

# Package ‘DDD’

July 21, 2016

**Type** Package

**Title** Diversity-Dependent Diversification

**Version** 3.3

**Date** 2016-7-13

**Depends** R (>= 3.0.2), deSolve, ape, ade4

**Imports** phytools, subplex, Matrix, expoRkit, SparseM

**Author** Rampal S. Etienne & Bart Haegeman

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

**License** GPL-3

**Description** Implements maximum likelihood methods based on the diversity-dependent birth-death process to test whether speciation or extinction are diversity-dependent.  
See Etienne et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, <DOI:10.1098/rspb.2011.1439>.  
Also implements maximum likelihood methods to detect various types of key innovations in the light of diversity-dependence.  
See Etienne & Haegeman 2012, Am. Nat. 180: E75-E89, <DOI:10.1086/667574>.  
Finally, DDD contains a function to simulate the diversity-dependent process.

**NeedsCompilation** no

## R topics documented:

|                        |    |
|------------------------|----|
| DDD-package . . . . .  | 2  |
| bd_loglik . . . . .    | 6  |
| bd_ML . . . . .        | 7  |
| dd_KI_loglik . . . . . | 10 |
| dd_KI_ML . . . . .     | 12 |
| dd_KI_sim . . . . .    | 15 |
| dd_loglik . . . . .    | 16 |
| dd_LR . . . . .        | 18 |
| dd_ML . . . . .        | 21 |
| dd_MS_loglik . . . . . | 24 |
| dd_MS_ML . . . . .     | 26 |
| dd_MS_sim . . . . .    | 29 |
| dd_sim . . . . .       | 30 |
| dd_SR_loglik . . . . . | 32 |
| dd_SR_ML . . . . .     | 33 |
| dd_SR_sim . . . . .    | 36 |

|                     |    |
|---------------------|----|
| L2phylo . . . . .   | 38 |
| optimizer . . . . . | 39 |
| roundn . . . . .    | 39 |
| sample2 . . . . .   | 40 |

|              |           |
|--------------|-----------|
| <b>Index</b> | <b>42</b> |
|--------------|-----------|

---

|             |  |
|-------------|--|
| DDD-package | <i>Diversity Dependent Diversification</i> |
|-------------|--|

---

## Description

This package is a likelihood-based statistical package to estimate parameters under diversity-dependent diversification models.

New in version 0.3:

- Models with shifts in the parameters at a time tshift.

New in version 0.4:

- Bug fix of prefactor in likelihood - bug did not affect parameter estimates.
- There is an option to condition or not on non-extinction of the phylogeny.
- The functions dd\_loglik and dd\_SR\_loglik now accept untransformed parameters; the transformation is done in the auxiliary functions.
- Improved help pages.
- Auxiliary functions no longer in help pages.

New in version 1.0:

- Another bug fix of prefactor in likelihood - bug did not affect parameter estimates.
- Models with key innovations defined as decoupling of diversity-dependent dynamics.
- Option to compute the likelihood of a set of branching times or of the phylogeny; these differ by a constant factor.

New in version 1.1:

- A model with diversity-dependence in speciation and extinction rate, currently only available for the likelihood without shifts or decoupling.

New in version 1.2:

- Possibility of automatically exploring multiple initial conditons for the SR models
- Suppressing unnecessary output
- Flushing computed likelihood values immediately (even with buffer on)
- Bug fix in default value of the initial time of decoupling

New in version 1.3:

- Fix of default length of idparsopt
- Fix of conversion problems for small branching times
- Default initial parameter values in dd\_ML based on birth-death model

New in version 1.4:

- Bug fix for when there is a large number of missing species

New in version 1.5:

- Bug fix for returned output of functions when there are fixed parameters.

Textual output of DDD was correct.

- Return to Nelder-Mead optimization for dd\_ML (was set to "nlm" in versions 1.3 & 1.4)
- New format for output of all ML-functions (dd\_ML, dd\_SR\_ML, dd\_KI\_ML)

New in version 1.6 (not on CRAN)

- Bug fix in returned result when optimization does not converge.
- Works also when only branching times are entered in ascending order (previously error was reported)

New in version 1.7

- Cleaning up code, fixing various small bugs
- New optimization algorithm because of occasional strange results with optim package

New in version 1.8

- Bug fix for missing number of species in KI model
- Bug fix for optimization of one parameter (due to R's inconsistent handling of matrices)
- Model 5 (dd\_ML and dd\_loglik) now also allows values of  $-1 < r < 0$ ; i.e. when extinction also decreases with diversity but less so than speciation

New in version 1.9

- Bug fix in optimizer (did not stop when likelihood was positive)
- Bug fix for large numbers of missing species
- Conv = -1 is now always reported when optimization is not started or does not converge
- An error is reported when parameter values cause numerical problems (NA or NaN) in likelihood computation, e.g. when parameter values are very high; loglik is then set to -Inf
- For dd\_KI\_loglik and dd\_KI\_ML: if the number of missing species in main clade and subclade is known, this can now be specified
- Fixing K at Inf is now possible in SR and KI models

New in version 1.10

- All non-convergences now get conv = -1 (this was conv = 1 for true nonconvergence and conv = -1 for optimization that did not start because of low likelihood values)
- Likelihoods can also be computed for a tree with a stem age rather than only a crown age

New in version 1.11

- Various small tweaks to make the code run faster and some bug fixes

New in version 1.12

- Allows conditioning on survival of crown lineages AND on the number of extant species by setting cond = 2

New in version 1.13

- Fast computation of dd\_loglik for the case K = Inf

New in version 1.14

- Fast computation of dd\_loglik for the case  $\mu = 0$
- Allows time-dependence instead of diversity-dependence; exponential decline in speciation rate and/or extinction
- Start of time-dependence can be specified to allow synchronisation across multiple clades
- Removed link to LASER

New in version 2.0

- Includes function to simulate trees under diversity-dependence

## New in version 2.1

- Fixes numerical problems with the convolve-function (used to deal with missing species)
- Fixes a bug in dd\_KI\_loglik (ddep unknown)
- Changed third element of the output of dd\_sim, such that all speciation and extinction times are in Mya

## New in version 2.2

- Allows conditioning on the number of taxa only for dd\_loglik, assuming a uniform prior on stem age; this is now the default
- An additional model of time-dependence, following deterministic logistic equation

## New in version 2.3

- Allows conditioning on the number of taxa only for dd\_SR\_loglik, assuming a uniform prior on stem age; this is now the default
- Fix of bug for stem age (soc = 1), introduced in v2.2.
- Allows Rabosky's exponential model, with  $la(n) = la_0 * n^{(-x(la_0, mu, K))}$  as ddmodel = 2.1,  $mu(n) = mu_0 * n^{x(la, mu_0, K)}$  as ddmodel = 4.1
- Allows specific case of Rabosky's exponential model, with  $la(n) = la_0 * n^{(-1)}$ , as ddmodel = 2.2, and  $mu(n) = mu_0 * n$ , as ddmodel = 4.2

## New in version 2.4

- Fix of bug in dd\_KI\_loglik ('m' was not found)

## New in version 2.5

- Fix of bug in bd\_loglik where NA was output to screen for a nonexistent parameter under diversity-dependence without extinction
- Fix of bug in dd\_KI\_loglik when m is a two-element vector
- Fix of bug in dd\_KI\_loglik, introduced in v2.3

## New in version 2.6

- Allows linear dependence of speciation with parameter  $K'$  (= diversity where speciation = 0)

## New in version 2.7

- Fixes bugs in dd\_KI\_loglik, for cond = 0 introduced in v2.3, and for cond = 1 introduced in v2.0
- Solves (some) memory problems with dd\_KI\_ML and dd\_KI\_loglik when cond = 1
- Provides option to set loglikelihood to -Inf if optimization did not converge

## New in version 3.0

- Includes a model of macroevolutionary succession in which there is no decoupling of diversity-dependence but parameters may change
- Computes LRb, a bootstrap version of the likelihood ratio
- Allows use of alternative ODE solvers
- Allows use of subplex as optimization algorithm.

