Package 'DAISIE'

April 13, 2017

Type Package

Title Dynamical Assembly of Islands by Speciation, Immigration and Extinction

Version 1.4

Date 2017-04-12

Depends R (>= 3.0.0)

Imports deSolve, grDevices, graphics, stats, utils, DDD (> 3.0), subplex, Matrix, tensor

Suggests testthat

Author Rampal S. Etienne, Luis M. Valente, Albert B. Phillimore & Bart Haegeman

Maintainer Rampal S. Etienne <r.s.etienne@rug.nl>

License GPL-2

Description Simulates and computes the (maximum) likelihood of a dynamical model of island biota assembly through speciation, immigration and extinction. See Valente et al. 2015. Ecology Letters 18: 844-852, <DOI:10.1111/ele.12461>.

NeedsCompilation no

Repository CRAN

Date/Publication 2017-04-13 17:50:28 UTC

R topics documented:

DAISIE-package	*Dynamical Assembly of Islands by Speciation, Immigration and Extinction*

Description

DAISIE is an island biogeography model that assumes island biota assembly is governed by immigration, extinction and speciation through cladogenesis and anagenesis. This likelihood-based statistical package can simulate islands and estimate parameters of the DAISIE model based on phylogenetic/phylogeographic data. Cladogenesis and immigration rates can be dependent on diversity.

Details

Package:	DAISIE
Type:	Package
Version:	1.4
Date:	2017-04-12
License:	GPL 2.0

DAISIE_ML computes the parameters that maximizes the loglikelihood computed by DAISIE_loglik_all for one island/archipelago

DAISIE_ML2 computes the parameters that maximizes the loglikelihood for a set of islands / archipelagoes

DAISIE_sim simulates islands with given parameters

DAISIE_tutorial opens a step-by-step tutorial with example from Galapagos islands

DAISIE_loglik_all computes the loglikelihood of a given parameterization of the DAISIE model for a data set of colonization and branching times

DAISIE_plot_sims produces species-through-time plots based on simulated islands

DAISIE_dataprep converts data in table format to the list format accepted by DAISIE_loglik_all and DAISIE_ML

DAISIE_ExpEIN computes the expected number of endemics, non-endemics and their sum for a given time and set of parameters

DAISIE_probdist calculates the joint distribution of the number of endemics and non-endemics for a given set of parameter values, a given mainland species pool size and a given set of times

DAISIE_margprobdist calculates the marginal distribution of the number of endemics and non-endemics and their sum for a given set of parameter values, a given mainland species pool size and a given set of times

DAISIE_numcol calculates expectation and marginal distribution of the number of colonizations (lineages) for a given set of parameter values, a given mainland species pool size and a given set of times

Author(s)

Rampal S. Etienne, Luis M. Valente, Albert B. Phillimore & Bart Haegeman
Maintainer: Rampal S. Etienne <r.s.etienne@rug.nl>rep(0:(D1 - 1), each = D2)

References

Valente, L., A.B. Phillimore & R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852, <DOI:10.1111/ele.12461>.

Valente, L., R.S. Etienne & L. Davalos (2017). Recent extinctions disturb path to equilibrium diversity in Caribbean bats. Nature Ecology and Evolution 1: 0026. <DOI:10.1038/s41559-016-0026>.

See Also

DDD

DAISIE_convertprobdist

Converts the joint distribution of endemics and non-endemics under the DAISIE model to list format

---

Description

This function converts the joint distribution of the number of endemics and non-endemics from the matrix format of DAISIE_probdist to a list format

Usage

```
DAISIE_convertprobdist(
    pb
    )
```

Arguments

pb   Probability distribution in matrix format as output by DAISIE_probdist

Value

A list of length nrow(pb) containing matrices of square dimensions of size sqrt(ncol - 1) containing the joint probabilities with endemics in the rows and non-endemics in the columns. The last element of the list is a vector a times at which the joint probability distribution is evaluated.

Author(s)

Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Examples

```
### Compute the probability distribution at t = 4 and t = 8, for a mainland pool
# size of 250 potential colonists and a vector of 5 parameters (cladogenesis, extinction,
# clade-level carrying capacity, immigration, anagenesis) starting from an empty
# island; output in list format

pb = DAISIE_probdist(
   pars1 = c(0.3,0.35,Inf,0.75,0.012),
   pars2 = c(100,250),
   tvec = c(4,8),
   initEI = c(0,0),
   initprobs = NULL
   )
DAISIE_convertprobdist(pb)
```

DAISIE_dataprep   *Prepare colonisation and branching time data to run in DAISIE.*

Description

This function produces a data object that can be run in DAISIE likelihood computation/optimization functions. The function converts a user-specified table to a DAISIE-compatible format. See Galapagos_datatable.Rdata for a template of an input table.)

Usage

```
DAISIE_dataprep(
datatable,
island_age,
M,
number_clade_types = 1,
list_type2_clades = NA,
prop_type2_pool = "proportional",
```

```
epss = 1E-5
)
```

Arguments

datatable   Data frame (table) with user-specified data. See file Galapagos_datatable.Rdata
                         for a template of an input table. Each row on the table represents and indepen-
                         dent colonisation event. Table has the following four columns.

                         `$Clade_name` - name of independent colonization event
                         `$Status` - One of the following categories:
                         * "Non_endemic": applies to non-endemic species for cases where both island
                         and non-island populations of the species have been sampled)
                         * "Non_endemic_MaxAge": applies to non-endemic species for cases where is-
                         land individuals of the species have not been sampled and only the age of the
                         species is available)
                         * "Endemic": applies to endemic species and is applicable for both cladogenetic
                         and anagenetic species
                         * "Endemic_MaxAge": applies to endemic species for cases where island indi-
                         viduals of the species have not been sampled and only the age of the species is
                         available. This could apply to endemic species that have recently gone extinct
                         because of antropogenic causes that are (evidently) not modelled, and for which
                         no DNA data is available.
                         * "Endemic&Non_Endemic": when endemic clade is present and its mainland
                         ancestor has re-colonized
                         `$Missing_species` - Number of island species that were not sampled for par-
                         ticular clade (only applicable for "Endemic" clades)
                         `$Branching_times` - Stem age of the population/species in the case of "Non-
                         endemic", "Non-endemic_MaxAge" and "Endemic" anagenetic species. For
                         "Endemic" cladogenetic species these should be branching times of the radia-
                         tion including the stem age of the radiation.

