

Package ‘DAISIE’

August 26, 2020

Type Package

Title Dynamical Assembly of Islands by Speciation, Immigration and Extinction

Version 3.0.1

Date 2020-08-25

Depends R (>= 3.5.0)

biocViews

Imports deSolve, graphics, stats, utils, DDD (>= 4.4), subplex, Matrix, tensor, pracma, testit, foreach, doParallel, magrittr, parallel

Suggests covr, testthat (>= 2.1.0), knitr, rmarkdown, gridExtra, dplyr, ggplot2, ggtree, tidytree, tidyr, purrr, ape, phytools, tibble

Enhances doMC

License GPL-3

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

NeedsCompilation yes

Encoding UTF-8

VignetteBuilder knitr

RoxygenNote 7.1.1

Author Rampal S. Etienne [aut, cre] (<<https://orcid.org/0000-0003-2142-7612>>), Luis Valente [aut] (<<https://orcid.org/0000-0003-4247-8785>>), Albert B. Phillimore [aut] (<<https://orcid.org/0000-0002-6553-1553>>), Bart Haegeman [aut] (<<https://orcid.org/0000-0003-2325-4727>>), Joshua W. Lambert [aut], Pedro Neves [aut] (<<https://orcid.org/0000-0003-2561-4677>>), Shu Xie [aut], Richèl J.C. Bilderbeek [aut] (<<https://orcid.org/0000-0003-1107-7049>>), Torsten Hauffe [ctb] (<<https://orcid.org/0000-0001-5711-9457>>),

Giovanni Laudanno [ctb],
 Nadiah Kristensen [ctb],
 Raphael Scherrer [ctb]

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

Repository CRAN

Date/Publication 2020-08-26 10:00:07 UTC

R topics documented:

archipelagos41	3
archipelago_data	3
Bats_GreaterAntilles	4
Biwa_datalist	5
create_area_pars	6
create_CS_version	7
create_hyper_pars	8
create_trait_pars	9
DAISIE_convertprobdist	10
DAISIE_create_island	11
DAISIE_dataprep	12
DAISIE_ExpEIN	15
DAISIE_format_CS_full_stt	17
DAISIE_get_brts_myra	18
DAISIE_IC	19
DAISIE_loglik_IW	21
DAISIE_margprobdist	23
DAISIE_ML	25
DAISIE_ML_IW	32
DAISIE_MW_ML	35
DAISIE_numcol	40
DAISIE_plot_age_diversity	41
DAISIE_plot_input	42
DAISIE_plot_island	44
DAISIE_plot_sims	46
DAISIE_probdist	48
DAISIE_sim	49
DAISIE_sim_constant_rate_shift	54
DAISIE_sim_MW	58
DAISIE_sim_relaxed_rate	61
DAISIE_sim_time_dependent	65
DAISIE_sim_trait_dependent	69
DAISIE_SR_loglik_CS	73
DAISIE_SR_ML	75
frogs_datalist	81
frogs_datatable	82
Galapagos_datalist	83
Galapagos_datalist_2types	84

Galapagos_datatable	85
islands_10reps_RAW	86
islands_1type_1000reps	87
islands_2types_1000reps	88
Macaronesia_datalist	89
NewZealand_birds_datalist	90

Index	92
--------------	-----------

archipelagos41	<i>DAISIE datalist object including bird phylogenetic data and physical data for 41 archipelagos</i>
----------------	--

Description

A datalist with 41 items representing the 41 archipelagos. Each archipelago can be called separately using `archipelagos41[[x]]` with `x` being a number between 1 and 41. Using `archipelagos41[[x]][[1]]` will show just the top part of the archipelago item where the archipelago name and physical features are displayed. The structure of each of the archipelagos is the same as regular DAISIE datalist generated using `DAISIE_dataprep`.

Format

A datalist containing data on the 41 archipelagos studied in Valente et al 2020 (Main Dataset D1). Contains colonisation and branching times for bird species in each of the archipelagos. It also contains information on archipelago name, area, age and distance from the nearest mainland.

Source

Valente L, Phillimore AB, Melo M, Warren BH, Clegg SM, Havenstein K, Tiedemann R, Illera JC, Thébaud C, Aschenbach T, Etienne RS. A simple dynamic model explains island bird diversity worldwide (2020) *Nature*, 579, 92-96

archipelago_data	<i>Physical data on 41 archipelagos</i>
------------------	---

Description

A dataframe with in subsequent columns the name of the archipelago (Archipelago) the area of the archipelago (Area), the age (Age) and the distance from the mainland (Distance)

Format

A dataframe containing information on archipelago name, area, age and distance from the mainland

Source

Valente L, Phillimore AB, Melo M, Warren BH, Clegg SM, Havenstein K, Tiedemann R, Illera JC, Thébaud C, Aschenbach T, Etienne RS. A simple dynamic model explains island bird diversity worldwide (2020) *Nature*, 579, 92-96

Bats_GreaterAntilles *Colonization and branching times of Noctilionoid bats from the Greater Antilles.*

Description

A list containing the colonization and branching times of the noctilionoid bats of the Greater Antilles. Main dataset used in Valente, Etienne and Dávalos (2017) *Nature Ecology and Evolution*. Island age 20 Myr and mainland pool size of 100 species.

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 elements of the list each contain 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

\$stac - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_endemic: 4

* Endemic_MaxAge: 5 or 6

\$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.

Format

A list with 17 elements, the first of which contains 2 elements and the following 16 containing 5 components.

Source

Valente L, Etienne RS, Dávalos (2017) Recent extinctions disturb path to equilibrium diversity in Caribbean bats. *Nature Ecology and Evolution*, 1, 26.

See Also

[DAISIE_dataprep](#), [DAISIE_ML](#), [DAISIE_SR_ML](#)

Biwa_datalist

Colonization and branching times of 68 fish clades.

Description

A list containing the colonization and branching times of the fishes of Lake Biwa (Japan). Main dataset used in Hauffe et al (2020). 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 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_endemic: 4

* Endemic_MaxAge: 5 or 6

`$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.

Format

A list with 69 elements, the first of which contains 2 elements and the following 68 containing 5 components.

Source

Hauffe, T., D. Delicado, R.S. Etienne and L. Valente. Lake expansion elevates equilibrium diversity via increasing colonisation. (2020) Journal of Biogeography

See Also

[DAISIE_dataprep](#), [DAISIE_ML](#), [DAISIE_SR_ML](#)

create_area_pars	<i>Create named list of area parameters</i>
------------------	---

Description

Create named list of area parameters

Usage

```
create_area_pars(
  max_area,
  current_area,
  proportional_peak_t,
  total_island_age,
  sea_level_amplitude,
  sea_level_frequency,
  island_gradient_angle
)
```

Arguments

max_area	Numeric defining maximum area.
current_area	A numeric with the current island area at present (i.e., at the end of the simulation).
proportional_peak_t	Numeric value from 0 to 1 indicating where in the island's history the peak area is achieved.
total_island_age	Numeric defining total island age.
sea_level_amplitude	Numeric defining amplitude of area fluctuation from sea level.

sea_level_frequency

Numeric defining frequency of sine wave of area change from sea level.

island_gradient_angle

Numeric defining the angle in degrees specifying the slope of the island.

Value

list of numerical values containing area and sea level parameters for island ontogeny simulation

Author(s)

Richel J.C Bilderbeek, Joshua Lambert, Pedro Neves

Examples

```
testit::assert(DAISIE:::are_area_pars(
  create_area_pars(
    max_area = 10,
    current_area = 1,
    proportional_peak_t = 0.5,
    total_island_age = 5,
    sea_level_amplitude = 5,
    sea_level_frequency = 10,
    island_gradient_angle = 0
  )
)
)
```

create_CS_version *Creates the list object for CS_version argument in DAISIE_ML_CS*

Description

Creates the list object for CS_version argument in DAISIE_ML_CS

Usage

```
create_CS_version(model = 1, relaxed_par = NULL)
```

Arguments

model	the CS model to run, options are 1 for single rate DAISIE model, 2 for multi-rate DAISIE, or 0 for IW test model
relaxed_par	the parameter to relax (integrate over). Options are "cladogenesis", "extinction", "carrying_capacity", "immigration", or "anagenesis"

Value

A list of two elements

- model: the CS model to run, options are 1 for single rate DAISIE model, 2 for multi-rate DAISIE, or 0 for IW test model
- relaxed_par: the parameter to relax (integrate over). Options are "cladogenesis", "extinction", "carrying_capacity", "immigration", or "anagenesis"

create_hyper_pars *Create list of hyperparameters*

Description

Create list of hyperparameters

Usage

```
create_hyper_pars(d, x)
```

Arguments

- | | |
|---|--|
| d | Numeric defining the scaling parameter for exponent for calculating cladogenesis rate. |
| x | Numeric defining the exponent for calculating extinction rate. |

Value

Named list with hyperparameters

Author(s)

Pedro Neves, Joshua Lambert

Examples

```
hyper_pars <- create_hyper_pars(d = 0.027, x = 0.15)
```

create_trait_pars	<i>Create named list of trait state parameters</i>
-------------------	--

Description

Create named list of trait state parameters

Usage

```
create_trait_pars(  
  trans_rate,  
  immig_rate2,  
  ext_rate2,  
  ana_rate2,  
  clado_rate2,  
  trans_rate2,  
  M2  
)
```

Arguments

trans_rate	A numeric with the per capita transition rate with state1
immig_rate2	A numeric with the per capita immigration rate with state2
ext_rate2	A numeric with the per capita extinction rate with state2
ana_rate2	A numeric with the per capita anagenesis rate with state2
clado_rate2	A numeric with the per capita cladogenesis rate with state2
trans_rate2	A numeric with the per capita transition rate with state2
M2	A numeric with the number of species with trait state 2 on mainland

Value

list of numerical values containing trait state parameters

Examples

```
testit::assert(DAISIE:::are_trait_pars(  
  create_trait_pars(  
    trans_rate = 0.5,  
    immig_rate2 = 0.1,  
    ext_rate2 = 0.2,  
    ana_rate2 = 0.3,  
    clado_rate2 = 0.4,  
    trans_rate2 = 0.5,  
    M2 = 1000)) == TRUE)
```

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; store 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
)
prob_dists <- DAISIE_convertprobdist(pb)
```

DAISIE_create_island *Converts simulation output into island output*

Description

Converts simulation output into island output

Usage

```
DAISIE_create_island(
  stt_table,
  totaltime,
  island_spec,
  mainland_n,
  trait_pars = NULL
)
```

Arguments

stt_table	Matrix with number of species at each time step.
totaltime	Numeric defining the length of the simulation in time units.
island_spec	Matrix with current state of simulation containing number of species.
mainland_n	A numeric stating the number of mainland species, that is the number of species that can potentially colonize the island. If using a clade-specific diversity dependence, this value is set to 1. If using an island-wide diversity dependence, this value is set to the number of mainland species.
trait_pars	A named list containing diversification rates considering two trait states created by create_trait_pars : <ul style="list-style-type: none"> • [1]:A numeric with the per capita transition rate with state1 • [2]:A numeric with the per capita immigration rate with state2 • [3]:A numeric with the per capita extinction rate with state2 • [4]:A numeric with the per capita anagenesis rate with state2 • [5]:A numeric with the per capita cladogenesis rate with state2 • [6]:A numeric with the per capita transition rate with state2 • [7]:A numeric with the number of species with trait state 2 on mainland

Value

list with the island information, composed stt table, branching times of extant species, status of species on the island and number of missing species.

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-05,
  verbose = TRUE
)
```

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 an independent 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 when an approximate colonisation time is known

* "Non_endemic_MaxAge": applies to non-endemic species for cases where colonisation time is unknown

* "Endemic": applies to endemic species or endemic clades when an approximate colonisation time is known

* "Endemic_MaxAge": applies to endemic species or endemic clades for cases where the colonisation time is unknown, or when the user wants to specify an upper bound for colonisation. This could for example apply to endemic species that have recently gone extinct because of anthropogenic causes, and which are not included in the phylogeny ("NA" should be given in the branching times column). It could also apply to insular radiations with long stem branches, for which the time of the first cladogenetic event is known, but the precise time of colonisation is not.

* "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 particular clade (only applicable for "Endemic" clades). If NA is given in branching times column, this should be equal to the number of species in the clade minus 1

\$Branching_times - Stem age of the population/species in the case of "Non_endemic", "Non_endemic_MaxAge" and "Endemic" species with no extant close relatives on the island. Set "NA" if colonisation time unknown and no upper bound is known. For "Endemic" cladogenetic species these should be branching times of the radiation, including the stem age of the radiation (colonisation time estimate).

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 considered 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 macroevolutionary process. Applies only if number_clade_types = 2. Default "proportional" 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.
verbose	Boolean. States if intermediate results should be printed to console. Defaults to TRUE.