## New in version 3.1

- Fixes bug with dd\_sim and dd\_KI\_sim for ddmodel 1.3 (Thanks to Giovanni Laudanno)
- Fixes bug with cond = 3
- Includes matrix exponentiation as an alternative to numerical ODE solvers

## New in version 3.2

- Fixes bug with dd\_loglik when lambda becomes infinite

- Changes default tolerance of subplex
- Fixes tiny bug in p-value computation in dd\_LR

New in version 3.3

- Documentation for several previously internal functions that were used by other packages
- Contains a function to simulate the SR model

Forthcoming:

- High-precision arithmetic in numerically unstable regions of bd\_loglik

## Details

Package: DDD  
 Type: Package  
 Version: 3.3  
 Date: 2016-7-13  
 License: GPL 3.0

bd\_loglik computes the likelihood under the constant-rate birth-death model of several specific time-dependent birth-death models.

bd\_ML finds the parameters that maximizes the likelihood computed by bd\_loglik.

dd\_loglik computes the likelihood of any of 4 different models of diversity-dependent diversification, given a set of parameters and a data set of phylogenetic branching times.

dd\_ML finds the parameters that maximizes the likelihood computed by dd\_loglik.

dd\_SR\_loglik is the same as dd\_loglik except that it allows a shift in the parameters at time tshift.

dd\_SR\_ML finds the parameters that maximizes the likelihood computed by dd\_SR\_loglik.

dd\_KI\_loglik is the same as dd\_loglik except that it allows decoupling of a subclade's dynamics from the main clade's dynamics potentially accompanied by a shift in parameters at time t\_d.

dd\_KI\_ML finds the parameters that maximizes the likelihood computed by dd\_KI\_loglik.

dd\_MS\_loglik is the same as dd\_KI\_loglik except that the subclade's dynamics do not decouple from the main clade's dynamics but they do get different parameters at time t\_d.

dd\_MS\_ML finds the parameters that maximizes the likelihood computed by dd\_MS\_loglik.

dd\_sim, dd\_SR\_sim, dd\_KI\_sim and dd\_MS\_sim simulate the diversity-dependent process according to the simple, rate shift, key innovation and macro-evolutionary models respectively.

dd\_LR performs a bootstrap likelihood ratio test of the standard diversity-dependent model against the constant-rates diversity-independent model.

## Author(s)

Rampal S. Etienne & Bart Haegeman  
 Maintainer: Rampal S. Etienne (r.s.etienne@rug.nl)

## References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**See Also**

TESS

---

|           |  |
|-----------|--|
| bd_loglik | <i>Loglikelihood for diversity-independent diversification model</i> |
|-----------|--|

---

**Description**

This function computes loglikelihood of a diversity-independent diversification model for a given set of branching times and parameter values.

**Usage**

```
bd_loglik(pars1, pars2, brts, missnumspec, methode = 'lsoda')
```

**Arguments**

**pars1**                Vector of parameters:

pars1[1] corresponds to lambda0 (speciation rate)  
 pars1[2] corresponds to mu0 (extinction rate)  
 pars1[3] corresponds to lambda1 (decline parameter in speciation rate) or K in diversity-dependence-like models  
 pars1[4] corresponds to mu1 (decline parameter in extinction rate)

**pars2**                Vector of model settings:

pars2[1] sets the model of time-dependence:  
 - pars2[1] == 0 no time dependence  
 - pars2[1] == 1 speciation and/or extinction rate is exponentially declining with time  
 - pars2[1] == 2 stepwise decline in speciation rate as in diversity-dependence without extinction  
 - pars2[1] == 3 decline in speciation rate following deterministic logistic equation for ddmodel = 1  
 - pars2[1] == 4 decline in speciation rate such that the expected number of species matches with that of ddmodel = 1 with the same mu

pars2[2] sets the conditioning:  
 - pars2[2] == 0 conditioning on stem or crown age  
 - pars2[2] == 1 conditioning on stem or crown age and non-extinction of the phylogeny  
 - pars2[2] == 2 conditioning on stem or crown age and on the total number of extant taxa (including missing species)  
 - pars2[2] == 3 conditioning on the total number of extant taxa (including missing species)

pars2[3] sets whether the likelihood is for the branching times (0) or the phylogeny (1)

pars2[4] sets whether the parameters and likelihood should be shown on screen

|            |   |
|------------|---|
|            | (1) or not (0)  |
|            | pars2[5] sets whether the first data point is stem age (1) or crown age (2) |
| brts       | A set of branching times of a phylogeny, all positive                       |
| misnumspec | The number of species that are in the clade but missing in the phylogeny    |
| methode    | The method used to solve the master equation, default is 'lsoda'.           |

**Value**

The loglikelihood

**Author(s)**

Rampal S. Etienne, Bart Haegeman & Cesar Martinez

**References**

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**See Also**

[bd\\_ML](#)

**Examples**

```
bd_loglik(pars1 = c(0.5,0.1), pars2 = c(0,1,1,0,2), brts = 1:10, misnumspec = 0)
```

---

|       |   |
|-------|---|
| bd_ML | <i>Maximization of the loglikelihood under the diversity-independent, possibly time-dependent diversification model</i> |
|-------|---|

---

**Description**

This function computes the maximum likelihood estimates of the parameters of a diversity-independent diversification model for a given set of phylogenetic branching times. It also outputs the corresponding loglikelihood that can be used in model comparisons.

**Usage**

```
bd_ML(
  brts,
  initparsopt = c(0.1,0.05 * (tdmodel <= 1)
+ 10 * (length(brts) + misnumspec) * (tdmodel > 1)),
  idparsopt = c(1,2 + (tdmodel > 1)),
  idparsfix = (1:4)[-idparsopt],
  parsfix = rep(0,4)[idparsfix],
  misnumspec = 0,
  tdmodel = 0,
  cond = 1,
  btorph = 1,
```

```

soc = 2,
tol = c(1E-3, 1E-4, 1E-6),
maxiter = 1000 * round((1.25)^length(idparsopt)),
changeloglikifnoconv = FALSE,
optimmethod = 'subplex',
methode = 'lsoda'
)

```

## Arguments

|                          |  |
|--------------------------|--|
| <code>brts</code>        | A set of branching times of a phylogeny, all positive  |
| <code>initparsopt</code> | The initial values of the parameters that must be optimized  |
| <code>idparsopt</code>   | The ids of the parameters that must be optimized, e.g. 1:3 for intrinsic speciation rate, extinction rate and carrying capacity. The ids are defined as follows:<br>id == 1 corresponds to $\lambda_0$ (speciation rate)<br>id == 2 corresponds to $\mu_0$ (extinction rate)<br>id == 3 corresponds to $\lambda_1$ (parameter controlling decline in speciation rate with time)<br>id == 4 corresponds to $\mu_1$ (parameter controlling decline in extinction rate with time)   |
| <code>idparsfix</code>   | The ids of the parameters that should not be optimized, e.g. c(1,3) if $\lambda_0$ and $\lambda_1$ should not be optimized, but only $\mu_0$ and $\mu_1$ . In that case <code>idparsopt</code> must be c(2,4). The default is to fix all parameters not specified in <code>idparsopt</code> .  |
| <code>parsfix</code>     | The values of the parameters that should not be optimized  |
| <code>missnumspec</code> | The number of species that are in the clade but missing in the phylogeny   |
| <code>tdmodel</code>     | Sets the model of time-dependence:<br>tdmodel == 0 : constant speciation and extinction rates<br>tdmodel == 1 : speciation and/or extinction rate is exponentially declining with time<br>tdmodel == 2 : stepwise decline in speciation rate as in diversity-dependence without extinction<br>tdmodel == 3 : decline in speciation rate following deterministic logistic equation for <code>ddmodel</code> = 1<br>tdmodel == 4 : decline in speciation rate such that the expected number of species matches with that of <code>ddmodel</code> = 1 with the same $\mu$ |
| <code>cond</code>        | Conditioning:<br>cond == 0 : conditioning on stem or crown age<br>cond == 1 : conditioning on stem or crown age and non-extinction of the phylogeny<br>cond == 2 : conditioning on stem or crown age and on the total number of extant taxa (including missing species)<br>cond == 3 : conditioning on the total number of extant taxa (including missing species)   |
| <code>btorph</code>      | Sets whether the likelihood is for the branching times (0) or the phylogeny (1)  |
| <code>soc</code>         | Sets whether stem or crown age should be used (1 or 2)   |
| <code>tol</code>         | Sets the tolerances in the optimization. Consists of:<br>reltolx = relative tolerance of parameter values in optimization<br>reltolf = relative tolerance of function value in optimization<br>abstolx = absolute tolerance of parameter values in optimization  |



|                      |  |
|----------------------|--|
| maxiter              | Sets the maximum number of iterations in the optimization  |
| changeloglikifnoconv | if TRUE the loglik will be set to -Inf if ML does not converge   |
| optimmethod          | Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions) |
| methode              | The method used to solve the master equation under tmodel = 4, default is 'lsoda'.   |

## Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor  $q! m! / (q + m)!$  where  $q$  is the number of species in the phylogeny and  $m$  is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

## Value

|         |  |
|---------|--|
| lambda0 | gives the maximum likelihood estimate of lambda0                           |
| mu0     | gives the maximum likelihood estimate of mu0                               |
| lambda1 | gives the maximum likelihood estimate of lambda1                           |
| mu1     | gives the maximum likelihood estimate of mu1                               |
| 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 & Bart Haegeman

## References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

## See Also

[bd\\_loglik](#)

## Examples

```
cat("Estimating parameters for a set of branching times brts with the default settings:")
brts = 1:20
bd_ML(brts = brts, cond = 1)
```

---

|              |  |
|--------------|--|
| dd_KI_loglik | <i>Loglikelihood for diversity-dependent diversification models with decoupling of a subclade from a main clade at time <math>t = t_d</math></i> |
|--------------|--|

---

## Description

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the diversity-dependent dynamics of a subclade decouple from the dynamics of the main clade at time  $t_d$ , potentially accompanied by a shift in parameters.

