procedure n°5 or 6 is selected, the "repulsion" parameter will be set to "NULL" by the fonction).

Value

In practice, the function generates a table containing the spatio-tremporal information corresponding to each randomised tree, and which are similar to the tables generated by the "treeExtractions" or "postTreeExtractions" function.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

---

variogramModel	*A function to estimate variogram model(s)*

Description

Using the functions "vgm" and "fit.variogram" of the "gstat" package (Pebesma 2004) to estimate the best variogram model(s) for one or several raster files.

Usage

```
variogramModel(envVariables)
```

Arguments

envVariables   list of environmental rasters for which a variogram model has to be estimated.

Value

The function returns a list of variogram models, one per considered raster file.

Note

This function involves graphics generation and thus cannot be launched from a terminal window. It has to be called from a R interface.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016a). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Rose R, Pybus OG (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Pebesma EJ (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30: 683-691.

---

vectorRandomisation   *A function to randomise the end position of a set of movement vectors*

Description

This function randomises the position of a set of movement vectors by rotating them around their start location. The randomisation (i.e. rotation) is performed on a template raster and while avoiding that new start and/or new end position fall on a cell with a "NA" value. During the randomisation, vectors length (i.e. euclidian distance between start and end locations) is not modified.

Usage

```
vectorRandomisation(rast, fromCoor, toCoor, showingPlots = FALSE)
```

Arguments

rast	template raster on which the position of vectors will be randomised.
fromCoor	matrix containing the coordinates of the start position of each movement vector (one line per vector, 1st column with the longitude and the 2nd column with the latitude).
toCoor	matrix containing the coordinates of the end position of each movement vector (one line per vector, 1st column with the longitude and the 2nd column with the latitude).
showingPlots	boolean variable specifying if the plot displaying the randomised vectors has to be displayed or not.

Details

This function can be used to randomise the movement vectors obtained, e.g., by CMR (capture-mark-recapture) or GPS data.

Value

The function returns a matrix of randomised vector coordinates with one line per vector: new starting longitude, new starting latitude, new ending longitude and new ending latitude.

Author(s)

Simon Dellicour

References

Dellicour S, Rose R, Pybus OG (2016a). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* 17: 82.

Dellicour S, Rose R, Faria N, Lemey P, Pybus OG (2016b). SERAPHIM: studying environmental rasters and phylogenetically-informed movements. *Bioinformatics* 32 (20): 3204-3206.

Dellicour S, Prunier JG, Piry S, Eloy MC, Bertouille S, Licoppe A, Frantz AC, Flamand MC (2019).

Landscape genetic analyses of Cervus elaphus and Sus scrofa: comparative study and analytical developments. *Heredity* 123: 228-241.

Dellicour S, Desmecht D, Paternostre J, Malengreaux C, Licoppe A, Gilbert M, Linden A (2020). Unravelling the dispersal dynamics and drivers of the African swine fever outbreak in Belgium. *Journal of Applied Ecology* 57: 1619-1629.

Index