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 (if only colonisation time was given)

* Endemic_MaxAge: 6 (if colonisation time and cladogenesis times were given)

`$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
```

```
utils::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.

  utils::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.

  utils::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	<p>A numeric vector containing the model parameters:</p> <ul style="list-style-type: none"> • pars[1]: λ^c (cladogenesis rate) • pars[2]: μ (extinction rate) • pars[3]: K (carrying capacity), set K=Inf for diversity independence. • pars[4]: γ (immigration rate) • pars[5]: λ^a (anagenesis rate) • pars[6]: λ^c (cladogenesis rate) for either type 2 species or rate set 2 in rate shift model • pars[7]: μ (extinction rate) for either type 2 species or rate set 2 in rate shift model • pars[8]: K (carrying capacity) for either type 2 species or rate set 2 in rate shift model, set K=Inf for diversity independence. • pars[9]: γ (immigration rate) for either type 2 species or rate set 2 in rate shift model • pars[10]: λ^a (anagenesis rate) for either type 2 species or rate set 2 in rate shift model <p>Elements 6:10 are required only when type 2 species are included or in the rate shift model. For <code>DAISIE_sim_relaxed_rate()</code> pars[6] is the standard deviation of the gamma distribution for the relaxed parameter and the parameter chosen by the <code>relaxed_par</code> argument is the mean of the gamma distribution for the relaxed parameter.</p>
M	Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island.
initEI	The initial values for the number of endemics and non-endemics. In <code>DAISIE_probdist()</code> or <code>DAISIE_margprobdist()</code> either this or <code>initprobs</code> must be NULL. In <code>DAISIE_numcol()</code> when it is NULL, it is assumed that the island is empty.

Value

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

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_format_CS_full_stt

Formats clade-specific simulation output into standard DAISIE list output

Description

Formats clade-specific simulation output into standard DAISIE list output

Usage

```
DAISIE_format_CS_full_stt(
  island_replicates,
  time,
  M,
  verbose = TRUE,
  trait_pars = NULL
)
```

Arguments

- | | |
|-------------------|---|
| island_replicates | List output from DAISIE_sim_core_constant_rate() , DAISIE_sim_core_time_dependent() , DAISIE_sim_core_constant_rate_shift() or DAISIE_sim_min_type2() functions. Minimally, this must be a list that has as many elements as replicates. Each element must be a list with the elements <code>island_age</code> , <code>not_present</code> and <code>stt_all</code> . <code>stt_all</code> must be a data frame with the column names <code>Time</code> , <code>nI</code> , <code>nA</code> , <code>nC</code> and <code>present</code> . |
| time | Numeric defining the length of the simulation in time units. For example, if an island is known to be 4 million years old, setting <code>time = 4</code> will simulate the entire life span of the island; setting <code>time = 2</code> will stop the simulation at the mid-life of the island. |
| M | Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island. |

verbose	In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
trait_pars	A named list containing diversification rates considering two trait states created by <code>create_trait_pars</code> : <ul style="list-style-type: none"> • [1]:A numeric with the per capita transition rate with state1 • [2]:A numeric with the per capita immigration rate with state2 • [3]:A numeric with the per capita extinction rate with state2 • [4]:A numeric with the per capita anagenesis rate with state2 • [5]:A numeric with the per capita cladogenesis rate with state2 • [6]:A numeric with the per capita transition rate with state2 • [7]:A numeric with the number of species with trait state 2 on mainland

Value

List with CS DAISIE simulation output

DAISIE_get_brts_my	<i>Extract the sorted branching times, in million years ago. from a data table</i>
--------------------	--

Description

Extract the sorted branching times, in million years ago. from a data table

Usage

```
DAISIE_get_brts_my(data_table)
```

Arguments

data_table data table

Value

the sorted branching times, in million years ago

Author(s)

Richel J.C. Bilderbeek

Examples

```
data(Galapagos_datatable)
brts_my <- DAISIE_get_brts_my(data_table = Galapagos_datatable)
testit::assert(length(brts_my) > 1)
testit::assert(all(brts_my > 0))
```

DAISIE_IC

Calculates information criterion from DAISIE ML estimates?

Description

Calculates information criterion from DAISIE ML estimates?

Usage

```
DAISIE_IC(
  datalist,
  initparsopt,
  idparsopt,
  parsfix,
  idparsfix,
  endmc = 1000,
  res = 100,
  cond = 0,
  ddmmodel = 0
)
```

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 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island

age followed by the branching times of the island clade including the stem age of the clade

\$stac - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_Endemic: 4

* Endemic_Singleton_MaxAge: 5

* Endemic_Clade_MaxAge: 6

* Endemic&Non_Endemic_Clade_MaxAge: 7

\$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

initparsopt	The initial values of the parameters that must be optimized, they are all positive.
idparsopt	The ids of the parameters that must be optimized. The ids are defined as follows: id = 1 corresponds to λ^c (cladogenesis rate) id = 2 corresponds to μ (extinction rate) id = 3 corresponds to K (clade-level carrying capacity) id = 4 corresponds to γ (immigration rate) id = 5 corresponds to λ^a (anagenesis rate) id = 6 corresponds to λ^c (cladogenesis rate) for an optional subset of the species id = 7 corresponds to μ (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 γ (immigration rate) for an optional subset of the species id = 10 corresponds to λ^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).
parsfix	The values of the parameters that should not be optimized.
idparsfix	The ids of the parameters that should not be optimized, e.g. c(1,3) if λ^c and K should not be optimized.
endmc	Numeric for how many simulations should run.
res	Sets the maximum number of species for which a probability must be computed, must be larger than the size of the largest clade.
cond	cond = 0 : conditioning on island age cond = 1 : conditioning on island age and non-extinction of the island biota .
ddmodel	Sets the model of diversity-dependence:

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

Value

List of two numerics with WIC and AICb

DAISIE_loglik_IW	<i>Computes the loglikelihood of the DAISIE model with island-wide diversity-dependence given data and a set of model parameters</i>
------------------	--

Description

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

Usage

```

DAISIE_loglik_IW(
  pars1,
  pars2,
  datalist,
  methode = "ode45",
  abstolint = 1e-16,
  reltolint = 1e-14,
  verbose = FALSE
)

```

Arguments

`pars1` Contains the model parameters:

`pars1[1]` corresponds to λ^c (cladogenesis rate)
`pars1[2]` corresponds to μ (extinction rate)
`pars1[3]` corresponds to K (clade-level carrying capacity)
`pars1[4]` corresponds to γ (immigration rate)
`pars1[5]` corresponds to λ^a (anagenesis rate)
`pars1[6]` is optional; it may contain M , the total number of species on the mainland

`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
Only `ddmodel = 11` is currently implemented

`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]` Specifies whether intermediate output should be provided, because computation may take long. Default is 0, no output. A value of 1 means the parameters and loglikelihood are printed. A value of 2 means also intermediate progress during loglikelihood computation is shown.

`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
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)
methode	Method of the ODE-solver. See package deSolve for details. Default is "ode45"
abstolint	Absolute tolerance of the integration
reltolint	Relative tolerance of the integration
verbose	Logical controlling if progress is printed to console.

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_IW](#), [DAISIE_loglik_CS](#), [DAISIE_sim_constant_rate](#)

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	<p>Vector of model parameters:</p> <p>pars1[1] corresponds to λ^c (cladogenesis rate) pars1[2] corresponds to μ (extinction rate) pars1[3] corresponds to K (clade-level carrying capacity) pars1[4] corresponds to γ (immigration rate) pars1[5] corresponds to λ^a (anagenesis rate).</p>
pars2	<p>Vector of settings:</p> <p>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.</p>
tvec	The times at which the probabilities need to be computed.
initEI	The initial values for the number of endemics and non-endemics. In DAISIE_probdist() or DAISIE_margprobdist() either this or <code>initprobs</code> must be <code>NULL</code> . In DAISIE_numcol() when it is <code>NULL</code> , it is assumed that the island is empty.
initprobs	The initial probability distribution for the number of endemics and non-endemics; either this or <code>initEI</code> must be <code>NULL</code> .
pb	Probability distribution in matrix format as output by DAISIE_probdist() .

Value

out	<p>A list of three vectors:</p> <p>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</p>
-----	--

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

marg_prob_dists <- 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 with clade-specific diversity-dependence

Description

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

Usage

```
DAISIE_ML_CS(
  datalist,
  datatype = "single",
  initparsopt,
  idparsopt,
  parsfix,
  idparsfix,
  idparsnoshift = 6:10,
  idparsmat = NULL,
```

```

res = 100,
ddmodel = 0,
cond = 0,
island_ontogeny = NA,
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",
CS_version = 1,
verbose = 0,
tolint = c(1e-16, 1e-10),
jitter = 0
)

```

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 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade

`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

	<ul style="list-style-type: none"> * Endemic&Non_Endemic: 3 * Non_Endemic: 4 * Endemic_Singleton_MaxAge: 5 * Endemic_Clade_MaxAge: 6 * Endemic&Non_Endemic_Clade_MaxAge: 7
	<p>\$missing_species - number of island species that were not sampled for particular clade (only applicable for endemic clades)</p> <p>\$type1or2 - whether the colonist belongs to type 1 or type 2</p>
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, they are all positive.
idparsopt	<p>The ids of the parameters that must be optimized. The ids are defined as follows:</p> <ul style="list-style-type: none"> id = 1 corresponds to λ^c (cladogenesis rate) id = 2 corresponds to μ (extinction rate) id = 3 corresponds to K (clade-level carrying capacity) id = 4 corresponds to γ (immigration rate) id = 5 corresponds to λ^a (anagenesis rate) id = 6 corresponds to λ^c (cladogenesis rate) for an optional subset of the species id = 7 corresponds to μ (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 γ (immigration rate) for an optional subset of the species id = 10 corresponds to λ^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).
parsfix	The values of the parameters that should not be optimized.
idparsfix	The ids of the parameters that should not be optimized, e.g. c(1,3) if λ^c and K 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 λ^c and μ have the same values for both groups.
idparsmat	<p>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:</p> <ul style="list-style-type: none"> * lac = (initial) cladogenesis rate * mu = extinction rate * K = maximum number of species possible in the clade * gam = (initial) immigration rate

	<p>* λ_{aa} = (initial) anagenesis rate Example: <code>idparsmat = rbind(c(1, 2, 3, 4, 5), c(1, 2, 3, 6, 7))</code> has different rates of immigration and anagenesis for the two islands.</p>
<code>res</code>	Sets the maximum number of species for which a probability must be computed, must be larger than the size of the largest clade.
<code>ddmodel</code>	<p>Sets the model of diversity-dependence:</p> <p><code>ddmodel = 0</code> : no diversity dependence <code>ddmodel = 1</code> : linear dependence in speciation rate <code>ddmodel = 11</code>: linear dependence in speciation rate and in immigration rate <code>ddmodel = 2</code> : exponential dependence in speciation rate <code>ddmodel = 21</code>: exponential dependence in speciation rate and in immigration rate</p>
<code>cond</code>	<p><code>cond = 0</code> : conditioning on island age <code>cond = 1</code> : conditioning on island age and non-extinction of the island biota</p>
<code>island_ontogeny</code>	<p>In <code>DAISIE_sim_time_dependent()</code>, <code>DAISIE_ML_CS</code> and plotting a string describing the type of island ontogeny. Can be "const", "beta" for a beta function describing area through time. String checked by <code>is_island_ontogeny_input()</code>. In all other functions a numeric describing the type of island ontogeny. Can be 0 for constant, 1 for a beta function describing area through time. In ML functions <code>island_ontogeny = NA</code> assumes constant ontogeny.</p>
<code>eqmodel</code>	<p>Sets the equilibrium constraint that can be used during the likelihood optimization. Only available for <code>datatype = 'single'</code>.</p> <p><code>eqmodel = 0</code> : no equilibrium is assumed <code>eqmodel = 13</code> : near-equilibrium is assumed on endemics using deterministic equation for endemics and immigrants. Endemics must be within x_E of the equilibrium value <code>eqmodel = 15</code> : 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.</p>
<code>x_E</code>	Sets the fraction of the equilibrium endemic diversity above which the endemics are assumed to be in equilibrium; only active for <code>eqmodel = 13</code> or <code>15</code> .
<code>x_I</code>	Sets the fraction of the equilibrium non-endemic diversity above which the system is assumed to be in equilibrium; only active for <code>eqmodel = 15</code> .
<code>tol</code>	<p>Sets the tolerances in the optimization. Consists of:</p> <p><code>reltolx</code> = relative tolerance of parameter values in optimization <code>reltolf</code> = relative tolerance of function value in optimization <code>abstolx</code> = absolute tolerance of parameter values in optimization.</p>
<code>maxiter</code>	Sets the maximum number of iterations in the optimization.
<code>methode</code>	Method of the ODE-solver. See package <code>deSolve</code> for details. Default is "Isodes".

optimmethod	Method used in likelihood optimization. Default is "subplex" (see subplex package). Alternative is 'simplex' which was the method in previous versions.
CS_version	a numeric or list. Default is 1 for the standard DAISIE model, for a relaxed-rate model a list with the following elements: <ul style="list-style-type: none"> • model: the CS model to run, options are 1 for single rate DAISIE model, 2 for multi-rate DAISIE, or 0 for IW test model • relaxed_par: the parameter to relax (integrate over). Options are "cladogenesis", "extinction", "carrying_capacity", "immigration", or "anagenesis"
verbose	In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
tolint	Vector of two elements containing the absolute and relative tolerance of the integration.
jitter	Numeric for <code>optimizer()</code> . Jitters the parameters being optimized by the specified amount which should be very small, e.g. $1e-5$. Jitter when <code>link[subplex]{subplex}()</code> produces incorrect output due to parameter transformation.