## Usage

```
dd_KI_loglik(pars1, pars2, brtsM, brtsS, missnumspec, methode = 'analytical')
```

## Arguments

|       |  |
|-------|--|
| pars1 | <p>Vector of parameters:</p> <p>pars1[1] corresponds to <math>\lambda_M</math> (speciation rate) of the main clade<br/> pars1[2] corresponds to <math>\mu_M</math> (extinction rate) of the main clade<br/> pars1[3] corresponds to <math>K_M</math> (clade-level carrying capacity) of the main clade<br/> pars1[4] corresponds to <math>\lambda_S</math> (speciation rate) of the subclade<br/> pars1[5] corresponds to <math>\mu_S</math> (extinction rate) of the subclade<br/> pars1[6] corresponds to <math>K_S</math> (clade-level carrying capacity) of the subclade<br/> pars1[7] corresponds to <math>t_d</math> (the time of decoupling)</p>  |
| pars2 | <p>Vector of model settings:</p> <p>pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than <math>1 + \text{missnumspec} + \text{length}(\text{brts})</math>.</p> <p>pars2[2] sets the model of diversity-dependence:</p> <ul style="list-style-type: none"> <li>- pars2[2] == 1 linear dependence in speciation rate with parameter <math>K</math> (= diversity where speciation = extinction)</li> <li>- pars2[2] == 1.3 linear dependence in speciation rate with parameter <math>K'</math> (= diversity where speciation = 0)</li> <li>- pars2[2] == 2 exponential dependence in speciation rate with parameter <math>K</math> (= diversity where speciation = extinction)</li> <li>- pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity</li> <li>- pars2[2] == 2.2 1/n dependence in speciation rate</li> <li>- pars2[2] == 2.3 exponential dependence in speciation rate with parameter <math>x</math> (= exponent)</li> <li>- pars2[2] == 3 linear dependence in extinction rate</li> <li>- pars2[2] == 4 exponential dependence in extinction rate</li> <li>- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity</li> <li>- pars2[2] == 4.2 1/n dependence in extinction rate</li> </ul> <p>pars2[3] sets the conditioning:</p> <ul style="list-style-type: none"> <li>- pars2[3] == 0 no conditioning</li> </ul> |

|             |  |
|-------------|--|
|             | - pars2[3] == 1 conditioning on non-extinction of the phylogeny  |
|             | pars2[4] sets the time of splitting of the branch that will decouple   |
|             | pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)   |
|             | pars2[6] sets whether the first data point is stem age (1) or crown age (2)  |
| brtsM       | A set of branching times of the main clade in the phylogeny, all positive  |
| brtsS       | A set of branching times of the subclade in the phylogeny, all positive  |
| missnumspec | The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector c(missnumspec_M,missnumspec_S) with missing species in main clade and subclade respectively. |
| methode     | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.                      |

**Value**

The loglikelihood

**Author(s)**

Rampal S. Etienne & Bart Haegeman

**References**

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**See Also**

[dd\\_KI\\_ML](#), [dd\\_loglik](#) [dd\\_SR\\_loglik](#)

**Examples**

```
pars1 = c(0.25,0.12,25.51,1.0,0.16,8.61,9.8)
pars2 = c(200,1,0,18.8,1,2)
missnumspec = 0
brtsM = c(25.2,24.6,24.0,22.5,21.7,20.4,19.9,19.7,18.8,17.1,15.8,11.8,9.7,8.9,5.7,5.2)
brtsS = c(9.6,8.6,7.4,4.9,2.5)
dd_KI_loglik(pars1,pars2,brtsM,brtsS,missnumspec,method = 'ode45')
```

---

|          |  |
|----------|--|
| dd_KI_ML | <i>Maximization of the loglikelihood under a diversity-dependent diversification model with decoupling of a subclade's diversification dynamics from the main clade's dynamics</i> |
|----------|--|

---

## Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model with decoupling of the diversification dynamics of a subclade from the dynamics of the main clade for a given set of phylogenetic branching times of main clade and subclade and the time of splitting of the lineage that will form the subclade. It also outputs the corresponding loglikelihood that can be used in model comparisons.

## Usage

```
dd_KI_ML(
  brtsM,
  brtsS,
  tsplit,
  initparsopt = c(0.5, 0.1, 2 * (1 + length(brtsM) + missnumspec[1]),
  2 * (1 + length(brtsS) + missnumspec[length(missnumspec)]), (tsplit + max(brtsS))/2),
  parsfix = NULL,
  idparsopt = c(1:3, 6:7),
  idparsfix = NULL,
  idparsnoshift = (1:7)[c(-idparsopt, (-1)^(length(idparsfix) != 0) * idparsfix)],
  res = 10 * (1 + length(c(brtsM, brtsS)) + sum(missnumspec)),
  ddmodel = 1,
  missnumspec = 0,
  cond = 1,
  soc = 2,
  tol = c(1E-3, 1E-4, 1E-6),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  changeloglikifnoconv = FALSE,
  optimmethod = 'subplex',
  methode = 'analytical'
)
```

## Arguments

|             |   |
|-------------|---|
| brtsM       | A set of branching times of the main clade in a phylogeny, all positive   |
| brtsS       | A set of branching times of the subclade in a phylogeny, all positive   |
| tsplit      | The branching time at which the lineage forming the subclade branches off, positive   |
| initparsopt | The initial values of the parameters that must be optimized   |
| parsfix     | The values of the parameters that should not be optimized   |
| idparsopt   | The ids of the parameters that must be optimized, e.g. 1:7 for all parameters.<br>The ids are defined as follows:<br>id == 1 corresponds to lambda_M (speciation rate) of the main clade<br>id == 2 corresponds to mu_M (extinction rate) of the main clade |

|                      |  |
|----------------------|--|
|                      | <p>id == 3 corresponds to K_M (clade-level carrying capacity) of the main clade</p> <p>id == 4 corresponds to lambda_S (speciation rate) of the subclade</p> <p>id == 5 corresponds to mu_S (extinction rate) of the subclade</p> <p>id == 6 corresponds to K_S (clade-level carrying capacity) of the subclade</p> <p>id == 7 corresponds to t_d (the time of decoupling)</p>   |
| idparsfix            | The ids of the parameters that should not be optimized, e.g. c(1,3,4,6) if lambda and K should not be optimized, but only mu. In that case idparsopt must be c(2,5,7). The default is to fix all parameters not specified in idparsopt.  |
| idparsnoshift        | The ids of the parameters that should not shift; This can only apply to ids 4, 5 and 6, e.g. idparsnoshift = c(4,5) means that lambda and mu have the same values before and after tshift  |
| res                  | sets the maximum number of species for which a probability must be computed, must be larger than 1 + max(length(brtsM),length(brtsS))  |
| ddmodel              | <p>sets the model of diversity-dependence:</p> <p>ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)</p> <p>ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)</p> <p>ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)</p> <p>ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity</p> <p>ddmodel == 2.2 : 1/n dependence in speciation rate</p> <p>ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)</p> <p>ddmodel == 3 : linear dependence in extinction rate</p> <p>ddmodel == 4 : exponential dependence in extinction rate</p> <p>ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity</p> <p>ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity</p> |
| misnumspec           | The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector c(misnumspec_M,misnumspec_S) with missing species in main clade and subclade respectively.   |
| cond                 | <p>Conditioning:</p> <p>cond == 0 : no conditioning</p> <p>cond == 1 : conditioning on non-extinction of the phylogeny</p>   |
| soc                  | Sets whether stem or crown age should be used (1 or 2); stem age only works when cond = 0  |
| tol                  | <p>Sets the tolerances in the optimization. Consists of:</p> <p>reltolx = relative tolerance of parameter values in optimization</p> <p>reltolf = relative tolerance of function value in optimization</p> <p>abstolx = absolute tolerance of parameter values in optimization</p>   |
| maxiter              | Sets the maximum number of iterations in the optimization  |
| changeloglikifnoconv | if TRUE the loglik will be set to -Inf if ML does not converge   |
| optimmethod          | Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)   |