island_age   Age of island in appropriate units

M   The size of the mainland pool, i.e the number of species that can potentially
                         colonize the island

number_clade_types
                         Number of clade types. Default: number_clade_types = 1 all species are con-
                         sidered to belong to same macroevolutionary process. If number_clade_types =
                         2, there are two types of clades with distinct macroevolutionary processes.

list_type2_clades
                         If number_clade_types = 2, list_type2_clades specifies the names of the clades
                         that have a distinct macroevolutionary process. The names must match those
                         in the $Clade_name column of the source data table (e.g. list_type2_clades =
                         "Finches"). If number_clade_types = 1, then list_type2_clades = NA should be
                         specified (default)

prop_type2_pool
                         Specifies the fraction of potential mainland colonists that have a distinct macroevo-
                         lutionary process. Applies only if number_clade_types = 2. Default "propor-

tional" sets the fraction to be proportional to the number of clades of distinct macroevolutionary process that have colonised the island. Alternatively, the user can specify a value between 0 and 1 (e.g. if mainland pool size is 1000 and prop_type2_pool = 0.02 then number of type2 species is 20).

epss   Default= 1E-5 should be appropriate in most cases. This value is used to set the maximum age of colonisation of "Non_endemic_MaxAge" and "Endemic_MaxAge" species to an age that is slightly younger than the island for cases when the age provided for that species is older than the island. The new maximum age is then used as an upper bound to integrate over all possible colonisation times.

Details

The output is an R list containing the data formatted to be run on other DAISIE functions.

Value

datalist   R list object containing data:
                  The first element of the list has two or three components:
                  $island_age - the island age
                  Then, depending on whether a distinction between species types is made, we have:
                  $not_present - the number of mainland lineages that are not present on the island
                  or:
                  $not_present_type1 - the number of mainland lineages of type 1 that are not present on the island
                  $not_present_type2 - the number of mainland lineages of type 2 that are not present on the island
                  The following elements of the list each contains information on a single colonist lineage on the island and has 5 components:
                  $colonist_name - the name of the species or clade that colonized the island
                  $branching_times - island age and stem age of the population/species in the case of "Non-endemic", "Non-endemic_MaxAge" and "Endemic" anagenetic species. For "Endemic" cladogenetic species these are island age and branching times of the radiation including the stem age of the radiation.
                  $stac - the status of the colonist
                  * Non_endemic_MaxAge: 1
                  * Endemic: 2
                  * Endemic&Non_Endemic: 3
                  * Non_endemic: 4
                  * Endemic_MaxAge: 5
                  $missing_species - number of island species that were not sampled for particular clade (only applicable for endemic clades)
                  $type_1or2 - whether the colonist belongs to type 1 or type 2

Author(s)

Luis M Valente

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Examples

```
### Create Galapagos data object where all taxa have the same macroevolutionary process

data(Galapagos_datatable)
DAISIE_dataprep(
   datatable = Galapagos_datatable,
   island_age = 4,
   M = 1000
   )

### Create Galapagos data object with a distinct macroevolutionary processes
# for the Darwin's finches. One process applies to type 1 species (all species
# except for Darwin's finches) and the other applies only to type 2 species
# (Darwin's finches). Set fraction of potential colonists of type 2 to be
# proportional to the number of type2 clades present on the island.

data(Galapagos_datatable)
DAISIE_dataprep(
   datatable = Galapagos_datatable,
   island_age = 4,
   M = 1000,
   number_clade_types = 2,
   list_type2_clades = "Finches"
   )

### Create Galapagos data object with a distinct macroevolutionary processes
# for the Darwin's finches. One process applies to type 1 species (all species
# except for Darwin's finches) and the other applies only to type 2 species
# (Darwin's finches). Set fraction of potential colonists of type 2 to be 0.163.

data(Galapagos_datatable)
DAISIE_dataprep(
   datatable = Galapagos_datatable,
   island_age = 4,
   M = 1000,
   number_clade_types = 2,
   list_type2_clades = "Finches",
   prop_type2_pool = 0.163
   )
```

DAISIE_ExpEIN   *The expected number of endemics and non-endemics under the DAISIE model*

Description

This function calculates the expected number of endemics, non-endemics and the sum of these for a given set of parameter values, a given mainland species pool size and a given time

Usage

```
DAISIE_ExpEIN(
    t,
    pars,
    M,
    initEI = c(0,0)
    )
```

Arguments

t	The time at which the expectations need to be computed
pars	Vector of parameters:
	pars[1] corresponds to lambda^c (cladogenesis rate)
	pars[2] corresponds to mu (extinction rate)
	pars[3] corresponds to K (clade-level carrying capacity)
	pars[4] corresponds to gamma (immigration rate)
	pars[5] corresponds to lambda^a (anagenesis rate)
M	The size of the mainland pool, i.e the number of species that can potentially colonize the island
initEI	The initial values of the number of endemics and non-endemics

Value

out	The output is a list with three elements:
	ExpE The number of endemic species
	ExpI The number of non-endemic species
	ExpN The sum of the number of endemics and non-endemics

Author(s)

Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Examples

```
### Compute the expected values at t = 4, for a mainland pool size of 1000 potential
# colonists and a vector of 5 parameters (cladogenesis, extinction, clade-level carrying
# capacity, immigration, anagenesis)

DAISIE_ExpEIN(
   t = 4,
   pars = c(0.5,0.1,Inf,0.01,0.4),
   M = 1000
   )
```

DAISIE_loglik_all   *Computes the loglikelihood of the DAISIE model given data and a set*
                          of model parameters

Description

Computes the loglikelihood of the DAISIE model given colonization and branching times for lineages on an island, and a set of model parameters

Usage

```
DAISIE_loglik_all(
   pars1,
   pars2,
   datalist,
   methode = "lsodes"
   )
```

Arguments

pars1   Contains the model parameters:

                  pars1[1] corresponds to lambda^c (cladogenesis rate)
                  pars1[2] corresponds to mu (extinction rate)
                  pars1[3] corresponds to K (clade-level carrying capacity)
                  pars1[4] corresponds to gamma (immigration rate)
                  pars1[5] corresponds to lambda^a (anagenesis rate)
                  pars1[6] corresponds to lambda^c (cladogenesis rate) for an optional subset of
                  the species
                  pars1[7] corresponds to mu (extinction rate) for an optional subset of the
                  species
                  pars1[8] corresponds to K (clade-level carrying capacity) for an optional sub-
                  set of the species
                  pars1[9] corresponds to gamma (immigration rate) for an optional subset of
                  the species
                  pars1[10] corresponds to lambda^a (anagenesis rate) for an optional subset of

the species

pars1[11] corresponds to p_f (fraction of mainland species that belongs to the second subset of species