Value

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

lambda_c	gives the maximum likelihood estimate of λ^c , the rate of cladogenesis
mu	gives the maximum likelihood estimate of μ , the extinction rate
K	gives the maximum likelihood estimate of K, the carrying-capacity
gamma	gives the maximum likelihood estimate of γ , the immigration rate
lambda_a	gives the maximum likelihood estimate of λ^a , the rate of anagenesis
lambda_c2	gives the maximum likelihood estimate of λ^{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 λ^{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 freedom
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_constant_rate](#), [DAISIE_sim_time_dependent](#), [DAISIE_sim_constant_rate_shift](#)

Examples

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

```
utils::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:
```

```
utils::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:
```

```
utils::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:
```

```
utils::data(Galapagos_datalist_2types)
DAISIE_ML(
  datalist = Galapagos_datalist_2types,
  dmodel = 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:
```

```
utils::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:
```

```
utils::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:
```

```
utils::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_ML_IW

Maximization of the loglikelihood under the DAISIE model with island-wide diversity-dependence

Description

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

Usage

```

DAISIE_ML_IW(
  datalist,
  initparsopt,
  idparsopt,
  parsfix,
  idparsfix,
  res = 100,
  ddmmodel = 11,
  cond = 0,
  tol = c(1e-04, 1e-05, 1e-07),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  methode = "ode45",
  optimmethod = "subplex",
  verbose = 0,
  tolint = c(1e-16, 1e-14),
  jitter = 0
)

```

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 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade

$\$stac$ - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_Endemic: 4

* Endemic_Singleton_MaxAge: 5

* Endemic_Clade_MaxAge: 6

* Endemic&Non_Endemic_Clade_MaxAge: 7

$\$missing_species$ - 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

`initparsopt` The initial values of the parameters that must be optimized, they are all positive.

`idparsopt` The ids of the parameters that must be optimized. The ids are defined as follows:

id = 1 corresponds to λ^c (cladogenesis rate)

id = 2 corresponds to μ (extinction rate)

id = 3 corresponds to K (clade-level carrying capacity)

id = 4 corresponds to γ (immigration rate)

id = 5 corresponds to λ^a (anagenesis rate)

id = 6 corresponds to λ^c (cladogenesis rate) for an optional subset of the species

id = 7 corresponds to μ (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).
parsfix	The values of the parameters that should not be optimized.
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.
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 .
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 "Isodes".
optimmethod	Method used in likelihood optimization. Default is "subplex" (see subplex package). Alternative is 'simplex' which was the method in previous versions.
verbose	In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
tolint	Vector of two elements containing the absolute and relative tolerance of the integration.
jitter	Numeric for <code>optimizer()</code> . Jitters the parameters being optimized by the specified amount which should be very small, e.g. 1e-5. Jitter when <code>link[subplex]{subplex}()</code> produces incorrect output due to parameter transformation.

Details

The result of `sort(c(idparsopt, idparsfix))` should be identical to `c(1:5)`. 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.

Value

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

lambda_c	gives the maximum likelihood estimate of λ^c , the rate of cladogenesis
mu	gives the maximum likelihood estimate of μ , the extinction rate
K	gives the maximum likelihood estimate of K, the carrying-capacity
gamma	gives the maximum likelihood estimate of γ , the immigration rate
lambda_a	gives the maximum likelihood estimate of λ^a , the rate of anagenesis
loglik	gives the maximum loglikelihood
df	gives the number of estimated parameters, i.e. degrees of freedom
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_IW](#), [DAISIE_ML_CS](#) [DAISIE_sim_constant_rate](#)

DAISIE_MW_ML	<i>Maximization of the loglikelihood under the DAISIE model with clade-specific diversity-dependence and explicit dependencies on island area and isolation as hypothesized by MacArthur & Wilson</i>
--------------	---

Description

This function computes the maximum likelihood estimates of the parameters of the relationships between parameters of the DAISIE model (with clade-specific diversity-dependence) and island area and distance of the island to the mainland for data from lineages colonizing several islands/archipelagos. It also outputs the corresponding loglikelihood that can be used in model comparisons.

A note on the sigmoidal functions used in distance_dep: For anagenesis and cladogenesis, the functional relationship is $k * (d/d0)^x / (1 + (d/d0)^x)$; for colonization the relationship is: $k - k * (d/d0)^x / (1 + (d/d0)^x)$. The d0 parameter is the 11th parameter entered. In 'sigmoidal_col_ana', the 11th parameter is the d0 for colonization and the 12th is the d0 for anagenesis.

Usage

```

DAISIE_MW_ML(
  datalist,
  initparsopt,
  idparsopt,
  parsfix,
  idparsfix,
  res = 100,
  ddmodel = 11,
  cond = 0,
  island_ontogeny = NA,
  tol = c(1e-04, 1e-05, 1e-07),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  methode = "lsodes",
  optimmethod = "subplex",
  CS_version = 1,
  verbose = 0,
  tolint = c(1e-16, 1e-10),
  distance_type = "continent",
  distance_dep = "power",
  parallel = "local",
  cpus = 3
)

```

Arguments

<code>datalist</code>	Data object containing information on colonisation and branching times of species for several islands or archipelagos, as well as the area, isolation and age of each of the islands/archipelagos. See <code>data(archipelagos41)</code> for an example.
<code>initparsopt</code>	The initial values of the parameters that must be optimized; they are all positive
<code>idparsopt</code>	The ids of the parameters that must be optimized. The ids are defined as follows (see Valente et al 2020 Supplementary Tables 1 and 2 a better explanation of the models and parameters):
	<ul style="list-style-type: none"> id = 1 corresponds to λ^{c0} (cladogenesis rate for unit area) id = 2 corresponds to y (exponent of area for cladogenesis rate) id = 3 corresponds to μ_0 (extinction rate for unit area) id = 4 corresponds to x (exponent of 1/area for extinction rate) id = 5 corresponds to K_0 (clade-level carrying capacity for unit area) id = 6 corresponds to z (exponent of area for clade-level carrying capacity) id = 7 corresponds to γ_0 (immigration rate for unit distance) id = 8 corresponds to α (exponent of 1/distance for immigration rate) id = 9 corresponds to λ^{a0} (anagenesis rate for unit distance) id = 10 corresponds to β (exponent of 1/distance for anagenesis rate) id = 11 corresponds to d_0 in models M15 to M19, and models with <code>distance_dep = 'sigmoidal_col'</code>, <code>'sigmoidal_ana'</code> or <code>'sigmoidal_clado'</code>; or d_0 for colonisation (when specifying <code>distance_dep = 'sigmoidal_col_ana'</code>)

id = 12 corresponds to d0 for anagenesis when specifying distance_dep = 'sigmoidal_col_ana'

parsfix	The values of the parameters that should not be optimized
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.
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
island_ontogeny	type of island ontogeny. If NA, then constant ontogeny is assumed
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.
CS_version	For internal testing purposes only. Default is 1, the original DAISIE code.
verbose	sets whether parameters and likelihood should be printed (1) or not (0)
tolint	Vector of two elements containing the absolute and relative tolerance of the integration
distance_type	Use 'continent' if the distance to the continent should be used, use 'nearest_big' if the distance to the nearest big landmass should be used, and use 'biologically_realistic' if the distance should take into account some biologically realism, e.g. an average of the previous two if both are thought to contribute.
distance_dep	Sets what type of distance dependence should be used. Default is a power law, denoted as 'power' (models M1-14 in Valente et al 2020). Alternatives are additive or interactive contributions of distance and area to the rate of cladogenesis ("area_additive_clado"; "area_interactive_clado", "area_interactive_clado1" and "area_interactive_clado2"). Other alternatives are exponential relationship

denoted by 'exp'; or sigmoids, either 'sigmoidal_col' for a sigmoid in the colonization, 'sigmoidal_ana' for sigmoidal anagenesis, 'sigmoidal_clado' for sigmoidal cladogenesis, and 'sigmoidal_col_ana' for sigmoids in both colonization and anagenesis.

A key for the different options of distance_dep that should be specified to run the models from Valente et al 2020 (Supplementary Data Table 1 and 2) is given below:

- * M1 to M14 - 'power'
- * M15 - 'area_additive_clado'
- * M16 and M19 - 'area_interactive_clado'
- * M17 - 'area_interactive_clado1'
- * M18 - 'area_interactive_clado2'
- * M20 and M24 - 'sigmoidal_col'
- * M21, M25 and M28 - 'sigmoidal_ana'
- * M22 and M26 - 'sigmoidal_clado'
- * M23 and M27 - 'sigmoidal_col_ana'

parallel	Sets whether parallel computation should be used. Use 'no' if no parallel computing should be used, 'cluster' for parallel computing on a unix/linux cluster, and 'local' for parallel computation on a local machine.
cpus	Number of cpus used in parallel computing. Default is 3. Will not have an effect if parallel = 'no'.

Value

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

lambda_c0	gives the maximum likelihood estimate of λ^c , the rate of cladogenesis for unit area
y	gives the maximum likelihood estimate of y , the exponent of area for the rate of cladogenesis
mu0	gives the maximum likelihood estimate of μ_0 , the extinction rate
x	gives the maximum likelihood estimate of x , the exponent of $1/\text{area}$ for the extinction rate
K0	gives the maximum likelihood estimate of K_0 , the carrying-capacity for unit area
z	gives the maximum likelihood estimate of z , the exponent of area for the carrying capacity
gamma0	gives the maximum likelihood estimate of γ_0 , the immigration rate for unit distance
y	gives the maximum likelihood estimate of α , the exponent of $1/\text{distance}$ for the rate of colonization
lambda_a0	gives the maximum likelihood estimate of λ^a , the rate of anagenesis for unit distance
beta	gives the maximum likelihood estimate of β , the exponent of $1/\text{distance}$ for the rate of anagenesis

loglik	gives the maximum loglikelihood
df	gives the number of estimated parameters, i.e. degrees of freedom
conv	gives a message on convergence of optimization; conv = 0 means convergence

Author(s)

Rampal S. Etienne & Luis Valente

References

Valente L, Phillimore AB, Melo M, Warren BH, Clegg SM, Havenstein K, Tiedemann R, Illera JC, Thébaud C, Aschenbach T, Etienne RS. A simple dynamic model explains island bird diversity worldwide (2020) *Nature*, 579, 92-96

See Also

[DAISIE_ML_CS](#),

Examples

```
cat("
### Fit the M19 model as in Valente et al 2020, using the ML
parameters as starting values (see Supplementary Tables 1 and 2).

utils::data(archipelagos41)

DAISIE_MW_ML(
  datalist= archipelagos41,
  initparsopt =
  c(0.040073803,1.945656546,0.150429656,
  67.25643672,0.293635061,0.059096872,0.382688527,
  0.026510781),
  idparsopt = c(1,3,4,7,8,9,10,11),
  parsfix = c(0,Inf,0) ,
  idparsfix = c(2,5,6),
  res = 100,
  ddmodel = 0,
  methode = 'lsodes',
  cpus = 4,
  parallel = 'local',
  optimmethod = 'subplex',
  tol = c(1E-4, 1E-5, 1E-7),
  distance_type = 'continent',
  distance_dep = 'area_interactive_clado'
)
")
```

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 λ^c (cladogenesis rate) pars1[2] corresponds to μ (extinction rate) pars1[3] corresponds to K (clade-level carrying capacity) pars1[4] corresponds to γ (immigration rate) pars1[5] corresponds to λ^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. In DAISIE_probdist() or DAISIE_margprobdist() either this or <code>initprobs</code> must be NULL. In DAISIE_numcol() 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

numcol <- 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 <code>Galapagos_datalist</code> and <code>DAISIE_data_prep</code> for examples) or in table format (see <code>Galapagos_datatable</code> for an example).
title	Title of the plot
island_age	Age of island in appropriate units. In <code>DAISIE_plot_age_diversity()</code> and <code>DAISIE_plot_island()</code> 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_constant_rate](#), [DAISIE_sim_time_dependent](#), [DAISIE_sim_constant_rate_shift](#), [DAISIE_plot_island](#)

Examples

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

DAISIE_plot_input *DAISIE tree plot*

Description

Shows the phylogenies of the multiple clades present on the island

Usage