**methode** The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.

### Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor  $q! \cdot m! / (q + m)!$  where  $q$  is the number of species in the phylogeny and  $m$  is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

### Value

|                 |  |
|-----------------|--|
| <b>lambda_M</b> | gives the maximum likelihood estimate of lambda of the main clade          |
| <b>mu_M</b>     | gives the maximum likelihood estimate of mu of the main clade              |
| <b>K_M</b>      | gives the maximum likelihood estimate of K of the main clade               |
| <b>lambda_2</b> | gives the maximum likelihood estimate of lambda of the subclade            |
| <b>mu_S</b>     | gives the maximum likelihood estimate of mu of the subclade                |
| <b>K_S</b>      | gives the maximum likelihood estimate of K of the subclade                 |
| <b>t_d</b>      | gives the time of the decoupling event                                     |
| <b>loglik</b>   | gives the maximum loglikelihood  |
| <b>df</b>       | gives the number of estimated parameters, i.e. degrees of freedom          |
| <b>conv</b>     | gives a message on convergence of optimization; conv = 0 means convergence |

### Note

The optimization may get trapped in local optima. Try different starting values to search for the global optimum.

### Author(s)

Rampal S. Etienne & Bart Haegeman

### References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

### See Also

[dd\\_KI\\_loglik](#), [dd\\_ML](#), [dd\\_SR\\_ML](#),

### Examples

```
cat("This will estimate parameters for two sets of branching times brtsM, brtsS\n")
cat("without conditioning.\n")
cat("The tolerance of the optimization is set high so runtime is fast in this example.\n")
cat("In real applications, use the default or more stringent settings for tol.\n")
brtsM = 4:10
brtsS = seq(0.1, 3.5, 0.7)
tsplit = 5
dd_KI_ML(brtsM = brtsM, brtsS = brtsS, tsplit = tsplit, idparsopt = c(1:3, 6, 7),
```

```
initparsopt = c(0.885, 2e-14, 6.999, 6.848, 4.001), idparsfix = NULL, parsfix = NULL,
idparsnoshift = c(4,5), cond = 0, tol = c(3E-1,3E-1,3E-1))
```

---

|           |  |
|-----------|--|
| dd_KI_sim | <i>Function to simulate a key innovation in macro-evolution with the innovative clade decoupling from the diversity-dependent diversification dynamics of the main clade</i> |
|-----------|--|

---

## Description

Simulating a diversity-dependent diversification process where at a given time a new clade emerges with different inherent speciation rate and extinction rate and clade-level carrying capacity and with decoupled dynamics

## Usage

```
dd_KI_sim(pars, age, ddmodel = 1)
```

## Arguments

|         |  |
|---------|--|
| pars    | <p>Vector of parameters:</p> <p>pars[1] corresponds to <math>\lambda_M</math> (speciation rate of the main clade)<br/> pars[2] corresponds to <math>\mu_M</math> (extinction rate of the main clade)<br/> pars[3] corresponds to <math>K_M</math> (clade-level carrying capacity of the main clade)<br/> pars[4] corresponds to <math>\lambda_S</math> (speciation rate of the subclade)<br/> pars[5] corresponds to <math>\mu_S</math> (extinction rate of the subclade)<br/> pars[6] corresponds to <math>K_S</math> (clade-level carrying capacity of the subclade)<br/> pars[7] tinn, the time the shift in rates occurs in the lineage leading to the subclade</p>  |
| age     | Sets the crown age for the simulation  |
| ddmodel | <p>Sets the model of diversity-dependence:</p> <p>ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)<br/> ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)<br/> ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)<br/> ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity<br/> ddmodel == 2.2 : 1/n dependence in speciation rate<br/> ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)<br/> ddmodel == 3 : linear dependence in extinction rate<br/> ddmodel == 4 : exponential dependence in extinction rate<br/> ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity<br/> ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity</p> |

**Value**

out

A list with the following elements: The first element is the tree of extant species in phylo format  
 The second element is the tree of all species, including extinct species, in phylo format  
 The third element is a matrix of all species where  
 - the first column is the time at which a species is born  
 - the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage  
 - the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage  
 - the fourth column is the time of extinction of the species  
 If the fourth element equals -1, then the species is still extant.  
 - the fifth column indicates whether the species belong to the main clade (0) or the subclade (1)  
 The fourth element is the subclade tree of extant species (without stem)  
 The fifth element is the subclade tree of all species (without stem)  
 The sixth element is the same as the first, except that it has attributed 0 for the main clade and 1 for the subclade  
 The seventh element is the same as the Second, except that it has attributed 0 for the main clade and 1 for the subclade  
 The sixth and seventh element will be NULL if the subclade does not exist (because it went extinct).

**Author(s)**

Rampal S. Etienne

**References**

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**Examples**

```
dd_KI_sim(c(0.2,0.1,20,0.1,0.05,30,4),10)
```

---

dd\_loglik

---

*Loglikelihood for diversity-dependent diversification models*


---

**Description**

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values.

**Usage**

```
dd_loglik(pars1, pars2, brts, misnumspec, methode = 'analytical')
```



**Arguments**

|                          |   |
|--------------------------|---|
| <code>pars1</code>       | <p>Vector of parameters:</p> <p><code>pars1[1]</code> corresponds to <math>\lambda</math> (speciation rate)<br/> <code>pars1[2]</code> corresponds to <math>\mu</math> (extinction rate)<br/> <code>pars1[3]</code> corresponds to <math>K</math> (clade-level carrying capacity)</p>   |
| <code>pars2</code>       | <p>Vector of model settings:</p> <p><code>pars2[1]</code> sets the maximum number of species for which a probability must be computed. This must be larger than <math>1 + \text{missnumspec} + \text{length}(\text{brts})</math>.</p> <p><code>pars2[2]</code> sets the model of diversity-dependence:</p> <ul style="list-style-type: none"> <li>- <code>pars2[2] == 1</code> linear dependence in speciation rate with parameter <math>K</math> (= diversity where speciation = extinction)</li> <li>- <code>pars2[2] == 1.3</code> linear dependence in speciation rate with parameter <math>K'</math> (= diversity where speciation = 0)</li> <li>- <code>pars2[2] == 2</code> exponential dependence in speciation rate with parameter <math>K</math> (= diversity where speciation = extinction)</li> <li>- <code>pars2[2] == 2.1</code> variant of exponential dependence in speciation rate with offset at infinity</li> <li>- <code>pars2[2] == 2.2</code> <math>1/n</math> dependence in speciation rate</li> <li>- <code>pars2[2] == 2.3</code> exponential dependence in speciation rate with parameter <math>x</math> (= exponent)</li> <li>- <code>pars2[2] == 3</code> linear dependence in extinction rate</li> <li>- <code>pars2[2] == 4</code> exponential dependence in extinction rate</li> <li>- <code>pars2[2] == 4.1</code> variant of exponential dependence in extinction rate with offset at infinity</li> <li>- <code>pars2[2] == 4.2</code> <math>1/n</math> dependence in extinction rate</li> <li>- <code>pars2[2] == 5</code> linear dependence in speciation and extinction rate</li> </ul> <p><code>pars2[3]</code> sets the conditioning:</p> <ul style="list-style-type: none"> <li>- <code>pars2[3] == 0</code> conditioning on stem or crown age</li> <li>- <code>pars2[3] == 1</code> conditioning on stem or crown age and non-extinction of the phylogeny</li> <li>- <code>pars2[3] == 2</code> conditioning on stem or crown age and on the total number of extant taxa (including missing species)</li> <li>- <code>pars2[3] == 3</code> conditioning on the total number of extant taxa (including missing species)</li> </ul> <p><code>pars2[4]</code> sets whether the likelihood is for the branching times (0) or the phylogeny (1)</p> <p><code>pars2[5]</code> sets whether the parameters and likelihood should be shown on screen (1) or not (0)</p> <p><code>pars2[6]</code> sets whether the first data point is stem age (1) or crown age (2)</p> |
| <code>brts</code>        | A set of branching times of a phylogeny, all positive   |
| <code>missnumspec</code> | The number of species that are in the clade but missing in the phylogeny  |
| <code>methode</code>     | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.   |

**Value**

The loglikelihood

**Author(s)**

Rampal S. Etienne & Bart Haegeman

**References**

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**See Also**

[dd\\_ML](#), [dd\\_SR\\_loglik](#), [dd\\_KI\\_loglik](#)

**Examples**

```
dd_loglik(pars1 = c(0.5,0.1,100), pars2 = c(100,1,1,1,0,2), brts = 1:10, misssnumspec = 0)
```

---

|       |   |
|-------|---|
| dd_LR | <i>Bootstrap likelihood ratio test of diversity-dependent diversification model</i> |
|-------|---|

---

**Description**

This function computes the maximum likelihood and the associated estimates of the parameters of a diversity-dependent diversification model for a given set of phylogenetic branching times. It then performs a bootstrap likelihood ratio test of the diversity-dependent (DD) model against the constant-rates (CR) birth-death model. Finally, it computes the power of this test.