The elements 6:10 and 11 are optional, that is, pars1 should either contain 5, 10 or 11 elements. If 10, then the fraction of potential colonists of type 2 is computed from the data. If 11, then pars1[11] is used, overruling any information in the data.

pars2   Contains the model settings

pars2[1] corresponds to lx = length of ODE variable x
pars2[2] corresponds to ddmodel = diversity-dependent model, model of diversity-dependence, which can be one of

ddmodel = 0 : no diversity dependence
ddmodel = 1 : linear dependence in speciation rate
ddmodel = 11: linear dependence in speciation rate and in immigration rate
ddmodel = 2 : exponential dependence in speciation rate
ddmodel = 21: exponential dependence in speciation rate and in immigration rate

pars2[3] corresponds to cond = setting of conditioning

cond = 0 : conditioning on island age
cond = 1 : conditioning on island age and non-extinction of the island biota

pars2[4] sets whether parameters and likelihood should be printed (1) or not (0)

datalist   Data object containing information on colonisation and branching times. This object can be generated using the DAISIE_dataprep function, which converts a user-specified data table into a data object, but the object can of course also be entered directly. It is an R list object with the following elements.
The first element of the list has two or three components:

$island_age - the island age
Then, depending on whether a distinction between types is made, we have:
$not_present - the number of mainland lineages that are not present on the island
or:
$not_present_type1 - the number of mainland lineages of type 1 that are not present on the island
$not_present_type2 - the number of mainland lineages of type 2 that are not present on the island

The remaining elements of the list each contains information on a single colonist lineage on the island and has 5 components:

$colonist_name - the name of the species or clade that colonized the island
$branching_times - island age and stem age of the population/species in the

case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be island age and branching times of the radiation including the stem age of the radiation.
$stac - the status of the colonist

* Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4
* Endemic_MaxAge: 5

$missing_species - number of island species that were not sampled for particular clade (only applicable for endemic clades)
$type1or2 - whether the colonist belongs to type 1 or type 2

methode   Method of the ODE-solver. See package deSolve for details. Default is "lsodes"

Details

The output is a loglikelihood value

Value

The loglikelihood

Author(s)

Rampal S. Etienne & Bart Haegeman

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_ML, DAISIE_sim

Examples

```
data(Galapagos_datalist_2types)
pars1 = c(0.195442017,0.087959583,Inf,0.002247364,0.873605049,
          3755.202241,8.909285094,14.99999923,0.002247364,0.873605049,0.163)
pars2 = c(100,11,0,1)
DAISIE_loglik_all(pars1,pars2,Galapagos_datalist_2types)
```

DAISIE_margprobdist   *The marginal distribution of endemics and non-endemics under the*
                              DAISIE model

Description

This function calculates the marginal distribution of the number of endemics and non-endemics and their sum for a given set of parameter values, a given mainland species pool size and a given set of times

Usage

```
DAISIE_margprobdist(
    pars1,
    pars2,
    tvec,
    initEI = c(0,0),
    initprobs = NULL,
    pb = NULL
    )
```

Arguments

pars1   Vector of model parameters:

                 pars1[1] corresponds to lambda^c (cladogenesis rate)
                 pars1[2] corresponds to mu (extinction rate)
                 pars1[3] corresponds to K (clade-level carrying capacity)
                 pars1[4] corresponds to gamma (immigration rate)
                 pars1[5] corresponds to lambda^a (anagenesis rate)

pars2   Vector of settings:

                 pars2[1] corresponds to res, the maximum number of endemics or non-endemics
                 for which the ODE system is solved; this must be much larger than the actual
                 number for which the probability needs to be calculated.)
                 pars2[2] corresponds to M, size of the mainland pool, i.e the number of species
                 that can potentially colonize the island.

tvec   The times at which the probabilities need to be computed.

initEI   The initial values for the number of endemics and non-endemics; either this or
                 initprobs must be NULL

initprobs   The initial probability distribution for the number of endemics and non-endemics;
                 either this or initEI must be NULL

pb   Rather than computing the joint distribution from given parameter values, one
                 can also specify a precomputed probability distribution in the matrix format of
                 DAISIE_probdist.

Value

out   A list of three vectors:

                       pE The probability distribution of the number of endemic species
                       pI The probability distribution of the number of non-endemic species
                       pN The probability distribution of the sum of the number of endemics and non-endemics

Author(s)

Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Examples

```
### Compute the marginal probability distributions at t = 4 and t = 8, for a mainland
# pool size of 250 potential colonists and a vector of 5 parameters (cladogenesis,
# extinction, clade-level carrying capacity, immigration, anagenesis) starting from
# an empty island

DAISIE_margprobdist(
   pars1 = c(0.3,0.35,Inf,0.75,0.012),
   pars2 = c(100,250),
   tvec = c(4,8),
   initEI = c(5,1),
   initprobs = NULL
   )
```

DAISIE_ML   *Maximization of the loglikelihood under the DAISIE model*

Description

This function computes the maximum likelihood estimates of the parameters of the DAISIE model for data from lineages colonizing an island. It also outputs the corresponding loglikelihood that can be used in model comparisons.

Usage

```
DAISIE_ML(
   datalist,
   datatype = 'single',
   initparsopt,
   idparsopt,
```

```
parsfix,
idparsfix,
idparsnoshift = 6:10,
idparsmat = NULL,
res = 100,
ddmodel = 0,
cond = 0,
eqmodel = 0,
x_E = 0.95,
x_I = 0.98,
tol = c(1e-04, 1e-05, 1e-07),
maxiter = 1000 * round((1.25)^length(idparsopt)),
methode = 'lsodes',
optimmethod = 'subplex'
)
```

Arguments

datalist   Data object containing information on colonisation and branching times. This
                 object can be generated using the DAISIE_dataprep function, which converts a
                 user-specified data table into a data object, but the object can of course also be
                 entered directly. It is an R list object with the following elements.
                 The first element of the list has two three components:

                 $island_age - the island age
                 Then, depending on whether a distinction between types is made, we have:
                 $not_present - the number of mainland lineages that are not present on the
                 island
                 or:
                 $not_present_type1 - the number of mainland lineages of type 1 that are not
                 present on the island
                 $not_present_type2 - the number of mainland lineages of type 2 that are not
                 present on the island