```
DAISIE_plot_input(  
  trees,  
  age = NULL,  
  tcols = NULL,  
  metadata = NULL,  
  mapping = NULL,  
  xlen = 0.001,  
  pargs = NULL,  
  bckgd = "white"  
)
```

Arguments

trees	A list of trees, in phylo format, named after their respective clades
age	The age of the island, on the same scale as the branch lengths of the trees. If unspecified, the depth of the deepest stem across trees.
tcols	A vector of colonization times, with one value per clade. Order must be the same as in 'trees'. If unspecified, all clades are assigned the age of the island as colonization time.
metadata	Optional data frame with clade-level metadata. One column must be named "clade".
mapping	Optional aesthetic mapping to apply to the trees, as returned by the 'ggplot2::aes' function. Mapped variables can be anything in the columns of the 'data' node-wise data frame associated to the 'ggtree' plot being created (e.g. node, label, clade, mrca) or anything in the columns of the clade-wise 'metadata', if provided (in this case the aesthetics is mapped to all nodes within each clade).
xlen	Length of the extra tips grafted to each tree at the island age. These are a hack for scaling the plot. Keep this value small.
pargs	Optional arguments to be passed to 'geom_point' when plotting points at colonization events (e.g. size, shape...).
bckgd	Optional background color of the figure. This is because we use rectangles as a hack to hide tree branches prior to island colonization. Default to white background.

Value

A 'ggtree' plot, which is also a 'ggplot' object. The output is fully customizable, as any 'ggplot' object.

Author(s)

Raphael Scherrer (github.com/rscherrer)

Examples

```

# check whether package tibble has been installed
if (!requireNamespace('tibble', quietly = TRUE)) {
  cat("Package tibble needed for this function to work. Please install it.")
} else {
  set.seed(42)

  # Random trees
  t1 <- ape::rtree(10)
  t1$tip.label <- gsub("t", "t1.", t1$tip.label)
  t2 <- ape::rtree(3)
  t2$tip.label <- gsub("t", "t2.", t2$tip.label)
  t3 <- DAISIE::DAISIE_single_branch(
    "t3.1",
    edge.length = 4.6
  ) # tree with one species
  trees <- list(t1, t2, t3)
  names(trees) <- c("A", "B", "C")

  # Toy colonization events for each clade
  tcols <- c(4.5, 5, 4.6)

  # Toy metadata
  metadata <- tibble::tibble(
    clade = names(trees),
    endemic = TRUE, # whether each clade is endemic
    uncertain = FALSE # whether colonization time is known for sure
  )
  metadata$endemic[3] <- FALSE
  metadata$uncertain[2] <- TRUE

  # Island age
  age <- 5

  # Make a plot
  p <- DAISIE_plot_input(
    trees,
    age,
    tcols,
    metadata,
    mapping = ggplot2::aes(color = endemic, linetype = uncertain),
    pargs = list(size = 3)
  )
  p
}

```

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 <code>Galapagos_datalist</code> and <code>DAISIE_data_prep</code> for examples) or in table format (see <code>Galapagos_datatable</code> for an example).
island_age	Age of island in appropriate units. In <code>DAISIE_plot_age_diversity()</code> and <code>DAISIE_plot_island()</code> 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_constant_rate](#), [DAISIE_sim_time_dependent](#), [DAISIE_sim_constant_rate_shift](#), [DAISIE_plot_age_diversity](#)

Examples

```
### Plot Galapagos islands dataset from data table (specify island age)

utils::data(Galapagos_datatable)
DAISIE_plot_island(Galapagos_datatable, island_age=4)
```

```
### Plot Galapagos islands dataset from datalist (no need to specify
## island age)
utils::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.

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.

Usage

```
DAISIE_plot_sims(
  island_replicates,
  plot_plus_one = TRUE,
  type = "all_species",
  sample_freq = 25,
  trait_pars = NULL
)
```

Arguments

island_replicates	List output from DAISIE_sim_core_constant_rate() , DAISIE_sim_core_time_dependent() , DAISIE_sim_core_constant_rate_shift() or DAISIE_sim_min_type2() functions. Minimally, this must be a list that has as many elements as replicates. Each element must be a list with the elements <code>island_age</code> , <code>not_present</code> and <code>stt_all</code> . <code>stt_all</code> must be a data frame with the column names <code>Time</code> , <code>nI</code> , <code>nA</code> , <code>nC</code> and <code>present</code> .
plot_plus_one	Boolean to indicate to plot all values plus one. Set to <code>TRUE</code> for default behavior. Set to <code>FALSE</code> to plot all values without adding one. Only works when there is one type of species.
type	String to indicate if stt of all species or all possible stt should be plotted. Default is <code>"all_species"</code> , <code>"type1_species"</code> or <code>"type2_species"</code> should be plotted.
sample_freq	Numeric specifying the number of units times should be divided by for plotting purposes. Larger values will lead to plots with higher resolution, but will also run slower.

`trait_pars` A named list containing diversification rates considering two trait states created by `create_trait_pars`:

- [1]:A numeric with the per capita transition rate with state1
- [2]:A numeric with the per capita immigration rate with state2
- [3]:A numeric with the per capita extinction rate with state2
- [4]:A numeric with the per capita anagenesis rate with state2
- [5]:A numeric with the per capita cladogenesis rate with state2
- [6]:A numeric with the per capita transition rate with state2
- [7]:A numeric with the number of species with trait state 2 on mainland

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_constant_rate](#), [DAISIE_sim_time_dependent](#), [DAISIE_sim_constant_rate_shift](#), [DAISIE_format_CS](#)

Examples

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

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

DAISIE_probdist	<i>The joint distribution of endemics and non-endemics under the DAISIE model</i>
-----------------	---

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 λ^c (cladogenesis rate) pars1[2] corresponds to μ (extinction rate) pars1[3] corresponds to K (clade-level carrying capacity) pars1[4] corresponds to γ (immigration rate) pars1[5] corresponds to λ^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. In <code>DAISIE_probdist()</code> or <code>DAISIE_margprobdist()</code> either this or <code>initprobs</code> must be <code>NULL</code> . In <code>DAISIE_numcol()</code> when it is <code>NULL</code> , it is assumed that the island is empty.
initprobs	The initial probability distribution for the number of endemics and non-endemics; either this or <code>initEI</code> must be <code>NULL</code> .

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 $1 + \text{length}(tvec)$ and $\text{pars}[2]^2 + 1$ 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

prob_dists <- 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 (non-)oceanic islands with given parameters under time-constant rates

Description

This function simulates islands with given cladogenesis, extinction, Kprime, immigration and anagenesis parameters, all of which modelled as time-constant parameters. If a single parameter set is provided (5 parameters) it simulates islands where all species have the same macro-evolutionary process. If two parameter 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. Further, it allows for the simulation of non-oceanic islands, generating islands for which the starting condition includes potential endemic and non-endemic species.

Usage

```
DAISIE_sim_constant_rate(
  time,
  M,
  pars,
  replicates,
  divdepmodel = "CS",
  nonoceanic_pars = c(0, 0),
```

```

num_guilds = NULL,
prop_type2_pool = NA,
replicates_apply_type2 = TRUE,
sample_freq = 25,
plot_sims = TRUE,
hyper_pars = create_hyper_pars(d = 0, x = 0),
area_pars = create_area_pars(max_area = 1, current_area = 1, proportional_peak_t = 0,
  total_island_age = 0, sea_level_amplitude = 0, sea_level_frequency = 0,
  island_gradient_angle = 0),
verbose = TRUE,
...
)

```

Arguments

time	Numeric defining the length of the simulation in time units. For example, if an island is known to be 4 million years old, setting time = 4 will simulate the entire life span of the island; setting time = 2 will stop the simulation at the mid-life of the island.
M	Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island.
pars	<p>A numeric vector containing the model parameters:</p> <ul style="list-style-type: none"> • pars[1]: λ^c (cladogenesis rate) • pars[2]: μ (extinction rate) • pars[3]: K (carrying capacity), set K=Inf for diversity independence. • pars[4]: γ (immigration rate) • pars[5]: λ^a (anagenesis rate) • pars[6]: λ^c (cladogenesis rate) for either type 2 species or rate set 2 in rate shift model • pars[7]: μ (extinction rate) for either type 2 species or rate set 2 in rate shift model • pars[8]: K (carrying capacity) for either type 2 species or rate set 2 in rate shift model, set K=Inf for diversity independence. • pars[9]: γ (immigration rate) for either type 2 species or rate set 2 in rate shift model • pars[10]: λ^a (anagenesis rate) for either type 2 species or rate set 2 in rate shift model <p>Elements 6:10 are required only when type 2 species are included or in the rate shift model. For <code>DAISIE_sim_relaxed_rate()</code> pars[6] is the standard deviation of the gamma distribution for the relaxed parameter and the parameter chosen by the <code>relaxed_par</code> argument is the mean of the gamma distribution for the relaxed parameter.</p>
replicates	Integer specifying number of island replicates to be simulated.
divdepmodel	Option <code>divdepmodel = 'CS'</code> runs a 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 a model with island-wide carrying capacity, where diversity-dependence operates within and among clades. Option `divdepmodel = 'GW'` runs a model with diversity-dependence operates within a guild.

<code>nonoceanic_pars</code>	A vector of length two with: <ul style="list-style-type: none"> • [1]: the probability of sampling a species from the mainland • [2]: the probability of the species sampled from the mainland being nonendemic
<code>num_guilds</code>	The number of guilds on the mainland. The number of mainland species is divided by the number of guilds when <code>divdepmodel = "GW"</code>
<code>prop_type2_pool</code>	Fraction of mainland species that belongs to the second subset of species (type 2). Applies only when two types of species are simulated (<code>length(pars) = 10</code>). For <code>DAISIE_dataprep()</code> applies only if <code>number_clade_types = 2</code> . In <code>DAISIE_dataprep()</code> the default "proportional" sets the fraction to be proportional to the number of clades of distinct macroevolutionary process that have colonised the island.
<code>replicates_apply_type2</code>	Applies only when two types of species are being simulated. Default <code>replicates_apply_type2 = TRUE</code> 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 <code>prop_type2_pool</code> is small or if the rate of immigration of type two species (<code>pars[9]</code>) is low, meaning that more replicates are needed to achieved an adequate sample size of islands with type 2 species. Setting <code>replicates_apply_type2 = FALSE</code> simulates islands up to the specified number of replicates regardless of whether type 2 species have colonised or not.
<code>sample_freq</code>	Numeric specifying the number of units times should be divided by for plotting purposes. Larger values will lead to plots with higher resolution, but will also run slower.
<code>plot_sims</code>	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.
<code>hyper_pars</code>	A named list of numeric hyperparameters for the rate calculations as returned by <code>create_hyper_pars()</code> : <ul style="list-style-type: none"> • [1]: is d the scaling parameter for exponent for calculating cladogenesis rate • [2]: is x the exponent for calculating extinction rate
<code>area_pars</code>	A named list containing area and sea level parameters as created by <code>create_area_pars()</code> : <ul style="list-style-type: none"> • [1]: maximum area • [2]: current area • [3]: value from 0 to 1 indicating where in the island's history the peak area is achieved • [4]: total island age • [5]: amplitude of area fluctuation from sea level

	<ul style="list-style-type: none"> • [6]: frequency of sine wave of area change from sea level • [7]: angle of the slope of the island
verbose	In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
...	Any arguments to pass on to plotting functions.

Value

A list. The highest level of the list corresponds to each individual replicate. The first element of each replicate is composed of island information containing:

- `$island_age`: A numeric with the island age.
- `$not_present`: the number of mainland lineages that are not present on the island. It is only present if only 1 type of species is simulated. Becomes `$not_present_type1`: the number of mainland lineages of type 1 that are not present on the island and `$not_present_type2`: the number of mainland lineages of type 2 that are not present on the island, if two types are simulated.
- `$stt_all`: STT table for all species on the island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)
- `$stt_stt_type1`: STT table for type 1 species on the island - only if 2 types of species were simulated (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present).
- `$stt_stt_type2`: STT table for type 2 species on the island - only if 2 types of species were simulated (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present).
- `$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 pertaining to each replicate contain information on a single colonist lineage on the island and have 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`: An integer ranging from 1 to 4 indicating the status of the colonist:
 1. Non_endemic_MaxAge
 2. Endemic
 3. Endemic&Non_Endemic
 4. Non_endemic_MaxAge
- `$missing_species`: number of island species that were not sampled for particular clade (only applicable for endemic clades)
- `$type_1or2`: whether the colonist belongs to type 1 or type 2

Author(s)

Luis Valente, Albert Phillimore, Joshua Lambert, Shu Xie, Pedro Neves, Richèl J. C. Bilderbeek, Rupal 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.