**Usage**

```
dd_LR(
  brts,
  initparsoptDD,
  initparsoptCR,
  misssnumspec,
  outputfilename = NULL,
  seed = 42,
  endmc = 1000,
  alpha = 0.05,
  plotit = TRUE,
  res = 10 * (1 + length(brts) + misssnumspec),
  ddmodel = 1,
  cond = 1,
  btorph = 1,
  soc = 2,
  tol = c(1E-3,1E-4,1E-6),
  maxiter = 2000,
  changeloglikifnoconv = FALSE,
```

```

    optimmethod = 'subplex',
    methode = 'analytical'
)

```

## Arguments

|                             |   |
|-----------------------------|---|
| <code>brts</code>           | A set of branching times of a phylogeny, all positive   |
| <code>initparsoptDD</code>  | The initial values of the parameters that must be optimized for the diversity-dependent (DD) model: <code>lambda_0</code> , <code>mu</code> and <code>K</code>  |
| <code>initparsoptCR</code>  | The initial values of the parameters that must be optimized for the constant-rates (CR) model: <code>lambda</code> and <code>mu</code>  |
| <code>missnumspec</code>    | The number of species that are in the clade but missing in the phylogeny  |
| <code>outputfilename</code> | The name (and location) of the file where the output will be saved. Default is no save.   |
| <code>seed</code>           | The seed for the pseudo random number generator for simulating the bootstrap data   |
| <code>endmc</code>          | The number of bootstraps  |
| <code>alpha</code>          | The significance level of the test  |
| <code>plotit</code>         | Boolean to plot results or not  |
| <code>res</code>            | Sets the maximum number of species for which a probability must be computed, must be larger than $1 + \text{length}(\text{brts})$   |
| <code>ddmodel</code>        | <p>Sets the model of diversity-dependence:</p> <p><code>ddmodel == 1</code> : linear dependence in speciation rate with parameter <code>K</code> (= diversity where speciation = extinction)</p> <p><code>ddmodel == 1.3</code> : linear dependence in speciation rate with parameter <code>K'</code> (= diversity where speciation = 0)</p> <p><code>ddmodel == 2</code> : exponential dependence in speciation rate with parameter <code>K</code> (= diversity where speciation = extinction)</p> <p><code>ddmodel == 2.1</code> : variant of exponential dependence in speciation rate with offset at infinity</p> <p><code>ddmodel == 2.2</code> : 1/n dependence in speciation rate</p> <p><code>ddmodel == 2.3</code> : exponential dependence in speciation rate with parameter <code>x</code> (= exponent)</p> <p><code>ddmodel == 3</code> : linear dependence in extinction rate</p> <p><code>ddmodel == 4</code> : exponential dependence in extinction rate</p> <p><code>ddmodel == 4.1</code> : variant of exponential dependence in extinction rate with offset at infinity</p> <p><code>ddmodel == 4.2</code> : 1/n dependence in extinction rate with offset at infinity</p> <p><code>ddmodel == 5</code> : linear dependence in speciation and extinction rate</p> |
| <code>cond</code>           | <p>Conditioning:</p> <p><code>cond == 0</code> : conditioning on stem or crown age</p> <p><code>cond == 1</code> : conditioning on stem or crown age and non-extinction of the phylogeny</p> <p><code>cond == 2</code> : conditioning on stem or crown age and on the total number of extant taxa (including missing species)</p> <p><code>cond == 3</code> : conditioning on the total number of extant taxa (including missing species)</p> <p>Note: <code>cond == 3</code> assumes a uniform prior on stem age, as is the standard in</p>  |

constant-rate birth-death models, see e.g. D. Aldous & L. Popovic 2004. Adv. Appl. Prob. 37: 1094-1115 and T. Stadler 2009. J. Theor. Biol. 261: 58-66. This conditioning turns out, for the diversity-dependent model, to provide the least-biased parameter estimates when extinction is low, but may still contain considerable bias if extinction is high. The default value has therefore been changed from 1 to 3 starting with DDD version 2.2.

|                      |   |
|----------------------|---|
| btorph               | Sets whether the likelihood is for the branching times (0) or the phylogeny (1)   |
| soc                  | Sets whether stem or crown age should be used (1 or 2)  |
| tol                  | Sets the tolerances in the optimization. Consists of:<br>reltolx = relative tolerance of parameter values in optimization<br>reltolf = relative tolerance of function value in optimization<br>abstolx = absolute tolerance of parameter values in optimization |
| maxiter              | Sets the maximum number of iterations in the optimization   |
| changeloglikifnoconv | if TRUE the loglik will be set to -Inf if ML does not converge  |
| optimmethod          | Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)  |
| methode              | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.                         |

## Details

The output is a list with 3 elements:

## Value

|        |   |
|--------|---|
| treeCR | a list of trees generated under the constant-rates model using the ML parameters under the CR model   |
| treeDD | a list of trees generated under the diversity-dependent model using the ML parameters under the diversity-dependent model   |
| out    | a dataframe with the parameter estimates and maximum likelihoods for diversity-dependent and constant-rates models \$model - the model used to generate the data. 0 = unknown (for real data), 1 = CR, 2 = DD<br>\$mc - the simulation number for each model<br>\$lambda_CR - speciation rate estimated under CR<br>\$mu_CR - extinction rate estimated under CR<br>\$LL_CR - maximum likelihood estimated under CR<br>\$conv_CR - convergence code for likelihood optimization; conv = 0 means convergence<br>\$lambda_DD1 - initial speciation rate estimated under DD for first set of initial values<br>\$mu_DD1 - extinction rate estimated under DD for first set of initial values<br>\$K_DD1 - clade-wide carrying-capacity estimated under DD for first set of initial values<br>\$LL_DD1 - maximum likelihood estimated under DD for first set of initial values<br>\$conv_DD1 - convergence code for likelihood optimization for first set of initial values; conv = 0 means convergence<br>\$lambda_DD2 - initial speciation rate estimated under DD for second set of initial values |

|              |  |
|--------------|--|
|              | \$mu_DD2 - extinction rate estimated under DD for second set of initial values   |
|              | \$K_DD2 - clade-wide carrying-capacity estimated under DD for second set of initial values                             |
|              | \$LL_DD2 - maximum likelihood estimated under DD for second set of initial values                                      |
|              | \$conv_DD2 - convergence code for likelihood optimization for second set of initial values; conv = 0 means convergence |
|              | \$LR - likelihood ratio between DD and CR  |
| pvalue       | p-value of the test  |
| LRalpha      | Likelihood ratio at the significance level alpha   |
| powerof test | power of the test for significance level alpha   |

**Author(s)**

Rampal S. Etienne & Bart Haegeman

**References**

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**See Also**

[dd\\_loglik](#), [dd\\_ML](#)

---

|       |  |
|-------|--|
| dd_ML | <i>Maximization of the loglikelihood under a diversity-dependent diversification model</i> |
|-------|--|

---

**Description**

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model for a given set of phylogenetic branching times. It also outputs the corresponding loglikelihood that can be used in model comparisons.