                 The remaining elements of the list each contains information on a single colonist
                 lineage on the island and has 5 components:

                 $colonist_name - the name of the species or clade that colonized the island
                 $branching_times - island age and stem age of the population/species in the
                 case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species.
                 For cladogenetic species these should be island age and branching times of the
                 radiation including the stem age of the radiation.
                 $stac - the status of the colonist

                 * Non_endemic_MaxAge: 1
                 * Endemic: 2
                 * Endemic&Non_Endemic: 3
                 * Non_endemic: 4

	$missing_species - number of island species that were not sampled for particular clade (only applicable for endemic clades) $type1or2 - whether the colonist belongs to type 1 or type 2
datatype	Sets the type of data: 'single' for a single island or archipelago treated as one, and 'multiple' for multiple archipelagoes potentially sharing the same parameters
initparsopt	The initial values of the parameters that must be optimized
idparsopt	The ids of the parameters that must be optimized. The ids are defined as follows: id = 1 corresponds to lambda^c (cladogenesis rate) id = 2 corresponds to mu (extinction rate) id = 3 corresponds to K (clade-level carrying capacity) id = 4 corresponds to gamma (immigration rate) id = 5 corresponds to lambda^a (anagenesis rate) id = 6 corresponds to lambda^c (cladogenesis rate) for an optional subset of the species id = 7 corresponds to mu (extinction rate) for an optional subset of the species id = 8 corresponds to K (clade-level carrying capacity) for an optional subset of the species id = 9 corresponds to gamma (immigration rate) for an optional subset of the species id = 10 corresponds to lambda^a (anagenesis rate) for an optional subset of the species id = 11 corresponds to p_f (fraction of mainland species that belongs to the second subset of species
idparsfix	The ids of the parameters that should not be optimized, e.g. c(1,3) if lambda^c and K should not be optimized.
parsfix	The values of the parameters that should not be optimized
idparsnoshift	For datatype = 'single' only: The ids of the parameters that should not be different between two groups of species; This can only apply to ids 6:10, e.g. idparsnoshift = c(6,7) means that lambda^c and mu have the same values for both groups
idparsmat	For datatype = 'multiple' only: Matrix containing the ids of the parameters, linking them to initparsopt and parsfix. Per island system we use the following order: * lac = (initial) cladogenesis rate * mu = extinction rate * K = maximum number of species possible in the clade * gam = (initial) immigration rate * laa = (initial) anagenesis rate Example: idparsmat = rbind(c(1,2,3,4,5),c(1,2,3,6,7)) has different rates of immigration and anagenesis for the two islands.
res	Sets the maximum number of species for which a probability must be computed, must be larger than the size of the largest clade

ddmodel	Sets the model of diversity-dependence:
	ddmodel = 0 : no diversity dependence ddmodel = 1 : linear dependence in speciation rate ddmodel = 11: linear dependence in speciation rate and in immigration rate ddmodel = 2 : exponential dependence in speciation rate ddmodel = 21: exponential dependence in speciation rate and in immigration rate
cond	cond = 0 : conditioning on island age cond = 1 : conditioning on island age and non-extinction of the island biota
eqmodel	Sets the equilibrium constraint that can be used during the likelihood optimization. Only available for datatype = 'single'. eqmodel = 0 : no equilibrium is assumed eqmodel = 13 : near-equilibrium is assumed on endemics using deterministic equation for endemics and immigrants. Endemics must be within x_E of the equilibrium value eqmodel = 15 : near-equilibrium is assumed on endemics and immigrants using deterministic equation for endemics and immigrants. Endemics must be within x_E of the equilibrium value, while non-endemics must be within x_I of the equilibrium value
x_E	Sets the fraction of the equlibrium endemic diversity above which the endemics are assumed to be in equilibrium; only active for eqmodel = 13 or 15
x_I	Sets the fraction of the equlibrium non-endemic diversity above which the system is assumed to be in equilibrium; only active for eqmodel = 15
tol	Sets the tolerances in the optimization. Consists of: reltolx = relative tolerance of parameter values in optimization reltolf = relative tolerance of function value in optimization abstolx = absolute tolerance of parameter values in optimization
maxiter	Sets the maximum number of iterations in the optimization
methode	Method of the ODE-solver. See package deSolve for details. Default is "lsodes"
optimmethod	Method used in likelihood optimization. Default is "subplex" (see subplex package). Alternative is 'simplex' which was the method in previous versions.

Details

The result of sort(c(idparsopt, idparsfix, idparsnoshift)) should be identical to c(1:10). If not, an error is reported that the input is incoherent. The same happens when the length of initparsopt is different from the length of idparsopt, and the length of parsfix is different from the length of idparsfix.

Including the 11th parameter (p_f) in either idparsopt or idparsfix (and therefore initparsopt or parsfix) is optional. If this parameter is not specified, then the information in the data is used, otherwise the information in the data is overruled.

Value

The output is a dataframe containing estimated parameters and maximum loglikelihood.

`lambda_c`	gives the maximum likelihood estimate of lambda^c, the rate of cladogenesis
`mu`	gives the maximum likelihood estimate of mu, the extinction rate
`K`	gives the maximum likelihood estimate of K, the carrying-capacity
`gamma`	gives the maximum likelihood estimate of gamma, the immigration rate
`lambda_a`	gives the maximum likelihood estimate of lambda^a, the rate of anagenesis
`lambda_c2`	gives the maximum likelihood estimate of lambda^c2, the rate of cladogenesis for the optional second group of species
`mu2`	gives the maximum likelihood estimate of mu2, the extinction rate for the optional second group of species
`K2`	gives the maximum likelihood estimate of K2, the carrying-capacity for the optional second group of species
`gamma2`	gives the maximum likelihood estimate of gamma2, the immigration rate for the optional second group of species
`lambda_a2`	gives the maximum likelihood estimate of lambda^a2, the rate of anagenesis for the optional second group of species
`loglik`	gives the maximum loglikelihood
`df`	gives the number of estimated parameters, i.e. degrees of feedom
`conv`	gives a message on convergence of optimization; conv = 0 means convergence

Author(s)

Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852. <DOI:10.1111/ele.12461>.