See Also

[DAISIE_plot_sims\(\)](#) for plotting STT of simulation outputs.

Other simulation models: [DAISIE_sim_constant_rate_shift\(\)](#), [DAISIE_sim_relaxed_rate\(\)](#), [DAISIE_sim_time_dependent\(\)](#)

Examples

```
## Simulate 2 islands for 1 million years, where all species have equal
## rates. Pool size 100.

clado_rate <- 0.5
ext_rate <- 0.2
carr_cap <- Inf
immig_rate <- 0.005
ana_rate <- 1
sim_pars <- c(clado_rate, ext_rate, carr_cap, immig_rate, ana_rate)
set.seed(1)
island_replicates <- DAISIE_sim_constant_rate(
  time = 1,
  M = 100,
  pars = sim_pars,
  replicates = 2,
  plot_sims = FALSE,
  verbose = FALSE
)

## Simulate 2 islands for 1 million years with two types of species (type1
## and type 2). Pool size 100
## Fraction of type 2 species in source pool is 0.15. Function will
## simulate until number of islands where type 2 species has colonised is
## equal to number specified in replicates.

clado_rate <- 0.5
ext_rate <- 0.2
carr_cap <- Inf
immig_rate <- 0.005
ana_rate <- 1
sim_pars_type1 <- c(clado_rate, ext_rate, carr_cap, immig_rate, ana_rate)
sim_pars_type2 <- sim_pars_type1 * 2
set.seed(1)
island_replicates_2types <- DAISIE_sim_constant_rate(
```

```

time = 1,
M = 100,
pars = c(sim_pars_type1, sim_pars_type2),
replicates = 2,
prop_type2_pool = 0.15,
plot_sims = FALSE,
verbose = FALSE
)
## Simulate two non-oceanic island for 1 million years.
## Pool size 500. Island area as a proportion
## of mainland is 0.1, proportion of native species is 0.9.
clado_rate <- 0.5
ext_rate <- 0.2
carr_cap <- Inf
immig_rate <- 0.005
ana_rate <- 1
sim_pars <- c(clado_rate, ext_rate, carr_cap, immig_rate, ana_rate)
set.seed(1)
island_replicates <- DAISIE_sim_constant_rate(
  time = 1,
  M = 500,
  pars = sim_pars,
  replicates = 2,
  nonoceanic_pars = c(0.1, 0.9),
  plot_sims = FALSE,
  verbose = FALSE
)

## Simulate 2 islands for 1 million years with a shift in immigration rate
## at 0.195 Ma, and plot the species-through-time plot. Pool size 296.

pars_before_shift <- c(0.079, 0.973, Inf, 0.136, 0.413)
pars_after_shift <- c(0.079, 0.973, Inf, 0.652, 0.413)
tshift <- 0.195
set.seed(1)
island_shift_replicates <- DAISIE_sim_constant_rate_shift(
  time = 1,
  M = 296,
  pars = c(pars_before_shift, pars_after_shift),
  replicates = 2,
  shift_times = tshift,
  plot_sims = FALSE,
  verbose = FALSE
)

```

DAISIE_sim_constant_rate_shift

Simulate (non-)oceanic islands with given parameters under a rate-shift regime

Description

This function simulates islands with given cladogenesis, extinction, K prime, immigration and anagenesis parameters, all of which modelled as time-constant parameters, which can be switched to a different diversification regime (i.e., different set of parameters) at one or more set times before the present. Further, it allows for the simulation of non-oceanic islands, generating islands for which the starting condition includes potential endemic and non-endemic species.

Usage

```
DAISIE_sim_constant_rate_shift(
  time,
  M,
  pars,
  replicates,
  shift_times,
  divdepmode = "CS",
  nonoceanic_pars = c(0, 0),
  num_guilds = NULL,
  sample_freq = 25,
  plot_sims = TRUE,
  hyper_pars = create_hyper_pars(d = 0, x = 0),
  area_pars = DAISIE::create_area_pars(max_area = 1, current_area = 1,
    proportional_peak_t = 0, total_island_age = 0, sea_level_amplitude = 0,
    sea_level_frequency = 0, island_gradient_angle = 0),
  verbose = TRUE,
  ...
)
```

Arguments

time	Numeric defining the length of the simulation in time units. For example, if an island is known to be 4 million years old, setting time = 4 will simulate the entire life span of the island; setting time = 2 will stop the simulation at the mid-life of the island.
M	Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island.
pars	A numeric vector containing the model parameters: <ul style="list-style-type: none"> • pars[1]: λ^c (cladogenesis rate) • pars[2]: μ (extinction rate) • pars[3]: K (carrying capacity), set $K=Inf$ for diversity independence. • pars[4]: γ (immigration rate) • pars[5]: λ^a (anagenesis rate) • pars[6]: λ^c (cladogenesis rate) for either type 2 species or rate set 2 in rate shift model • pars[7]: μ (extinction rate) for either type 2 species or rate set 2 in rate shift model

- pars[8]: K (carrying capacity) for either type 2 species or rate set 2 in rate shift model, set K=Inf for diversity independence.
- pars[9]: gamma (immigration rate) for either type 2 species or rate set 2 in rate shift model
- pars[10]: lambda^a (anagenesis rate) for either type 2 species or rate set 2 in rate shift model

Elements 6:10 are required only when type 2 species are included or in the rate shift model. For `DAISIE_sim_relaxed_rate()` pars[6] is the standard deviation of the gamma distribution for the relaxed parameter and the parameter chosen by the `relaxed_par` argument is the mean of the gamma distribution for the relaxed parameter.

<code>replicates</code>	Integer specifying number of island replicates to be simulated.
<code>shift_times</code>	a numeric vector specifying when the rate shifts occur before the present.
<code>divdepmode</code>	Option <code>divdepmode = 'CS'</code> runs a 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 <code>divdepmode = 'IW'</code> runs a model with island-wide carrying capacity, where diversity-dependence operates within and among clades. Option <code>divdepmode = 'GW'</code> runs a model with diversity-dependence operates within a guild.
<code>nonoceanic_pars</code>	A vector of length two with: <ul style="list-style-type: none"> • [1]: the probability of sampling a species from the mainland • [2]: the probability of the species sampled from the mainland being nonendemic
<code>num_guilds</code>	The number of guilds on the mainland. The number of mainland species is divided by the number of guilds when <code>divdepmode = "GW"</code>
<code>sample_freq</code>	Numeric specifying the number of units times should be divided by for plotting purposes. Larger values will lead to plots with higher resolution, but will also run slower.
<code>plot_sims</code>	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.
<code>hyper_pars</code>	A named list of numeric hyperparameters for the rate calculations as returned by <code>create_hyper_pars()</code> : <ul style="list-style-type: none"> • [1]: is d the scaling parameter for exponent for calculating cladogenesis rate • [2]: is x the exponent for calculating extinction rate
<code>area_pars</code>	A named list containing area and sea level parameters as created by <code>create_area_pars()</code> : <ul style="list-style-type: none"> • [1]: maximum area • [2]: current area • [3]: value from 0 to 1 indicating where in the island's history the peak area is achieved • [4]: total island age

- [5]: amplitude of area fluctuation from sea level
 - [6]: frequency of sine wave of area change from sea level
 - [7]: angle of the slope of the island
- verbose In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
- ... Any arguments to pass on to plotting functions.

Value

A list. The highest level of the least corresponds to each individual replicate. The first element of each replicate is composed of island information containing:

- \$island_age: A numeric with the island age.
- \$not_present: A numeric with 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)
- \$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 pertaining to each replicate contain information on a single colonist lineage on the island and have 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: An integer ranging from 1 to 4 indicating the status of the colonist:
 1. Non_endemic_MaxAge
 2. Endemic
 3. Endemic&Non_Endemic
 4. Non_endemic_MaxAge
- \$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 Valente, Albert Phillimore, Torsten Hauffe

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.

Hauffe, T., D. Delicado, R.S. Etienne and L. Valente (2020). Lake expansion elevates equilibrium diversity via increasing colonization.

See Also

[DAISIE_plot_sims\(\)](#) for plotting STT of simulation outputs.

Other simulation models: [DAISIE_sim_relaxed_rate\(\)](#), [DAISIE_sim_time_dependent\(\)](#), [DAISIE_sim\(\)](#)

DAISIE_sim_MW

Simulate multiple islands based on hyperparameters that describe relationships between area, isolation and local parameters

Description

This function simulates islands of given age, area and isolation. A list of islands/archipelagos with this information is provided as a data frame object (in the example, the `archipelago_data` from Valente et al 2020 Nature). For each island, local parameters are first calculated based on hyperparameters describing the dependency of rates on island area and isolation. Simulations are then run for each island, from island birth until the given island age based on the local parameters for each island (using the `DAISIE_sim` function).

Returns R list object that contains the simulated islands.

Usage

```
DAISIE_sim_MW(
  archipelago_data,
  M,
  pars,
  replicates,
  divdepmodel = "CS",
  distance_dep = "power",
  cladogenesis_dep = "NULL",
  sigmoidal_par = "NULL"
)
```

Arguments

`archipelago_data`

Object (data.frame) containing a table with columns with the following names: Archipelago: names of the archipelagos or islands to be simulated; Area - the area of the archipelago/island; Age: Age of the oldest island in the archipelago, or age of the island if single island. Distance - distance to the mainland (or other applicable isolation metric). As an example, the dataset from Valente et al 2020 for 41 archipelagos is provided in `data(archipelago_data)`.

M	The size of the mainland pool, i.e the number of species that can potentially colonize the island(s).
pars	<p>Contains the model hyperparameters:</p> <p>pars[1] corresponds to λ^c_0 (initial cladogenesis rate)</p> <p>pars[2] corresponds to γ (dependency of cladogenesis on area) 0 - no dependency, positive value - cladogenesis increases with area, negative value - cladogenesis decreases with area</p> <p>pars[3] corresponds to μ_0 (initial extinction rate)</p> <p>pars[4] corresponds to α (dependency of extinction on area) 0 - no dependency, positive value - extinction increases with area, negative value - extinction decreases with area.</p> <p>pars[5] corresponds to K_0 (initial carrying capacity). Set $K=\text{Inf}$ for non-diversity dependence.</p> <p>pars[6] corresponds to β (dependency of K on area). 0 - no dependency, positive value - K increases with area, negative value - K decreases with area.</p> <p>pars[7] corresponds to γ_0 (initial immigration rate)</p> <p>pars[8] corresponds to δ (dependency of immigration on island isolation). 0 - no dependency, positive value - immigration increases with distance, negative value - immigration decreases with distance</p> <p>pars[9] corresponds to λ^a_0 (initial anagenesis rate)</p> <p>pars[10] corresponds to β (dependency of anagenesis on island isolation). 0 - no dependency, positive value - anagenesis increases with distance, negative value - anagenesis decreases with distance.</p> <p>pars[11] corresponds to d_0, parameter that describes interactive/additive effect of isolation and area on cladogenesis (power models); or that describes the shape of the sigmoidal relationship between isolation and a parameter (sigmoidal models).</p>
replicates	Number of island replicates to be simulated per island in the table archipelago_data. If there are 5 islands in archipelago_data, and number of replicates is set to 10, 50 islands will be simulated.
divdepmodel	<p>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.</p> <p>Option divdepmodel = 'IW' runs model with island-wide carrying capacity, where diversity-dependence operates within and among clades.</p>
distance_dep	Shape of dependency of parameter with distance. Can be of type 'power' or type 'sigmoidal'.
cladogenesis_dep	<p>When distance_dep = 'power' is specified, cladogenesis_dep specifies the type of effect of area and isolation on cladogenesis. The possible options are:</p> <p>'NULL' (no additive or interactive effect of area on isolation, same as models M1-M14 in Valente et al 2020)</p> <p>'additive' (additive effect of area and isolation on cladogenesis, same as model M15 in Valente et al 2020)</p> <p>'interactive' (interactive effect of area and isolation on cladogenesis, same as</p>

models M16 and M19 in Valente et al 2020)
 'interactive1' (interactive effect of area and isolation on cladogenesis, same as model M17 in Valente et al 2020)
 'interactive2' (interactive effect of area and isolation on cladogenesis, same as model M18 in Valente et al 2020)

sigmoidal_par When distance_dep = 'sigmoidal', sigmoidal_par specifies to which parameter the sigmoidal relationship with distance is applied. Options:
 'cladogenesis'
 'anagenesis'
 'colonisation'.