**Usage**

```
dd_ML(
  brts,
  initparsopt = if(ddmodel < 5) {c(0.2,0.1,2*(length(brts) + missnumspec))}
else {c(0.2,0.1,2*(length(brts) + missnumspec),0.01)},
  idparsopt = 1:length(initparsopt),
  idparsfix = (1:(3 + (ddmodel == 5)))[-idparsopt],
  parsfix = (ddmodel < 5) * c(0.2,0.1, 2 * (length(brts) + missnumspec))[-idparsopt]
+ (ddmodel == 5) * c(0.2, 0.1, 2 * (length(brts) + missnumspec), 0)[-idparsopt],
  res = 10 * (1 + length(brts) + missnumspec),
  ddmodel = 1,
  missnumspec = 0,
  cond = 1,
  btorph = 1,
  soc = 2,
```

```

tol = c(1E-3,1E-4,1E-6),
maxiter = 1000 * round((1.25)^length(idparsopt)),
changeloglikifnoconv = FALSE,
optimmethod = 'subplex',
methode = 'analytical'
)

```

## Arguments

|                    |   |
|--------------------|---|
| <b>brts</b>        | A set of branching times of a phylogeny, all positive   |
| <b>initparsopt</b> | The initial values of the parameters that must be optimized   |
| <b>idparsopt</b>   | The ids of the parameters that must be optimized, e.g. 1:3 for intrinsic speciation rate, extinction rate and carrying capacity. The ids are defined as follows:<br>id == 1 corresponds to lambda (speciation rate)<br>id == 2 corresponds to mu (extinction rate)<br>id == 3 corresponds to K (clade-level carrying capacity)<br>id == 4 corresponds to r ( $r = b/a$ where $\mu = \mu_0 + b * N$ and $\lambda = \lambda_0 - a * N$ ) (This is only available when <code>ddmodel = 5</code> )  |
| <b>idparsfix</b>   | The ids of the parameters that should not be optimized, e.g. c(1,3) if lambda and K should not be optimized, but only mu. In that case <code>idparsopt</code> must be 2. The default is to fix all parameters not specified in <code>idparsopt</code> .   |
| <b>parsfix</b>     | The values of the parameters that should not be optimized   |
| <b>res</b>         | Sets the maximum number of species for which a probability must be computed, must be larger than $1 + \text{length}(\text{brts})$   |
| <b>ddmodel</b>     | Sets the model of diversity-dependence:<br>ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)<br>ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)<br>ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)<br>ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity<br>ddmodel == 2.2 : 1/n dependence in speciation rate<br>ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)<br>ddmodel == 3 : linear dependence in extinction rate<br>ddmodel == 4 : exponential dependence in extinction rate<br>ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity<br>ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity<br>ddmodel == 5 : linear dependence in speciation and extinction rate |
| <b>misnumspec</b>  | The number of species that are in the clade but missing in the phylogeny  |
| <b>cond</b>        | Conditioning:<br>cond == 0 : conditioning on stem or crown age<br>cond == 1 : conditioning on stem or crown age and non-extinction of the phylogeny<br>cond == 2 : conditioning on stem or crown age and on the total number of extant taxa (including missing species)   |

|                      |  |
|----------------------|--|
|                      | cond == 3 : conditioning on the total number of extant taxa (including missing species)<br>Note: cond == 3 assumes a uniform prior on stem age, as is the standard in constant-rate birth-death models, see e.g. D. Aldous & L. Popovic 2004. Adv. Appl. Prob. 37: 1094-1115 and T. Stadler 2009. J. Theor. Biol. 261: 58-66. This conditioning turns out, for the diversity-dependent model, to provide the least-biased parameter estimates when extinction is low, but may still contain considerable bias if extinction is high. The default value has therefore been changed from 1 to 3 starting with DDD version 2.2. |
| btorph               | Sets whether the likelihood is for the branching times (0) or the phylogeny (1)  |
| soc                  | Sets whether stem or crown age should be used (1 or 2)   |
| tol                  | Sets the tolerances in the optimization. Consists of:<br>reltolx = relative tolerance of parameter values in optimization<br>reltolf = relative tolerance of function value in optimization<br>abstolx = absolute tolerance of parameter values in optimization  |
| maxiter              | Sets the maximum number of iterations in the optimization  |
| changeloglikifnoconv | if TRUE the loglik will be set to -Inf if ML does not converge   |
| optimmethod          | Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)   |
| methode              | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.  |

## Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor  $q! / (q + m)!$  where  $q$  is the number of species in the phylogeny and  $m$  is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

## Value

|        |  |
|--------|--|
| lambda | gives the maximum likelihood estimate of lambda  |
| mu     | gives the maximum likelihood estimate of mu  |
| K      | gives the maximum likelihood estimate of K   |
| r      | (only if ddmodel == 5) gives the ratio of linear dependencies in speciation and extinction rates |
| 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 & Bart Haegeman

## References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

**See Also**

[dd\\_loglik](#), [dd\\_SR\\_ML](#), [dd\\_KI\\_ML](#),

**Examples**

```
cat("Estimating the intrinsic speciation rate lambda and the carrying capacity K")
cat("for a fixed extinction rate of 0.1, conditioning on clade survival and two missing species:")
brts = 1:5
dd_ML(brts = brts, initparsopt = c(1.3078, 7.4188), idparsopt = c(1, 3), parsfix = 0.1,
      cond = 1, missnumspec = 2, tol = c(1E-3, 1E-3, 1E-4), optimmethod = 'simplex')
```

---

|              |   |
|--------------|---|
| dd_MS_loglik | <i>Loglikelihood for macro-evolutionary succession under diversity-dependent diversification with the key innovation at time <math>t = t_d</math></i> |
|--------------|---|

---

**Description**

This function computes the loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the diversity-dependent dynamics of an innovative subclade have different parameters from the dynamics of the main clade from time  $t_d$ , but both are governed by the same carrying capacity and experience each other's diversity.

**Usage**

```
dd_MS_loglik(pars1, pars2, brtsM, brtsS, missnumspec, methode = 'analytical')
```

**Arguments**

|       |  |
|-------|--|
| pars1 | Vector of parameters:<br><br>pars1[1] corresponds to $\lambda_M$ (speciation rate) of the main clade<br>pars1[2] corresponds to $\mu_M$ (extinction rate) of the main clade<br>pars1[3] corresponds to $K_M$ (clade-level carrying capacity) of the main clade<br>pars1[4] corresponds to $\lambda_m$ (speciation rate) of the subclade<br>pars1[5] corresponds to $\mu_s$ (extinction rate) of the subclade<br>pars1[6] corresponds to $t_d$ (the time of the key innovation)   |
| pars2 | Vector of model settings:<br><br>pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than $1 + \text{missnumspec} + \text{length}(\text{brts})$ .<br><br>pars2[2] sets the model of diversity-dependence:<br>- pars2[2] == 1 linear dependence in speciation rate with parameter $K$ (= diversity where speciation = extinction)<br>- pars2[2] == 1.3 linear dependence in speciation rate with parameter $K'$ (= diversity where speciation = 0)<br>- pars2[2] == 2 exponential dependence in speciation rate with parameter $K$ (= diversity where speciation = extinction)<br>- pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity<br>- pars2[2] == 2.2 $1/n$ dependence in speciation rate |



- pars2[2] == 2.3 exponential dependence in speciation rate with parameter  $x$  (= exponent)
- pars2[2] == 3 linear dependence in extinction rate
- pars2[2] == 4 exponential dependence in extinction rate
- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
- pars2[2] == 4.2  $1/n$  dependence in extinction rate

pars2[3] sets the conditioning:

- pars2[3] == 0 no conditioning
- pars2[3] == 1 conditioning on non-extinction of the phylogeny

pars2[4] sets the time of splitting of the branch that will undergo the key innovation leading to different parameters

pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)

pars2[6] sets whether the first data point is stem age (1) or crown age (2)

|             |  |
|-------------|--|
| brtsM       | A set of branching times of the main clade in the phylogeny, all positive  |
| brtsS       | A set of branching times of the subclade in the phylogeny, all positive  |
| missnumspec | The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector <code>c(missnumspec_M, missnumspec_S)</code> with missing species in main clade and subclade respectively. |
| methode     | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.                                    |

## Value

The loglikelihood

## Author(s)

Rampal S. Etienne & Bart Haegeman

## References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

## See Also

[dd\\_MS\\_ML](#), [dd\\_loglik](#), [dd\\_KI\\_loglik](#), [dd\\_SR\\_loglik](#)

## Examples

```
pars1 = c(0.2, 0.1, 40, 1.0, 0.1, 9.8)
pars2 = c(200, 1, 0, 18.8, 1, 2)
missnumspec = 0
brtsM = c(25.2, 24.6, 24.0, 22.5, 21.7, 20.4, 19.9, 19.7, 18.8, 17.1, 15.8, 11.8, 9.7, 8.9, 5.7, 5.2)
brtsS = c(9.6, 8.6, 7.4, 4.9, 2.5)
```

```
dd_MS_loglik(pars1,pars2,brtsM,brtsS,missnumspec,methode = 'ode45')
```

---

|          |  |
|----------|--|
| dd_MS_ML | <i>Maximization of the loglikelihood under a diversity-dependent diversification model with decoupling of a subclade's diversification dynamics from the main clade's dynamics</i> |
|----------|--|

---

## Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model where the diversity-dependent dynamics of an innovative subclade have different parameters from the dynamics of the main clade from time  $t_d$ , but both are governed by the same carrying capacity and experience each other's diversity. Required is a given set of phylogenetic branching times of main clade and subclade and the time of splitting of the lineage that will form the subclade. The function also outputs the corresponding loglikelihood that can be used in model comparisons.