See Also

DAISIE_loglik_all, DAISIE_sim

Examples

```
cat("
### When all species have the same rates, and we want to optimize all 5 parameters,
# we use:

data(Galapagos_datalist)
DAISIE_ML(
   datalist = Galapagos_datalist,
   initparsopt = c(2.5,2.7,20,0.009,1.01),
   ddmodel = 11,
```

```
   idparsopt = 1:5,
   parsfix = NULL,
   idparsfix = NULL
)

### When all species have the same rates, and we want to optimize all parameters
# except K (which we set equal to Inf), we use:

data(Galapagos_datalist)
DAISIE_ML(
   datalist = Galapagos_datalist,
   initparsopt = c(2.5,2.7,0.009,1.01),
   idparsopt = c(1,2,4,5),
   parsfix = Inf,
   idparsfix = 3
   )

### When all species have the same rates except that the finches have a different
# rate of cladogenesis, and we want to optimize all parameters except K (which we
# set equal to Inf), fixing the proportion of finch-type species at 0.163, we use:

data(Galapagos_datalist_2types)
DAISIE_ML(
   datalist = Galapagos_datalist_2types,
   initparsopt = c(0.38,0.55,0.004,1.1,2.28),
   idparsopt = c(1,2,4,5,6),
   parsfix = c(Inf,Inf,0.163),
   idparsfix = c(3,8,11),
   idparsnoshift = c(7,9,10)
   )

### When all species have the same rates except that the finches have a different
# rate of cladogenesis, extinction and a different K, and we want to optimize all
# parameters, fixing the proportion of finch-type species at 0.163, we use:

data(Galapagos_datalist_2types)
DAISIE_ML(
   datalist = Galapagos_datalist_2types,
   ddmodel = 11,
   initparsopt = c(0.19,0.09,0.002,0.87,20,8.9,15),
   idparsopt = c(1,2,4,5,6,7,8),
   parsfix = c(Inf,0.163),
   idparsfix = c(3,11),
   idparsnoshift = c(9,10)
   )


### When all species have the same rates except that the finches have a different
# rate of extinction, and we want to optimize all parameters except K (which we
# set equal to Inf), and we also# want to estimate the fraction of finch species
# in the mainland pool. we use:

data(Galapagos_datalist_2types)
```

```
DAISIE_ML(
   datalist = Galapagos_datalist_2types,
   initparsopt = c(2.48,2.7,0.009,1.01,2.25,0.163),
   idparsopt = c(1,2,4,5,7,11),
   parsfix = c(Inf,Inf),
   idparsfix = c(3,8),
   idparsnoshift = c(6,9,10)
   )

### When we have two islands with the same rates except for immigration and anagenesis rate,
# and we want to optimize all parameters, we use:

data(Galapagos_datalist)
DAISIE_ML(
   datalist = list(Galapagos_datalist,Galapagos_datalist),
   datatype = 'multiple',
   initparsopt = c(2.5,2.7,20,0.009,1.01,0.009,1.01),
   idparsmat = rbind(1:5,c(1:3,6,7)),
   idparsopt = 1:7,
   parsfix = NULL,
   idparsfix = NULL
)

### When we consider the four Macaronesia archipelagoes and set all parameters the same
# except for rates of cladogenesis, extinction and immigration for Canary Islands,
# rate of cladogenesis is fixed to 0 for the other archipelagoes,
# diversity-dependence is assumed to be absent
# and we want to optimize all parameters, we use:

data(Macaronesia_datalist)
DAISIE_ML(
   datalist = Macaronesia_datalist,
   datatype = 'multiple',
   initparsopt = c(1.053151832,0.052148979,0.512939011,0.133766934,0.152763179),
   idparsmat = rbind(1:5,c(6,2,3,7,5),1:5,1:5),
   idparsopt = c(2,4,5,6,7),
   parsfix = c(0,Inf),
   idparsfix = c(1,3)
)

")
```

| DAISIE_numcol | *The expectation and marginal distribution of the number of colonizations (lineages) under the DAISIE model* |

Description

This function calculates expectation and marginal distribution of the number of colonizations (lineages) for a given set of parameter values, a given mainland species pool size and a given set of times

Usage

```
DAISIE_numcol(
    pars1,
    pars2,
    tvec,
    initEI = NULL
    )
```

Arguments

pars1   Vector of model parameters:

                 pars1[1] corresponds to lambda^c (cladogenesis rate)
                 pars1[2] corresponds to mu (extinction rate)
                 pars1[3] corresponds to K (clade-level carrying capacity)
                 pars1[4] corresponds to gamma (immigration rate)
                 pars1[5] corresponds to lambda^a (anagenesis rate)

pars2   Vector of settings:

                 pars2[1] corresponds to res, the maximum number of endemics or non-endemics for which the ODE system is solved; this must be much larger than the actual number for which the probability needs to be calculated.)
                 pars2[2] corresponds to M, size of the mainland pool, i.e the number of species that can potentially colonize the island.

tvec   The times at which the probabilities need to be computed.

initEI   A list with the initial values for the number of endemics and non-endemics in each colonizing lineage; when it is NULL, it is assumed that the island is empty

Value

out   A list of three vectors:

                 expC The expectation of the number of colonizations/lineages at the given times
                 pC The probability distribution of the number of colonizations (lineages) at the given times

Author(s)

Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Examples

```
### Compute the marginal probability distributions at t = 4 and t = 8, for a mainland
# pool size of 250 potential colonists and a vector of 5 parameters (cladogenesis,
# extinction, clade-level carrying capacity, immigration, anagenesis) starting from
# an empty island

DAISIE_numcol(
   pars1 = c(0.3,0.35,Inf,0.75,0.012),
   pars2 = c(100,250),
   tvec = c(4,8),
   initEI = list(c(0,1),c(0,2),c(3,1))
   )
```

```
DAISIE_plot_age_diversity
```
                    Plot clade age against clade diversity.

Description

Plots clade age against clade diversity for all clades for which colonisation time is known.

Usage

```
DAISIE_plot_age_diversity(
   island,
   title = 'Clade age vs clade diversity',
   island_age = NA)
```

Arguments

island	Island data object. Can be in DAISIE list format (see Galapagos_datalist and DAISIE_data_prep for examples) or in table format (see Galapagos_datatable for an example)
title	Title of the plot
island_age	If island input is in table format, the age of the island must be specified. If island input is in DAISIE list format, this option will override the island age specified in the island list.

Details

R plot showing for each clade in the island object the time of colonisation and the diversity of the clade. Only clades for which colonisation time is known are plotted. Blue - endemic; black - non-endemic. C = number of independent colonisations on island, N = number of species on the island

Value

R plot.