Value

For each island listed in the rows of `archipelago_data`, a given number of islands is simulated depending on the number of replicates specified. The simulations for each island in `archipelago_data` are an element of the list, which can be called using `[[x]]`. Individual replicates for an island can be called with `[[x]][[x]]`. For example if the object is called `global_sims`, the 1st replicate for the first island can be called using `global_sims[[1]][[1]]`. The 3rd replicate of the 4th island is called with `global_sims[[4]][[3]]`. Each of the island replicates is a list in itself. The first (e.g. `global_sims[[x]][[x]][[1]]`) element of that list has the following components:

`$island_age` - the island or archipelago age
`$not_present` - the number of mainland lineages that are not present on the island
`$stt_all` - STT table for all species on the archipelago/island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)
`$area` Area of the island or archipelago
`$distance` Distance to the mainland
`$name` Name of the island or archipelago

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 with no close extant relatives on the islands 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

* 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)

Author(s)

Luis Valente, Albert Phillimore, Rampal Etienne

References

Valente, LM, Phillimore AB, Melo M, Warren B, Clegg S, Havenstein K, Tiedemann R, Illera JC, Thebaud C, Aschenbach T and Etienne RS (2020). A simple dynamic model explain island bird diversity worldwide. *Nature*, 579, 92-96.

See Also

[DAISIE_sim_constant_rate\(\)](#).

Examples

```
## Simulate 10 replicates for the each of the 41 archipelagos in the archipelago_data table,
##based on the hyperparameters of the M19 model (preferred model in Valente et al 2020)
## and the age, area and isolation values given in the archipelago_data table.
## Mainland pool size of 1000, clade-specific carrying capacity. The M19 model is a
## power model with an interactive effect of area and isolation on cladogenesis.
##
## data(archipelago_data)
## result <- DAISIE_sim_MW(
## archipelago_data = archipelago_data,
## M = 1000,
## pars = c(0.040073803,0,1.945656546,0.150429656,Inf,0,67.25643672,
## 0.293635061,0.059096872,0.382688527,0.026510781),
## replicates = 10,
## distance_dep = 'power',
## cladogenesis_dep = 'interactive',
## sigmoidal_par = 'NULL',
## divdepmodel = 'CS')
```

DAISIE_sim_relaxed_rate

Simulate (non-)oceanic islands with given parameters under a relaxed-rate model

Description

This function simulates islands with given cladogenesis, extinction, Kprime, immigration and anagenesis parameters, all of which can be modelled as time-constant parameters with variation between clades in one or multiple parameters. Further, it allows for the simulation of non-oceanic islands, generating islands for which the starting condition includes potential endemic and non-endemic species.

Usage

```
DAISIE_sim_relaxed_rate(
  time,
  M,
  pars,
```

```

replicates,
relaxed_par,
nonoceanic_pars = c(0, 0),
sample_freq = 25,
plot_sims = TRUE,
hyper_pars = create_hyper_pars(d = 0, x = 0),
area_pars = create_area_pars(max_area = 1, current_area = 1, proportional_peak_t = 0,
  total_island_age = 0, sea_level_amplitude = 0, sea_level_frequency = 0,
  island_gradient_angle = 0),
verbose = TRUE,
...
)

```

Arguments

time	Numeric defining the length of the simulation in time units. For example, if an island is known to be 4 million years old, setting time = 4 will simulate the entire life span of the island; setting time = 2 will stop the simulation at the mid-life of the island.
M	Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island.
pars	<p>A numeric vector containing the model parameters:</p> <ul style="list-style-type: none"> • pars[1]: λ^c (cladogenesis rate) • pars[2]: μ (extinction rate) • pars[3]: K (carrying capacity), set K=Inf for diversity independence. • pars[4]: γ (immigration rate) • pars[5]: λ^a (anagenesis rate) • pars[6]: λ^c (cladogenesis rate) for either type 2 species or rate set 2 in rate shift model • pars[7]: μ (extinction rate) for either type 2 species or rate set 2 in rate shift model • pars[8]: K (carrying capacity) for either type 2 species or rate set 2 in rate shift model, set K=Inf for diversity independence. • pars[9]: γ (immigration rate) for either type 2 species or rate set 2 in rate shift model • pars[10]: λ^a (anagenesis rate) for either type 2 species or rate set 2 in rate shift model <p>Elements 6:10 are required only when type 2 species are included or in the rate shift model. For <code>DAISIE_sim_relaxed_rate()</code> pars[6] is the standard deviation of the gamma distribution for the relaxed parameter and the parameter chosen by the <code>relaxed_par</code> argument is the mean of the gamma distribution for the relaxed parameter.</p>
replicates	Integer specifying number of island replicates to be simulated.
relaxed_par	A string determining which parameter is relaxed in a relaxed rate model.
nonoceanic_pars	A vector of length two with:

	<ul style="list-style-type: none"> • [1]: the probability of sampling a species from the mainland • [2]: the probability of the species sampled from the mainland being nonendemic
sample_freq	Numeric specifying the number of units times should be divided by for plotting purposes. Larger values will lead to plots with higher resolution, 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.
hyper_pars	A named list of numeric hyperparameters for the rate calculations as returned by <code>create_hyper_pars()</code> : <ul style="list-style-type: none"> • [1]: is d the scaling parameter for exponent for calculating cladogenesis rate • [2]: is x the exponent for calculating extinction rate
area_pars	A named list containing area and sea level parameters as created by <code>create_area_pars()</code> : <ul style="list-style-type: none"> • [1]: maximum area • [2]: current area • [3]: value from 0 to 1 indicating where in the island's history the peak area is achieved • [4]: total island age • [5]: amplitude of area fluctuation from sea level • [6]: frequency of sine wave of area change from sea level • [7]: angle of the slope of the island
verbose	In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
...	Any arguments to pass on to plotting functions.

Value

A list. The highest level of the least corresponds to each individual replicate. The first element of each replicate is composed of island information containing:

- `$island_age`: A numeric with the island age.
- `$not_present`: the number of mainland lineages that are not present on the island. It is only present if only 1 type of species is simulated. Becomes `$not_present_type1`: the number of mainland lineages of type 1 that are not present on the island and `$not_present_type2`: the number of mainland lineages of type 2 that are not present on the island, if two types are simulated.
- `$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)

- `$stt_stt_type1`: STT table for type 1 species on the island - only if 2 types of species were simulated (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present).
- `$stt_stt_type2`: STT table for type 2 species on the island - only if 2 types of species were simulated (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 pertaining to each replicate contain information on a single colonist lineage on the island and have 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`: An integer ranging from 1 to 4 indicating the status of the colonist:
 1. Non_endemic_MaxAge
 2. Endemic
 3. Endemic&Non_Endemic
 4. Non_endemic_MaxAge
- `$missing_species`: number of island species that were not sampled for particular clade (only applicable for endemic clades)
- `$type_1or2`: whether the colonist belongs to type 1 or type 2

Author(s)

Luis Valente, Albert Phillimore, Joshua Lambert, Shu Xie, Pedro Neves, Richèl J. C. Bilderbeek, Rupal Etienne

See Also

[DAISIE_plot_sims\(\)](#) for plotting STT of simulation outputs.

Other simulation models: [DAISIE_sim_constant_rate_shift\(\)](#), [DAISIE_sim_time_dependent\(\)](#), [DAISIE_sim\(\)](#)

Examples

```
## Simulate an island for 1 million years, with a relaxed the rate of
## cladogenesis between clades. Pool size 500.

clado_rate <- 0.5
ext_rate <- 0.2
carr_cap <- Inf
immig_rate <- 0.005
ana_rate <- 1
sd <- 1
sim_pars <- c(clado_rate, ext_rate, carr_cap, immig_rate, ana_rate, sd)
```

```

set.seed(1)
island_replicates <- DAISIE_sim_relaxed_rate(
  time = 1,
  M = 500,
  pars = sim_pars,
  replicates = 2,
  relaxed_par = "cladogenesis",
  plot_sims = FALSE,
  verbose = FALSE
)

```

DAISIE_sim_time_dependent

Simulate (non-)oceanic islands with given parameters under a time-dependent regime

Description

This function simulates islands with given cladogenesis, extinction, Kprime, immigration and anagenesis parameters, all of which modelled as time-dependent parameters.

Time dependency aims to capture the effect of area changes islands undergo from their emergence until subsidence. Thus, oceanic, volcanic island ontogeny scenarios can be modelled (by a beta function), as well as the effect of sea level fluctuations (modelled through a sine function). See parameter entry `area_pars` for details. Both island ontogeny and sea level fluctuations are allowed to operate simultaneously.

This function also allows for the simulation of non-oceanic islands, generating islands for which the starting condition includes potential endemic and non-endemic species.

Usage

```

DAISIE_sim_time_dependent(
  time,
  M,
  pars,
  replicates,
  area_pars,
  hyper_pars,
  divdepmodel = "CS",
  nonoceanic_pars = c(0, 0),
  num_guilds = NULL,
  sample_freq = 25,
  plot_sims = TRUE,
  island_ontogeny = "const",
  sea_level = "const",
  extcutoff = 1000,
  verbose = TRUE,
)

```

...
)

Arguments

time	Numeric defining the length of the simulation in time units. For example, if an island is known to be 4 million years old, setting time = 4 will simulate the entire life span of the island; setting time = 2 will stop the simulation at the mid-life of the island.
M	Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island.
pars	<p>A numeric vector containing the model parameters:</p> <ul style="list-style-type: none"> • pars[1]: λ^c (cladogenesis rate) • pars[2]: μ (extinction rate) • pars[3]: K (carrying capacity), set K=Inf for diversity independence. • pars[4]: γ (immigration rate) • pars[5]: λ^a (anagenesis rate) • pars[6]: λ^c (cladogenesis rate) for either type 2 species or rate set 2 in rate shift model • pars[7]: μ (extinction rate) for either type 2 species or rate set 2 in rate shift model • pars[8]: K (carrying capacity) for either type 2 species or rate set 2 in rate shift model, set K=Inf for diversity independence. • pars[9]: γ (immigration rate) for either type 2 species or rate set 2 in rate shift model • pars[10]: λ^a (anagenesis rate) for either type 2 species or rate set 2 in rate shift model <p>Elements 6:10 are required only when type 2 species are included or in the rate shift model. For <code>DAISIE_sim_relaxed_rate()</code> pars[6] is the standard deviation of the gamma distribution for the relaxed parameter and the parameter chosen by the <code>relaxed_par</code> argument is the mean of the gamma distribution for the relaxed parameter.</p>
replicates	Integer specifying number of island replicates to be simulated.
area_pars	<p>A named list containing area and sea level parameters as created by <code>create_area_pars()</code>:</p> <ul style="list-style-type: none"> • [1]: maximum area • [2]: current area • [3]: value from 0 to 1 indicating where in the island's history the peak area is achieved • [4]: total island age • [5]: amplitude of area fluctuation from sea level • [6]: frequency of sine wave of area change from sea level • [7]: angle of the slope of the island
hyper_pars	A named list of numeric hyperparameters for the rate calculations as returned by <code>create_hyper_pars()</code> :

	<ul style="list-style-type: none"> • [1]: is d the scaling parameter for exponent for calculating cladogenesis rate • [2]: is x the exponent for calculating extinction rate
divdepmodel	Option <code>divdepmodel = 'CS'</code> runs a 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 <code>divdepmodel = 'IW'</code> runs a model with island-wide carrying capacity, where diversity-dependence operates within and among clades. Option <code>divdepmodel = 'GW'</code> runs a model with diversity-dependence operates within a guild.
nonoceanic_pars	A vector of length two with: <ul style="list-style-type: none"> • [1]: the probability of sampling a species from the mainland • [2]: the probability of the species sampled from the mainland being nonendemic
num_guilds	The number of guilds on the mainland. The number of mainland species is divided by the number of guilds when <code>divdepmodel = "GW"</code>
sample_freq	Numeric specifying the number of units times should be divided by for plotting purposes. Larger values will lead to plots with higher resolution, 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.
island_ontogeny	In <code>DAISIE_sim_time_dependent()</code> , <code>DAISIE_ML_CS</code> and plotting a string describing the type of island ontogeny. Can be "const", "beta" for a beta function describing area through time. String checked by <code>is_island_ontogeny_input()</code> . In all other functions a numeric describing the type of island ontogeny. Can be 0 for constant, 1 for a beta function describing area through time. In ML functions <code>island_ontogeny = NA</code> assumes constant ontogeny.
sea_level	In <code>DAISIE_sim_time_dependent()</code> and plotting a string describing the type of sea level. Can be "const" or "sine" for a sine function describing area through time. String checked by <code>is_sea_level_input()</code> . In all other functions a numeric describing the type of sea level. Can be 0 for constant, 1 for a sine function describing area through time.
extcutoff	A numeric with the cutoff for the the maximum extinction rate preventing it from being too large and slowing down simulation.
verbose	In simulation and dataprep functions a logical, Default = TRUE gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, Default = 0 does not print, verbose = 1 prints intermediate output of the parameters and loglikelihood, verbose = 2 means also intermediate progress during loglikelihood computation is shown.
...	Any arguments to pass on to plotting functions.