## Usage

```
dd_MS_ML(
  brtsM,
  brtsS,
  tsplit,
  initparsopt = c(0.5, 0.1, 2 * (1 + length(brtsM) + length(brtsS) + sum(missnumspec)),
    (tsplit + max(brtsS))/2),
  parsfix = NULL,
  idparsopt = c(1:3, 6),
  idparsfix = NULL,
  idparsnoshift = (1:6)[c(-idparsopt, (-1)^(length(idparsfix) != 0) * idparsfix)],
  res = 10 * (1 + length(c(brtsM, brtsS)) + sum(missnumspec)),
  ddmodel = 1.3,
  missnumspec = 0,
  cond = 0,
  soc = 2,
  tol = c(1E-3, 1E-4, 1E-6),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  changeloglikifnoconv = FALSE,
  optimmethod = 'subplex',
  methode = 'analytical'
)
```

## Arguments

|             |   |
|-------------|---|
| brtsM       | A set of branching times of the main clade in a phylogeny, all positive             |
| brtsS       | A set of branching times of the subclade in a phylogeny, all positive               |
| tsplit      | The branching time at which the lineage forming the subclade branches off, positive |
| initparsopt | The initial values of the parameters that must be optimized                         |
| parsfix     | The values of the parameters that should not be optimized                           |

|               |  |
|---------------|--|
| idparsopt     | <p>The ids of the parameters that must be optimized, e.g. 1:7 for all parameters. The ids are defined as follows:</p> <p>id == 1 corresponds to lambda_M (speciation rate) of the main clade</p> <p>id == 2 corresponds to mu_M (extinction rate) of the main clade</p> <p>id == 3 corresponds to K_M (clade-level carrying capacity) of the main clade</p> <p>id == 4 corresponds to lambda_S (speciation rate) of the subclade</p> <p>id == 5 corresponds to mu_S (extinction rate) of the subclade</p> <p>id == 6 corresponds to t_d (the time of the key innovation)</p>   |
| idparsfix     | <p>The ids of the parameters that should not be optimized, e.g. c(1,3,4,6) if lambda and K should not be optimized, but only mu. In that case idparsopt must be c(2,5,7). The default is to fix all parameters not specified in idparsopt.</p>   |
| idparsnoshift | <p>The ids of the parameters that should not shift; This can only apply to ids 4, 5 and 6, e.g. idparsnoshift = c(4,5) means that lambda and mu have the same values before and after tshift</p>   |
| res           | <p>sets the maximum number of species for which a probability must be computed, must be larger than 1 + max(length(brtsM),length(brtsS))</p>   |
| ddmodel       | <p>sets the model of diversity-dependence:</p> <p>ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)</p> <p>ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)</p> <p>ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)</p> <p>ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity</p> <p>ddmodel == 2.2 : 1/n dependence in speciation rate</p> <p>ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)</p> <p>ddmodel == 3 : linear dependence in extinction rate</p> <p>ddmodel == 4 : exponential dependence in extinction rate</p> <p>ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity</p> <p>ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity</p> |
| missnumspec   | <p>The number of species that are in the clade but missing in the phylogeny. One can specify the sum of the missing species in main clade and subclade or a vector c(missnumspec_M,missnumspec_S) with missing species in main clade and subclade respectively.</p>  |
| cond          | <p>Conditioning:</p> <p>cond == 0 : no conditioning</p> <p>cond == 1 : conditioning on non-extinction of the phylogeny</p>   |
| soc           | <p>Sets whether stem or crown age should be used (1 or 2); stem age only works when cond = 0</p>   |
| tol           | <p>Sets the tolerances in the optimization. Consists of:</p> <p>reltolx = relative tolerance of parameter values in optimization</p> <p>reltolf = relative tolerance of function value in optimization</p> <p>abstolx = absolute tolerance of parameter values in optimization</p>   |
| maxiter       | <p>Sets the maximum number of iterations in the optimization</p>   |

|                      |  |
|----------------------|--|
| changeloglikifnoconv | if TRUE the loglik will be set to -Inf if ML does not converge   |
| optimmethod          | Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions) |
| methode              | The method used in the ode solver, default is ode45  |

### Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor  $q! / m!(q + m)!$  where  $q$  is the number of species in the phylogeny and  $m$  is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

### Value

|          |  |
|----------|--|
| lambda_M | gives the maximum likelihood estimate of lambda of the main clade          |
| mu_M     | gives the maximum likelihood estimate of mu of the main clade              |
| K_M      | gives the maximum likelihood estimate of K of the main clade               |
| lambda_2 | gives the maximum likelihood estimate of lambda of the subclade            |
| mu_S     | gives the maximum likelihood estimate of mu of the subclade                |
| t_d      | gives the time of the key innovation event                                 |
| 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 |

### Note

The optimization may get trapped in local optima. Try different starting values to search for the global optimum.

### Author(s)

Rampal S. Etienne & Bart Haegeman

### References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

### See Also

[dd\\_MS\\_loglik](#), [dd\\_ML](#), [dd\\_KI\\_ML](#), [dd\\_SR\\_ML](#),

### Examples

```
cat("This will estimate parameters for two sets of branching times brtsM, brtsS\n")
cat("without conditioning.\n")
cat("The tolerance of the optimization is set high so runtime is fast in this example.\n")
cat("In real applications, use the default or more stringent settings for tol.\n")
brtsM = 4:10
brtsS = seq(0.1, 3.5, 0.7)
tsplit = 5
```

```
dd_MS_ML(brtsM = brtsM, brtsS = brtsS, tsplit = tsplit, idparsopt = c(1:3,6),
  initparsopt = c(0.885, 2e-14, 10, 4.001), idparsfix = NULL, parsfix = NULL,
  idparsnoshift = c(4,5), cond = 0, tol = c(3E-1,3E-1,3E-1))
```

dd\_MS\_sim

*Function to simulate the macro-evolutionary succession process assuming diversity-dependent diversification*

## Description

Simulating a diversity-dependent diversification process where at a given time a new clade emerges with different inherent speciation rate and extinction rate

## Usage

```
dd_MS_sim(pars, age, ddmodel = 1.3)
```

## Arguments

**pars** Vector of parameters:

pars[1] corresponds to lambda\_M (speciation rate of the main clade)  
 pars[2] corresponds to mu\_M (extinction rate of the main clade)  
 pars[3] corresponds to K' (maximum number of species or a proxy for it in case of exponential decline in speciation rate) pars[4] corresponds to lambda\_S (speciation rate of the novel subclade)  
 pars[5] corresponds to mu\_S (extinction rate)  
 pars[6] tinn, the time the shift in rates occurs in the lineage leading to the subclade

**age** Sets the crown age for the simulation

**ddmodel** Sets the model of diversity-dependence:  
 ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0); ddmodel = 1 will be interpreted as this model  
 ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity; ddmodel = 2 will be interpreted as this model  
 ddmodel == 2.2 : 1/n dependence in speciation rate  
 ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)

## Value

**out** A list with the following elements: The first element is the tree of extant species in phylo format  
 The second element is the tree of all species, including extinct species, in phylo format  
 The third element is a matrix of all species where  
 - the first column is the time at which a species is born  
 - the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage

- the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage
  - the fourth column is the time of extinction of the species
- If the fourth element equals -1, then the species is still extant.
- the fifth column indicates whether the species belong to the main clade (0) or the subclade (1)
- The fourth element is the subclade tree of extant species (without stem)  
 The fifth element is the subclade tree of all species (without stem)  
 The sixth element is the same as the first, except that it has attributed 0 for the main clade and 1 for the subclade  
 The seventh element is the same as the Second, except that it has attributed 0 for the main clade and 1 for the subclade  
 The sixth and seventh element will be NULL if the subclade does not exist (because it went extinct).