Author(s)

Luis Valente

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_sim DAISIE_plot_island

Examples

```
### Plot Galapagos age-diversity for Galapagos dataset
data(Galapagos_datalist)
DAISIE_plot_age_diversity(Galapagos_datalist)
```

DAISIE_plot_island	*Plot colonisation and branching time of species found on an island dataset.*

Description

Produces an image with the times of colonisation, branching times, clade name and clade status for all clades found in a given dataset.

Usage

```
DAISIE_plot_island(
island,
island_age=NA)
```

Arguments

island   Island data object. Can be in DAISIE list format (see Galapagos_datalist and DAISIE_data_prep for examples) or in table format (see Galapagos_datatable for an example)

island_age   If island input is in table format, the age of the island must be specified. If island input is in DAISIE list format, this option will override the island age specified in the island list.

Details

R plot showing for each clade in the island object: time of colonisation, branching times, species status (endemic or non-endemic), total number of species in clade (n), number of species not sampled (m).

Value

R plot.

Author(s)

Luis Valente

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_sim DAISIE_plot_age_diversity

Examples

```
### Plot Galapagos islands dataset
data(Galapagos_datalist)
DAISIE_plot_island(Galapagos_datalist)
```

DAISIE_plot_sims   *Plot island species-through-time (STT) plots*

Description

Produces STT plots. If only one type of species is present in the simulated islands, STT is plotted for all species. If two types are present, three plots are produced: STT for all, STT for type 1 and STT for type 2.

Usage

```
DAISIE_plot_sims(
island_replicates
)
```

Arguments

island_replicates

        Island replicates in DAISIE format (Produced in DAISIE_sim with format=TRUE option, or in DAISIE_format_sim)

Details

R plots with number of total, endemic and non-endemic STTs for different types of species for the entire time span the islands were simulated. 2.5-97.5th percentiles are plotted in light grey, 25-75th percentiles plotted in dark grey.

Value

R plot.

Author(s)

Luis Valente

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_sim DAISIE_format_CS

Examples

```
### Plot islands with single process (only one type of species)
data(islands_1type_1000reps)
DAISIE_plot_sims(island_replicates = islands_1type_1000reps)


### Plot island with type 1 and type 2
data(islands_2types_1000reps)
DAISIE_plot_sims(island_replicates = islands_2types_1000reps)
```

DAISIE_probdist	*The joint distribution of endemics and non-endemics under the DAISIE model*

Description

This function calculates the joint distribution of the number of endemics and non-endemics for a given set of parameter values, a given mainland species pool size and a given set of times

Usage

```
DAISIE_probdist(
    pars1,
    pars2,
    tvec,
    initEI = c(0,0),
    initprobs = NULL
    )
```

Arguments

pars1   Vector of model parameters:

                pars1[1] corresponds to lambda^c (cladogenesis rate)
                pars1[2] corresponds to mu (extinction rate)
                pars1[3] corresponds to K (clade-level carrying capacity)
                pars1[4] corresponds to gamma (immigration rate)
                pars1[5] corresponds to lambda^a (anagenesis rate)

pars2   Vector of settings:

                pars2[1] corresponds to res, the maximum number of endemics or non-endemics for which the ODE system is solved; this must be much larger than the actual number for which the probability needs to be calculated.)
                pars2[2] corresponds to M, size of the mainland pool, i.e the number of species that can potentially colonize the island.

tvec   The times at which the probabilities need to be computed.

initEI   The initial values for the number of endemics and non-endemics; either this or initprobs must be NULL

initprobs   The initial probability distribution for the number of endemics and non-endemics; either this or initEI must be NULL

Details

To obtain a matrix of probabilities with endemics in rows and non-endemics in columns for a certain time, one can run DAISIE_convertprobdist

Value

A matrix of dimensions 1 + length(tvec) and pars[2]^2 + 1] where the first column contains the times at which the probabilities are evaluated and the other columns contain the joint probabilities.

Author(s)

Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Examples

```
### Compute the probability distribution at t = 4 and t = 8, for a mainland pool
# size of 250 potential colonists and a vector of 5 parameters (cladogenesis,
# extinction, clade-level carrying capacity, immigration, anagenesis) starting
# from an empty island

DAISIE_probdist(
   pars1 = c(0.3,0.35,Inf,0.75,0.012),
   pars2 = c(100,250),
   tvec = c(4,8),
   initEI = c(0,0),
   initprobs = NULL
   )
```

DAISIE_sim   *Simulate islands with given parameters.*

Description

This function simulates islands with given cladogenesis, extinction, Kprime, immigration and anagenesis parameters. If a single parameter set is provided (5 parameters) it simulates islands where all species have the same macro-evolutionary process. If two paramater sets (10 parameters) are provided, it simulates islands where two different macro-evolutionary processes operate, one applying to type 1 species and other to type 2 species.

Usage

```
DAISIE_sim(
   time,
   M,
   pars,
   replicates,
   divdepmodel = 'CS',
   prop_type2_pool = NA,
```

```
        replicates_apply_type2 = TRUE,
        sample_freq = 25,
        plot_sims = TRUE,
        ...)
```

Arguments

time	Length of the simulation in time units. For example, if an island is know to be 4 million years old, setting time = 4 will simulate entire life span of the island; setting time = 2 will stop the simulation at the mid-life of the island.
M	The size of the mainland pool, i.e the number of species that can potentially colonize the island
pars	Contains the model parameters: pars[1] corresponds to lambda^c (cladogenesis rate) pars[2] corresponds to mu (extinction rate) pars[3] corresponds to K (clade-level carrying capacity). Set K=Inf for non-diversity dependence. pars[4] corresponds to gamma (immigration rate) pars[5] corresponds to lambda^a (anagenesis rate) pars[6] corresponds to lambda^c (cladogenesis rate) for type 2 species pars[7] corresponds to mu (extinction rate) for type 2 species pars[8] corresponds to K (clade-level carrying capacity) for type 2 species. Set K=Inf for non-diversity dependence. pars[9] corresponds to gamma (immigration rate) for type 2 species pars[10] corresponds to lambda^a (anagenesis rate) for type 2 species The elements 6:10 are optional and are required only when type 2 species are included.
replicates	Number of island replicates to be simulated.
divdepmodel	Option divdepmodel='CS' runs model with clade-specific carrying capacity, where diversity-dependence operates only within single clades, i.e. only among species originating from the same mainland colonist. Option divdepmodel='IW' runs model with island-wide carrying capacity, where diversity-dependence operates within and among clades.
prop_type2_pool	Fraction of mainland species that belongs to the second subset of species (type 2). Applies only when two types of species are simulated (length(pars)=10).
replicates_apply_type2	Applies only when two types of species are being simulated. Default replicates_apply_type2=TRUE runs simulations until the number of islands where a type 2 species has colonised is equal to the specified number of replicates. This is recommended if prop_type2_pool is small or if the rate of immigration of type two species (pars[9]) is low, meaning that more replicates are needed to achieve an adequate sample size of islands with type 2 species. Setting replicates_apply_type2=FALSE, simulates islands up to specified number of replicates regardless of whether type 2 species have colonised or not.