Value

A list. The highest level of the list corresponds to each individual replicate. The first element of each replicate is composed of island information containing:

- `$island_age`: A numeric with the island age.
- `$not_present`: A numeric with the number of mainland lineages that are not present on the island.
- `$stt_all`: STT table for all species on the island (nI - number of non-endemic species; nA - number of anagenetic species, nC - number of cladogenetic species, present - number of independent colonisations present)
- `$brts_table`: Only for simulations under "TW". Table containing information on order of events in the data, for use in maximum likelihood optimization.).

The subsequent elements of the list pertaining to each replicate contain information on a single colonist lineage on the island and have 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`: An integer ranging from 1 to 4 indicating the status of the colonist:
 1. Non_endemic_MaxAge
 2. Endemic
 3. Endemic&Non_Endemic
 4. Non_endemic_MaxAge
- `$missing_species`: number of island species that were not sampled for particular clade (only applicable for endemic clades)
- `$type_1or2`: whether the colonist belongs to type 1 or type 2

Author(s)

Luis 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.
- Valente, L.M., Etienne, R.S. and Phillimore, A.B. (2014). The effects of island ontogeny on species diversity and phylogeny. *Proceedings of the Royal Society B: Biological Sciences* 281(1784), p.20133227.

See Also

[DAISIE_plot_sims\(\)](#) for plotting STT of simulation outputs.

Other simulation models: [DAISIE_sim_constant_rate_shift\(\)](#), [DAISIE_sim_relaxed_rate\(\)](#), [DAISIE_sim\(\)](#)

 DAISIE_sim_trait_dependent

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 parameter 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. If two parameter sets and a time shift (11 parameters) are provided, it simulates islands where at the given time a shift between the parameter sets will occur.

Returns R list object that contains the simulated islands

Usage

```
DAISIE_sim_trait_dependent(
  time,
  M,
  pars,
  replicates,
  divdepmodel = "CS",
  nonoceanic_pars = c(0, 0),
  num_guilds = NULL,
  sample_freq = 25,
  plot_sims = TRUE,
  island_ontogeny = "const",
  sea_level = "const",
  hyper_pars = create_hyper_pars(d = 0, x = 0),
  area_pars = DAISIE::create_area_pars(max_area = 1, current_area = 1,
    proportional_peak_t = 0, total_island_age = 0, sea_level_amplitude = 0,
    sea_level_frequency = 0, island_gradient_angle = 0),
  extcutoff = 1000,
  verbose = TRUE,
  trait_pars = NULL,
  ...
)
```

Arguments

time	Numeric defining the length of the simulation in time units. For example, if an island is known to be 4 million years old, setting time = 4 will simulate the entire life span of the island; setting time = 2 will stop the simulation at the mid-life of the island.
M	Numeric defining the size of mainland pool, i.e. the number of species that can potentially colonize the island.

<code>pars</code>	<p>A numeric vector containing the model parameters:</p> <ul style="list-style-type: none"> • <code>pars[1]</code>: λ^c (cladogenesis rate) • <code>pars[2]</code>: μ (extinction rate) • <code>pars[3]</code>: K (carrying capacity), set $K=\text{Inf}$ for diversity independence. • <code>pars[4]</code>: γ (immigration rate) • <code>pars[5]</code>: λ^a (anagenesis rate) • <code>pars[6]</code>: λ^c (cladogenesis rate) for either type 2 species or rate set 2 in rate shift model • <code>pars[7]</code>: μ (extinction rate) for either type 2 species or rate set 2 in rate shift model • <code>pars[8]</code>: K (carrying capacity) for either type 2 species or rate set 2 in rate shift model, set $K=\text{Inf}$ for diversity independence. • <code>pars[9]</code>: γ (immigration rate) for either type 2 species or rate set 2 in rate shift model • <code>pars[10]</code>: λ^a (anagenesis rate) for either type 2 species or rate set 2 in rate shift model <p>Elements 6:10 are required only when type 2 species are included or in the rate shift model. For <code>DAISIE_sim_relaxed_rate()</code> <code>pars[6]</code> is the standard deviation of the gamma distribution for the relaxed parameter and the parameter chosen by the <code>relaxed_par</code> argument is the mean of the gamma distribution for the relaxed parameter.</p>
<code>replicates</code>	Integer specifying number of island replicates to be simulated.
<code>divdepmodel</code>	Option <code>divdepmodel = 'CS'</code> runs a 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 <code>divdepmodel = 'IW'</code> runs a model with island-wide carrying capacity, where diversity-dependence operates within and among clades. Option <code>divdepmodel = 'GW'</code> runs a model with diversity-dependence operates within a guild.
<code>nonoceanic_pars</code>	<p>A vector of length two with:</p> <ul style="list-style-type: none"> • [1]: the probability of sampling a species from the mainland • [2]: the probability of the species sampled from the mainland being nonendemic
<code>num_guilds</code>	The number of guilds on the mainland. The number of mainland species is divided by the number of guilds when <code>divdepmodel = "GW"</code>
<code>sample_freq</code>	Numeric specifying the number of units times should be divided by for plotting purposes. Larger values will lead to plots with higher resolution, but will also run slower.
<code>plot_sims</code>	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.
<code>island_ontogeny</code>	In <code>DAISIE_sim_time_dependent()</code> , <code>DAISIE_ML_CS</code> and plotting a string describing the type of island ontogeny. Can be "const", "beta" for a beta function

	describing area through time. String checked by <code>is_island_ontogeny_input()</code> . In all other functions a numeric describing the type of island ontogeny. Can be 0 for constant, 1 for a beta function describing area through time. In ML functions <code>island_ontogeny = NA</code> assumes constant ontogeny.
<code>sea_level</code>	In <code>DAISIE_sim_time_dependent()</code> and plotting a string describing the type of sea level. Can be "const" or "sine" for a sine function describing area through time. String checked by <code>is_sea_level_input()</code> . In all other functions a numeric describing the type of sea level. Can be 0 for constant, 1 for a sine function describing area through time.
<code>hyper_pars</code>	A named list of numeric hyperparameters for the rate calculations as returned by <code>create_hyper_pars()</code> : <ul style="list-style-type: none"> • [1]: is d the scaling parameter for exponent for calculating cladogenesis rate • [2]: is x the exponent for calculating extinction rate
<code>area_pars</code>	A named list containing area and sea level parameters as created by <code>create_area_pars()</code> : <ul style="list-style-type: none"> • [1]: maximum area • [2]: current area • [3]: value from 0 to 1 indicating where in the island's history the peak area is achieved • [4]: total island age • [5]: amplitude of area fluctuation from sea level • [6]: frequency of sine wave of area change from sea level • [7]: angle of the slope of the island
<code>extcutoff</code>	A numeric with the cutoff for the the maximum extinction rate preventing it from being too large and slowing down simulation.
<code>verbose</code>	In simulation and dataprep functions a logical, <code>Default = TRUE</code> gives intermediate output should be printed. For ML functions a numeric determining if intermediate output should be printed, <code>Default = 0</code> does not print, <code>verbose = 1</code> prints intermediate output of the parameters and loglikelihood, <code>verbose = 2</code> means also intermediate progress during loglikelihood computation is shown.
<code>trait_pars</code>	A named list containing diversification rates considering two trait states created by <code>create_trait_pars</code> : <ul style="list-style-type: none"> • [1]:A numeric with the per capita transition rate with state1 • [2]:A numeric with the per capita immigration rate with state2 • [3]:A numeric with the per capita extinction rate with state2 • [4]:A numeric with the per capita anagenesis rate with state2 • [5]:A numeric with the per capita cladogenesis rate with state2 • [6]:A numeric with the per capita transition rate with state2 • [7]:A numeric with the number of species with trait state 2 on mainland
<code>...</code>	Any arguments to pass on to plotting functions.

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 last 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 - 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 - only if 2 types of species were simulated (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 simulated (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

* ndemic: 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

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. Hauffe, T., D. Delicado, R.S. Etienne and L. Valente (submitted). Lake expansion increases equilibrium diversity via the target effect of island biogeography.

See Also

[DAISIE_format_CS](#) [DAISIE_plot_sims](#)

DAISIE_SR_loglik_CS *Computes the loglikelihood of the DAISIE model with clade-specific diversity-dependence given data and a set of model parameters that may shift at some time*

Description

Computes the loglikelihood of the DAISIE model with clade-specific diversity-dependence given colonization and branching times for lineages on an island, and a set of model parameters that may shift at some time

Usage

```
DAISIE_SR_loglik_CS(
  pars1,
  pars2,
  datalist,
  methode = "lsodes",
  CS_version = 1,
  abstolint = 1e-16,
  reltolint = 1e-10,
  verbose = FALSE
)
```

Arguments

pars1	<p>Contains the model parameters:</p> <p>pars1[1] corresponds to λ^c (cladogenesis rate) pars1[2] corresponds to μ (extinction rate) pars1[3] corresponds to K (clade-level carrying capacity) pars1[4] corresponds to γ (immigration rate) pars1[5] corresponds to λ^a (anagenesis rate) pars1[6] corresponds to λ^c (cladogenesis rate) after the shift pars1[7] corresponds to μ (extinction rate) after the shift pars1[8] corresponds to K (clade-level carrying capacity) after the shift pars1[9] corresponds to γ (immigration rate) after the shift pars1[10] corresponds to λ^a (anagenesis rate) after the shift pars1[11] corresponds to the time of shift</p>
pars2	<p>Contains the model settings</p> <p>pars2[1] corresponds to lx = length of ODE variable x</p>

	<p><code>pars2[2]</code> corresponds to <code>ddmodel</code> = diversity-dependent model, model of diversity-dependence, which can be one of</p> <p><code>ddmodel</code> = 0 : no diversity dependence <code>ddmodel</code> = 1 : linear dependence in speciation rate <code>ddmodel</code> = 11: linear dependence in speciation rate and in immigration rate <code>ddmodel</code> = 2 : exponential dependence in speciation rate <code>ddmodel</code> = 21: exponential dependence in speciation rate and in immigration rate</p> <p><code>pars2[3]</code> corresponds to <code>cond</code> = setting of conditioning</p> <p><code>cond</code> = 0 : conditioning on island age <code>cond</code> = 1 : conditioning on island age and non-extinction of the island biota</p> <p><code>pars2[4]</code> sets whether parameters and likelihood should be printed (1) or not (0)</p>
<code>datalist</code>	<p>Data object containing information on colonisation and branching times. This object can be generated using the <code>DAISIE_dataprep</code> 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.</p> <p>The first element of the list has two or three components:</p> <p><code>\$island_age</code> - the island age Then, depending on whether a distinction between types is made, we have: <code>\$not_present</code> - the number of mainland lineages 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:</p> <p><code>\$colonist_name</code> - the name of the species or clade that colonized the island <code>\$branching_times</code> - 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. <code>\$stac</code> - the status of the colonist</p> <p>* Non_endemic_MaxAge: 1 * Endemic: 2 * Endemic&Non_Endemic: 3 * Non_endemic: 4 * Endemic_MaxAge: 5</p> <p><code>\$missing_species</code> - number of island species that were not sampled for particular clade (only applicable for endemic clades)</p>
<code>methode</code>	Method of the ODE-solver. See package <code>deSolve</code> for details. Default is "Isodes"
<code>CS_version</code>	For internal testing purposes only. Default is 1, the original DAISIE code.

abstolint	Absolute tolerance of the integration
reltolint	Relative tolerance of the integration
verbose	Logical controlling if progress is printed to console.