sample_freq   Specifies the number of units time should be divided by for plotting purposes.
                   Larger values will lead to plots with higher definition, but will also run slower.

plot_sims   Default=TRUE plots species-through-time (STT) plots. It detects how many
                   types of species are present. If only one type of species is present, STT is plotted
                   for all species. If two types are present, three plots are produced: STT for all,
                   STT for type 1 and STT for type 2.

...   Any arguments to pass on to plotting functions.

Details

Returns R list object that contains the simulated islands.

Value

Each simulated dataset is an element of the list, which can be called using [[x]]. For example if the
object is called island_replicates, the 1st replicate can be called using island_replicates[[1]] Each of
the island replicates is a list in itself. The first (e.g. island_replicates[[x]][[1]]) element of that list
has the following components:
$island_age - the island age
Then, depending on whether a distinction between types is made, we have:
$not_present - the number of mainland lineages that are not present on the island
or:
$not_present_type1 - the number of mainland lineages of type 1 that are not present on the island
$not_present_type2 - the number of mainland lineages of type 2 that are not present on the island
$stt_all - STT table for all species on the island (nI - number of non-endemic species; nA - num-
ber of anagenetic species, nC - number of cladogenetic species, present - number of independent
colonisations present)
$stt_stt_type1 - STT table for type 1 species on the island - only if 2 types of species were sim-
ulated (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of
cladogenetic species, present - number of independent colonisations present)
$stt_stt_type2 - STT table for type 2 species on the island - only if 2 types of species were sim-
ulated (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of
cladogenetic species, present - number of independent colonisations present)
$brts_table - Only for simulations under 'IW'. Table containing information on order of events
in the data, for use in maximum likelihood optimization.)

The subsequent elements of the list each contain information on a single colonist lineage on the
island and has 4 components:
$branching_times - island age and stem age of the population/species in the case of Non-endemic,
Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be
island age and branching times of the radiation including the stem age of the radiation.
$stac - the status of the colonist
* Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4
$missing_species - number of island species that were not sampled for particular clade (only ap-
plicable for endemic clades)

`$type_1or2` - whether the colonist belongs to type 1 or type 2

Author(s)

Luis Valente and Albert Phillimore

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_format_CS DAISIE_plot_sims

Examples

```
cat("
## Simulate 40 islands for 4 million years, where all species have equal
## rates, and plot the species-through-time plot. Pool size 1000.

pars_equal = c(2.550687345,2.683454548,Inf,0.00933207,1.010073119)
island_replicates_equal = DAISIE_sim(
   time = 4,
   M = 1000,
   pars = pars_equal,
   replicates = 40
   )

## Simulate 15 islands for 4 million years with two types of species (type1
## and type 2), and plot the species-through-time plot. Pool size 1000. Fraction
## of type 2 species in source pool is 0.163. Function will simulate until number of islands
## where type 2 species has colonised is equal to number specified in replicates.

pars_type1 = c(0.195442017,0.087959583,Inf,0.002247364,0.873605049)
pars_type2 = c(3755.202241,8.909285094,14.99999923,0.002247364,0.873605049)
island_replicates_2types = DAISIE_sim(
   time = 4,
   M = 1000,
   pars = c(pars_type1,pars_type2),
   replicates = 15,
   prop_type2_pool = 0.163
   )
")
```

DAISIE_tutorial   *Opens DAISIE tutorial*

Description

This function opens a PDF file that contains a step-by-step tutorial on how to run DAISIE functions, using the Galapagos avian dataset from Valente, Phillimore & Etienne 2015 as an example

Usage

```
DAISIE_tutorial()
```

Details

Opens a PDF file containing the tutorial

Author(s)

Luis Valente and Rampal S. Etienne

References

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

Galapagos_datalist   *Colonization and branching times of 8 terrestrial avifaunal clades in*
                            list format, accepted by DAISIE_ML and DAISIE_loglik_all

---

Description

A list containing the colonization and branching times of the terrestrial avifauna in the Galapagos where no distinction is made between types of colonists. This list can be generated using the DAISIE_dataprep function, which converts a user-specified data table into a data object, but the object can of course also be entered directly. It is an R list object with the following elements.

The first element of the list has two components:

$island_age - the island age
$not_present - the number of mainland lineages that are not present on the island

The following 8 elements of the list each contains information on a single colonist lineage on the island and has 5 components:

$colonist_name - the name of the species or clade that colonized the island
$branching_times - island age and stem age of the population/species in the case of Non-endemic,

Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be island age and branching times of the radiation including the stem age of the radiation.
`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4

`$missing_species` - number of island species that were not sampled for particular clade (only applicable for endemic clades)
`$type1or2` - whether the colonist belongs to type 1 or type 2. In this dataset all are equal to 1.

Usage

```
data(Galapagos_datalist)
```

Format

A list with 9 elements the first of which contains 2 elements and the following 8 containing 5 components.

Source

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

[DAISIE_dataprep](), [DAISIE_ML]()

Galapagos_datalist_2types

Colonization and branching times of 8 terrestrial avifaunal clades in list format, accepted by DAISIE_ML and DAISIE_loglik_all

---

Description

A list containing the colonization and branching times of the terrestrial avifauna in the Galapagos. This list can be generated using the DAISIE_dataprep function, which converts a user-specified data table into a data object, but the object can of course also be entered directly. It is an R list object with the following elements.

The first element of the list has three components:

`$island_age` - the island age

$not_present_type1 - the number of mainland lineages of type 1 that are not present on the island
$not_present_type2 - the number of mainland lineages of type 2 that are not present on the island

The following 8 elements of the list each contains information on a single colonist lineage on the island and has 5 components:

$colonist_name - the name of the species or clade that colonized the island
$branching_times - island age and stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be island age and branching times of the radiation including the stem age of the radiation.
$stac - the status of the colonist

* Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4

$missing_species - number of island species that were not sampled for particular clade (only applicable for endemic clades)
$type1or2 - whether the colonist belongs to type 1 or type 2. In this dataset only the finches are type 2

Usage

    data(Galapagos_datalist_2types)

Format

A list with 9 elements the first of which contains 3 elements and the following 8 containing 5 components.

Source

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_dataprep, DAISIE_ML

Galapagos_datatable   *Colonization and branching times of 8 terrestrial avifaunal clades in table format*

Description

A table containing the colonization and branching times of the terrestrial avifauna in the Galapagos. Each row on the table represents and independent colonisation event. The table has four columns.

$Clade_name - name of independent colonization event
$Status - One of the following categories:
* Non_endemic: for cases where both island and non-island populations of the species have been sampled)
* Non_endemic_MaxAge: for cases where island population of the species has not been sampled and only the age of the species is available)
* Endemic: applicable for both cladogenetic or anagenetic species
* Endemic&Non_Endemic: when endemic clade and mainland ancestor has re-colonized
$Missing_species - Number of island species that were not sampled for particular clade (only applicable for endemic clades)
$Branching_times - Stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be branching times of the radiation including the stem age of the radiation.

Usage

    data(Galapagos_datatable)

Format

A table with 8 rows and 4 columns.

Source

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

[DAISIE_dataprep](), [DAISIE_ML]()

---

islands_10reps_RAW *1000 islands in RAW format simulated with the ML parameters of the CR model for the Galapagos data*

Description

Each simulated dataset is an element of the list, which can be called using e.g. islands_10reps_RAW[[1]] Each of the island replicates is a list in itself. The first (e.g. islands_10reps_RAW[[x]][[1]]) element of that list has the following components:
The following elements of the RAW list each contain information on a single colonist lineage on the island and has 5 components:

`$branching_times` - island age and stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be island age and branching times of the radiation including the stem age of the radiation.
`$stac` - the status of the colonist
* Not_present: 0 \ cr * Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4
`$stt_table` - Species-through-time table for the descendants of the mainland species (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species)
`$missing_species` - number of island species that were not sampled for particular clade (only applicable for endemic clades)

Usage

    data(islands_10reps_RAW)

Format

A list with 10 items.

Source

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_sim, DAISIE_plot_sims

---

islands_1type_1000reps

1000 islands in DAISIE format simulated with the ML parameters of the CR model for the Galapagos data

---

Description

Each simulated dataset is an element of the list, which can be called using e.g.islands_1type_1000reps[[1]]
Each of the island replicates is a list in itself. The first (e.g. islands_1type_1000reps[[x]][[1]]) element of that list has the following components:
`$island_age` - the island age
`$not_present` - the number of mainland lineages that are not present on the island
`$stt_all` - STT table for all species on the island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)

The subsequent elements of the list each contain information on a single colonist lineage on the island and has 3 components:

`$branching_times` - island age and stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be island age and branching times of the radiation including the stem age of the radiation.

`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_endemic: 4

`$missing_species` - number of island species that were not sampled for particular clade (only applicable for endemic clades)

Usage

```
data(islands_1type_1000reps)
```

Format

A list with 1000 items.

Source

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_sim, DAISIE_plot_sims

---

islands_2types_1000reps

1000 islands in DAISIE format simulated with the ML parameters of the CR_lamc_mu_K model for the Galapagos data (2 types of species)

Description

Each simulated dataset is an element of the list, which can be called using e.g. islands_2types_1000reps[[1]] Each of the island replicates is a list in itself. The first (e.g. islands_2types_1000reps[[x]][[1]]) element of that list has the following components:

`$island_age` - the island age

`$not_present_type1` - the number of mainland lineages of type 1 that are not present on the island

`$not_present_type2` - the number of mainland lineages of type 2 that are not present on the island

`$stt_all` - STT table for all species on the island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)

`$stt_stt_type1` - STT table for type 1 species on the island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)

`$stt_stt_type2` - STT table for type 2 species on the island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)

The subsequent elements of the list each contain information on a single colonist lineage on the island and has 4 components:

`$branching_times` - island age and stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge and Endemic anagenetic species. For cladogenetic species these should be island age and branching times of the radiation including the stem age of the radiation.

`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4

`$missing_species` - number of island species that were not sampled for particular clade (only applicable for endemic clades)

`$type_1or2` - whether the colonist belongs to type 1 or type 2

Usage

```
data(islands_2types_1000reps)
```

Format

A list with 1000 items.

Source

Valente, L.M., A.B. Phillimore and R.S. Etienne (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. Ecology Letters 18: 844-852.

See Also

DAISIE_sim, DAISIE_plot_sims

| Macaronesia_datalist | *Colonization and branching times of terrestrial avifaunal clades from Azores, Canary Islands, Cape Verde and Madeira in list format, accepted by DAISIE_ML and DAISIE_loglik_all* |

Description

A list containing the colonization and branching times of the terrestrial avifauna in 4 archipelagos: Azores, Canary Islands, Cape Verde and Madeira. It is an R list object with the 4 main elements corresponding to each of the archipelagos (e.g. Macaronesia_datalist[[1]] calls the Azores data). Each of the four elements is then made of several elemants:

The first element of the list for an archipelago has two components:

`$island_age` - the island age
`$not_present` - the number of mainland lineages that are not present on the island

The following elements of the list each contains information on a single colonist lineage on the island and has 5 components:

`$colonist_name` - the name of the species or clade that colonized the island
`$branching_times` - island age and stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge, Endemic_MaxAge and Endemic anagenetic species. For cladogenetic species the island age and branching times of the radiation including the stem age of the radiation are shown.
`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1
* Endemic: 2
* Endemic&Non_Endemic: 3
* Non_endemic: 4

* Endemic_MaxAge: 5

`$missing_species` - number of island species that were not sampled for particular clade (only applicable for endemic clades)
`$type1or2` - whether the colonist belongs to type 1 or 2. In this dataset all are equal to 1.

Usage

```
data(Macaronesia_datalist)
```

Format

A list with 4 main elements for each archipelago. Each element has several sub-elements.

Source

Valente L., Illera J.C, Havenstein K., Pallien T., Etienne R.S., Tiedemann R. Macroevolutionary dynamics in Atlantic island avifaunas: were MacArthur & Wilson right about equilibrium? Under review.

See Also

DAISIE_dataprep, DAISIE_ML

Index