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_constant_rate](#)

Examples

```
utils::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_SR_ML

Maximization of the loglikelihood under the DAISIE model with clade-specific diversity-dependence

Description

This function computes the maximum likelihood estimates of the parameters of the DAISIE model with clade-specific diversity-dependence and a shift in parameters for data from lineages colonizing an island. It also outputs the corresponding loglikelihood that can be used in model comparisons.

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.

Usage

```
DAISIE_SR_ML_CS(
  datalist,
  initparsopt,
  idparsopt,
  parsfix,
  idparsfix,
  idparsnoshift = 6:10,
  res = 100,
  ddmodel = 0,
  cond = 0,
  island_ontogeny = NA,
  tol = c(1e-04, 1e-05, 1e-07),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  methode = "lsodes",
  optimmethod = "subplex",
  CS_version = 1,
  verbose = 0,
  tolint = c(1e-16, 1e-10),
  jitter = 0
)
```

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

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

	<ul style="list-style-type: none"> * 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)
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: <ul style="list-style-type: none"> id = 1 corresponds to λ^c (cladogenesis rate) id = 2 corresponds to μ (extinction rate) id = 3 corresponds to K (clade-level carrying capacity) id = 4 corresponds to γ (immigration rate) id = 5 corresponds to λ^a (anagenesis rate) id = 6 corresponds to λ^c (cladogenesis rate) after the shift id = 7 corresponds to μ (extinction rate) after the shift id = 8 corresponds to K (clade-level carrying capacity) after the shift id = 9 corresponds to γ (immigration rate) after the shift id = 10 corresponds to λ^a (anagenesis rate) after the shift id = 11 corresponds to the time of shift
parsfix	The values of the parameters that should not be optimized
idparsfix	The ids of the parameters that should not be optimized, e.g. c(1,3) if λ^c and K should not be optimized.
idparsnoshift	The ids of the parameters that should not be different before and after the shift
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: <ul style="list-style-type: none"> 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	<ul style="list-style-type: none"> cond = 0 : conditioning on island age cond = 1 : conditioning on island age and non-extinction of the island biota
island_ontogeny	type of island ontogeny. If NA, then constant ontogeny is assumed
tol	Sets the tolerances in the optimization. Consists of: <ul style="list-style-type: none"> 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"
optimmethode	Method used in likelihood optimization. Default is "subplex" (see subplex package). Alternative is 'simplex' which was the method in previous versions.
CS_version	For internal testing purposes only. Default is 1, the original DAISIE code.
verbose	sets whether parameters and likelihood should be printed (1) or not (0)
tolint	Vector of two elements containing the absolute and relative tolerance of the integration
jitter	Numeric for <code>optimizer()</code> . Jitters the parameters being optimized by the specified amount which should be very small, e.g. 1e-5. Jitter when <code>link[subplex]{subplex}()</code> produces incorrect output due to parameter transformation.

Value

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

lambda_c	gives the maximum likelihood estimate of λ^c , the rate of cladogenesis
mu	gives the maximum likelihood estimate of μ , the extinction rate
K	gives the maximum likelihood estimate of K, the carrying-capacity
gamma	gives the maximum likelihood estimate of γ , the immigration rate
lambda_a	gives the maximum likelihood estimate of λ^a , the rate of anagenesis
lambda_c2	gives the maximum likelihood estimate of λ^c_2 , the rate of cladogenesis for the optional second group of species
mu2	gives the maximum likelihood estimate of μ_2 , 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 γ_2 , the immigration rate for the optional second group of species
lambda_a2	gives the maximum likelihood estimate of λ^a_2 , 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 freedom
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_constant_rate](#)

Examples

```
## Not run:
### When all species have the same rates, and we want to optimize all 5 parameters,
# we use:
```

```
utils::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:
```

```
utils::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:
```

```
utils::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:
```

```
utils::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:

utils::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:

utils::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:

utils::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)
  )

## End(Not run)

```

frogs_datalist	<i>Colonization and branching times of 5 Eleutherodactylus clades from Hispaniola island.</i>
----------------	---

Description

A list containing the colonization and branching times of the Eleutherodactylus frogs of Hispaniola. Main dataset used in Etienne et al. 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 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

\$stac - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_endemic: 4

* Endemic_MaxAge: 5 or 6

\$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.

Format

A list with 6 elements, the first of which contains 2 elements and the following 5 elements containing 5 components.

Source

Etienne RS, Haegeman B, Dugo-Cota A, Vila C, Gonzalez-Voyer A & Valente L. The limits to ecological limits to diversification.

See Also

[DAISIE_dataprep](#), [DAISIE_ML](#), [DAISIE_SR_ML](#)

frogs_datatable	<i>Colonization and branching times of 5 Eleutherodactylus (frogs) clades from the island of Hispaniola.</i>
-----------------	--

Description

A table containing the colonization and branching times of the Eleutherodactylus frogs of the island of Hispaniola (Greater Antilles). Each row on the table represents an independent colonisation event. The table has four columns.

\$Clade_name - name of independent colonization event

\$Status - One of the following categories:

* Non_endemic: for non-endemic island species when an approximate time of colonisation is known

* Non_endemic_MaxAge: for non-endemic island species when colonisation time is unknown

* Endemic: for endemic species when an approximate colonisation time is known

* "Endemic_MaxAge": applies to endemic species or endemic clades for cases where the colonisation time is unknown, or when the user wants to specify an upper bound for colonisation. This could for example apply to endemic species that have recently gone extinct because of anthropogenic causes, and which are not included in the phylogeny ("NA" should be given in the branching times column). It could also apply to insular radiations with long stem branches, for which the time of the first cladogenetic event is known, but the precise time of colonisation is not.

* Endemic&Non_Endemic: when endemic clade and mainland ancestor has re-colonized

Format

A table with 5 rows and 4 columns.

Source

Etienne RS, Haegeman B, Dugo-Cota A, Vila C, Gonzalez-Voyer A & Valente L. The limits to ecological limits to diversification.

Galapagos_datalist	<i>Colonization and branching times of 8 terrestrial avifaunal clades in list format, accepted by DAISIE_ML and DAISIE_loglik_all</i>
--------------------	---

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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

`$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.

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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

`$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

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	<i>Colonization and branching times of 8 terrestrial avifaunal Galápagos clades in table format.</i>
---------------------	--

Description

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

\$Clade_name - name of independent colonization event

\$Status - One of the following categories:

* Non_endemic: for non-endemic island species when an approximate time of colonisation is known

* Non_endemic_MaxAge: for non-endemic island species when colonisation time is unknown

* Endemic: for endemic species when an approximate colonisation time is known

* "Endemic_MaxAge": applies to endemic species or endemic clades for cases where the colonisation time is unknown, or when the user wants to specify an upper bound for colonisation. This could for example apply to endemic species that have recently gone extinct because of anthropogenic causes, and which are not included in the phylogeny ("NA" should be given in the branching times column). It could also apply to insular radiations with long stem branches, for which the time of the first cladogenetic event is known, but the precise time of colonisation is not.

* 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" species with no extant close relatives on the island. Set "NA" if colonisation time unknown and no upper bound is known. For "Endemic" cladogenetic species these should be branching times of the radiation, including the stem age of the radiation (colonisation time estimate).

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	<i>1000 islands in RAW format simulated with the ML parameters of the CR model for the Galapagos data.</i>
--------------------	--

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

* 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)

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_constant_rate\(\)](#), [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)

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_constant_rate\(\)](#), [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

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_constant_rate\(\)](#), [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 elements:

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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

* 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. In this dataset all are equal to 1.

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. Equilibrium bird species diversity in Atlantic islands. 2017 Current Biology, 27, 1660-1666.

See Also

[DAISIE_dataprep](#), [DAISIE_ML](#)

NewZealand_birds_datalist

Colonization and branching times of New Zealand birds.

Description

A list containing the colonization and branching times of the birds of New Zealand. Main dataset used in Valente, Etienne, Garcia-R (2019) Current Biology. Island age 52 Myr and mainland pool size of 1000 species.

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 elements of the list each contain 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 followed by stem age of the population/species in the case of Non-endemic, Non-endemic_MaxAge species and Endemic species with no close relatives on the island. For endemic clades with more than one species on the island (cladogenetic clades/ radiations) these should be island age followed by the branching times of the island clade including the stem age of the clade.

`$stac` - the status of the colonist

* Non_endemic_MaxAge: 1

* Endemic: 2

* Endemic&Non_Endemic: 3

* Non_endemic: 4

* Endemic_MaxAge: 5 or 6

`$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.

Format

A list with 40 elements, the first of which contains 2 elements and the following 39 containing 5 components.

Source

Valente L, Etienne RS, Garcia-R JC (2019) Deep Macroevolutionary Impact of Humans on New Zealand's Unique Avifauna. *Current Biology*, 29, 2563–2569.

See Also

[DAISIE_dataprep](#), [DAISIE_ML](#), [DAISIE_SR_ML](#)

Index

* datasets

archipelago_data, 3
archipelagos41, 3
Bats_GreaterAntilles, 4
Biwa_datalist, 5
frogs_datalist, 81
frogs_datatable, 82
Galapagos_datalist, 83
Galapagos_datalist_2types, 84
Galapagos_datatable, 85
islands_10reps_RAW, 86
islands_1type_1000reps, 87
islands_2types_1000reps, 88
Macaronesia_datalist, 89
NewZealand_birds_datalist, 90

* models

DAISIE_convertprobdist, 10
DAISIE_dataprep, 12
DAISIE_ExpEIN, 15
DAISIE_margprobdist, 23
DAISIE_ML, 25
DAISIE_ML_IW, 32
DAISIE_MW_ML, 35
DAISIE_numcol, 40
DAISIE_plot_age_diversity, 41
DAISIE_plot_island, 44
DAISIE_plot_sims, 46
DAISIE_probdist, 48
DAISIE_sim, 49
DAISIE_sim_constant_rate_shift, 54
DAISIE_sim_MW, 58
DAISIE_sim_relaxed_rate, 61
DAISIE_sim_time_dependent, 65
DAISIE_sim_trait_dependent, 69
DAISIE_SR_loglik_CS, 73
DAISIE_SR_ML, 75

* simulation models

DAISIE_sim, 49
DAISIE_sim_constant_rate_shift, 54

DAISIE_sim_relaxed_rate, 61
DAISIE_sim_time_dependent, 65

archipelago_data, 3
archipelagos41, 3

Bats_GreaterAntilles, 4
Biwa_datalist, 5

create_area_pars, 6, 51, 56, 63, 66, 71
create_CS_version, 7
create_hyper_pars, 8, 51, 56, 63, 66, 71
create_trait_pars, 9, 11, 18, 47, 71

DAISIE_convertprobdist, 10
DAISIE_create_island, 11
DAISIE_dataprep, 5, 6, 12, 51, 82, 84–86, 90, 91

DAISIE_ExpEIN, 15
DAISIE_format_CS, 47, 73
DAISIE_format_CS_full_stt, 17
DAISIE_get_brts_my, 18
DAISIE_IC, 19
DAISIE_loglik_all, 30, 79
DAISIE_loglik_CS, 23
DAISIE_loglik_IW, 21, 35
DAISIE_margprobdist, 16, 23, 24, 40, 48
DAISIE_ML, 5, 6, 25, 75, 82, 84–86, 90, 91
DAISIE_ML_CS, 28, 35, 39, 67, 70
DAISIE_ML_CS (DAISIE_ML), 25
DAISIE_ML_IW, 23, 32
DAISIE_MW_ML, 35
DAISIE_numcol, 16, 24, 40, 40, 48
DAISIE_plot_age_diversity, 41, 41, 45
DAISIE_plot_input, 42
DAISIE_plot_island, 41, 42, 44, 45
DAISIE_plot_sims, 46, 53, 58, 64, 68, 73, 86, 87, 89
DAISIE_probdist, 10, 16, 24, 40, 48, 48
DAISIE_sim, 49, 58, 64, 68

DAISIE_sim_constant_rate, [23](#), [30](#), [35](#), [42](#),
[45](#), [47](#), [61](#), [75](#), [79](#), [86](#), [87](#), [89](#)
DAISIE_sim_constant_rate (DAISIE_sim),
[49](#)
DAISIE_sim_constant_rate_shift, [30](#), [42](#),
[45](#), [47](#), [53](#), [54](#), [64](#), [68](#)
DAISIE_sim_core_constant_rate, [17](#), [46](#)
DAISIE_sim_core_constant_rate_shift,
[17](#), [46](#)
DAISIE_sim_core_time_dependent, [17](#), [46](#)
DAISIE_sim_min_type2, [17](#), [46](#)
DAISIE_sim_MW, [58](#)
DAISIE_sim_relaxed_rate, [16](#), [50](#), [53](#), [56](#),
[58](#), [61](#), [62](#), [66](#), [68](#), [70](#)
DAISIE_sim_time_dependent, [28](#), [30](#), [42](#), [45](#),
[47](#), [53](#), [58](#), [64](#), [65](#), [67](#), [70](#), [71](#)
DAISIE_sim_trait_dependent, [69](#)
DAISIE_SR_loglik_all
(DAISIE_SR_loglik_CS), [73](#)
DAISIE_SR_loglik_CS, [73](#)
DAISIE_SR_ML, [5](#), [6](#), [75](#), [82](#), [91](#)
DAISIE_SR_ML_CS (DAISIE_SR_ML), [75](#)

frogs_datalist, [81](#)
frogs_datatable, [82](#)

Galapagos_datalist, [83](#)
Galapagos_datalist_2types, [84](#)
Galapagos_datatable, [85](#)

is_island_ontogeny_input, [28](#), [67](#), [71](#)
is_sea_level_input, [67](#), [71](#)
islands_10reps_RAW, [86](#)
islands_1type_1000reps, [87](#)
islands_2types_1000reps, [88](#)

Macaronesia_datalist, [89](#)

NewZealand_birds_datalist, [90](#)

optimizer, [29](#), [34](#), [78](#)