### Author(s)

Rampal S. Etienne

### References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

### Examples

```
dd_MS_sim(c(0.2,0.1,20,0.1,0.05,4),10)
```

---

dd\_sim

*Function to simulate the diversity-dependent diversification process*

---

### Description

Simulating the diversity-dependent diversification process

### Usage

```
dd_sim(pars, age, ddmodel = 1)
```

### Arguments

|      |   |
|------|---|
| pars | Vector of parameters:<br><br>pars[1] corresponds to lambda (speciation rate)<br>pars[2] corresponds to mu (extinction rate)<br>pars[3] corresponds to K (clade-level carrying capacity) |
| age  | Sets the crown age for the simulation   |

ddmodel Sets the model of diversity-dependence:  
 ddmodel == 1 : linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)  
 ddmodel == 1.3 : linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)  
 ddmodel == 2 : exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)  
 ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity  
 ddmodel == 2.2 : 1/n dependence in speciation rate  
 ddmodel == 2.3 : exponential dependence in speciation rate with parameter x (= exponent)  
 ddmodel == 3 : linear dependence in extinction rate  
 ddmodel == 4 : exponential dependence in extinction rate  
 ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity  
 ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity  
 ddmodel == 5 : linear dependence in speciation and extinction rate

### Value

out A list with the following three elements: The first element is the tree of extant species in phylo format  
 The second element is the tree of all species, including extinct species, in phylo format  
 The third element is a matrix of all species where  
 - the first column is the time at which a species is born  
 - the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage  
 - the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage  
 - the fourth column is the time of extinction of the species  
 If the fourth element equals -1, then the species is still extant.

### Author(s)

Rampal S. Etienne

### References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

### Examples

```
dd_sim(c(0.2, 0.1, 20), 10)
```

---

|              |  |
|--------------|--|
| dd_SR_loglik | <i>Loglikelihood for diversity-dependent diversification models with a shift in the parameters at time <math>t = tshift</math></i> |
|--------------|--|

---

### Description

This function computes loglikelihood of a diversity-dependent diversification model for a given set of branching times and parameter values where the parameters are allowed to shift at time  $t = tshift$

### Usage

```
dd_SR_loglik(pars1, pars2, brts, missnumspec, methode = 'analytical')
```

### Arguments

**pars1**                Vector of parameters:

pars1[1] corresponds to lambda (speciation rate) before the shift  
 pars1[2] corresponds to mu (extinction rate) before the shift  
 pars1[3] corresponds to K (clade-level carrying capacity) before the shift  
 pars1[4] corresponds to lambda (speciation rate) after the shift  
 pars1[5] corresponds to mu (extinction rate) after the shift  
 pars1[6] corresponds to K (clade-level carrying capacity) after the shift  
 pars1[7] corresponds to tshift (the time of shift)

**pars2**                Vector of model settings:

pars2[1] sets the maximum number of species for which a probability must be computed. This must be larger than  $1 + missnumspec + \text{length}(brts)$ .

pars2[2] sets the model of diversity-dependence:

- pars2[2] == 1 linear dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 1.3 linear dependence in speciation rate with parameter K' (= diversity where speciation = 0)
- pars2[2] == 2 exponential dependence in speciation rate with parameter K (= diversity where speciation = extinction)
- pars2[2] == 2.1 variant of exponential dependence in speciation rate with offset at infinity
- pars2[2] == 2.2 1/n dependence in speciation rate
- pars2[2] == 2.3 exponential dependence in speciation rate with parameter x (= exponent)
- pars2[2] == 3 linear dependence in extinction rate
- pars2[2] == 4 exponential dependence in extinction rate
- pars2[2] == 4.1 variant of exponential dependence in extinction rate with offset at infinity
- pars2[2] == 4.2 1/n dependence in extinction rate

pars2[3] sets the conditioning:

- pars2[3] == 0 no conditioning
- pars2[3] == 1 conditioning on non-extinction of the phylogeny
- pars2[3] == 2 conditioning on non-extinction of the phylogeny and on the



|             |   |
|-------------|---|
|             | total number of extant taxa (including missing species)   |
|             | pars2[4] sets whether the likelihood is for the branching times (0) or the phylogeny (1)  |
|             | pars2[5] sets whether the parameters and likelihood should be shown on screen (1) or not (0)  |
|             | pars2[6] sets whether the first data point is stem age (1) or crown age (2)   |
| brts        | A set of branching times of a phylogeny, all positive   |
| missnumspec | The number of species that are in the clade but missing in the phylogeny  |
| methode     | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1. |

### Value

The loglikelihood

### Author(s)

Rampal S. Etienne & Bart Haegeman

### References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

### See Also

[dd\\_SR\\_ML](#), [dd\\_loglik](#), [dd\\_KI\\_loglik](#)

### Examples

```
dd_SR_loglik(pars1 = c(0.2,0.1,50,0.2,0.1,70,5), pars2 = c(100,1,1,1,0,2),
  brts = 1:10, missnumspec = 0)
```

---

dd\_SR\_ML

*Maximization of the loglikelihood under a diversity-dependent diversification model with a shift in the parameters*

---

### Description

This function computes the maximum likelihood estimates of the parameters of a diversity-dependent diversification model with shifting parameters at time  $t = tshift$  for a given set of phylogenetic branching times. It also outputs the corresponding loglikelihood that can be used in model comparisons.

**Usage**

```
dd_SR_ML(
  brts,
  initparsopt = c(0.5, 0.1, 2 * (1 + length(brts) + missnumspec),
  2 * (1 + length(brts) + missnumspec), max(brts)/2),
  parsfix = NULL,
  idparsopt = c(1:3, 6:7),
  idparsfix = NULL,
  idparsnoshift = (1:7)[c(-idparsopt, (-1)^(length(idparsfix) != 0) * idparsfix)],
  res = 10 * (1 + length(brts) + missnumspec),
  ddmodel = 1,
  missnumspec = 0,
  cond = 1,
  btorph = 1,
  soc = 2,
  allbp = FALSE,
  tol = c(1E-3, 1E-4, 1E-6),
  maxiter = 1000 * round((1.25)^length(idparsopt)),
  changeloglikifnoconv = FALSE,
  optimmethod = 'subplex',
  methode = 'analytical'
)
```

**Arguments**

|                            |   |
|----------------------------|---|
| <code>brts</code>          | A set of branching times of a phylogeny, all positive   |
| <code>initparsopt</code>   | The initial values of the parameters that must be optimized   |
| <code>parsfix</code>       | The values of the parameters that should not be optimized   |
| <code>idparsopt</code>     | The ids of the parameters that must be optimized, e.g. 1:7 for all parameters.<br>The ids are defined as follows:<br>id == 1 corresponds to lambda (speciation rate) before the shift<br>id == 2 corresponds to mu (extinction rate) before the shift<br>id == 3 corresponds to K (clade-level carrying capacity) before the shift<br>id == 4 corresponds to lambda (speciation rate) after the shift<br>id == 5 corresponds to mu (extinction rate) after the shift<br>id == 6 corresponds to K (clade-level carrying capacity) after the shift<br>id == 7 corresponds to tshift (the time of shift) |
| <code>idparsfix</code>     | The ids of the parameters that should not be optimized, e.g. c(1,3,4,6) if lambda and K should not be optimized, but only mu. In that case idparsopt must be c(2,5,7). The default is to fix all parameters not specified in idparsopt.   |
| <code>idparsnoshift</code> | The ids of the parameters that should not shift; This can only apply to ids 4, 5 and 6, e.g. idparsnoshift = c(4,5) means that lambda and mu have the same values before and after tshift   |
| <code>res</code>           | sets the maximum number of species for which a probability must be computed, must be larger than 1 + length(brts)   |
| <code>ddmodel</code>       | sets the model of diversity-dependence:<br>ddmodel == 1 : linear dependence in speciation rate<br>ddmodel == 2 : exponential dependence in speciation rate<br>ddmodel == 2.1 : variant of exponential dependence in speciation rate with offset at infinity   |

|                      |   |
|----------------------|---|
|                      | ddmodel == 2.2 : 1/n dependence in speciation rate  |
|                      | ddmodel == 3 : linear dependence in extinction rate   |
|                      | ddmodel == 4 : exponential dependence in extinction rate  |
|                      | ddmodel == 4.1 : variant of exponential dependence in extinction rate with offset at infinity   |
|                      | ddmodel == 4.2 : 1/n dependence in extinction rate with offset at infinity  |
| missnumspec          | The number of species that are in the clade but missing in the phylogeny  |
| cond                 | Conditioning:<br>cond == 0 : no conditioning<br>cond == 1 : conditioning on non-extinction of the phylogeny<br>cond == 2 : conditioning on non-extinction of the phylogeny and on the total number of extant taxa (including missing species)<br>cond == 3 : conditioning on the total number of extant taxa (including missing species)<br>(including missing species)<br>Note: cond == 3 assumes a uniform prior on stem age, as is the standard in constant-rate birth-death models, see e.g. D. Aldous & L. Popovic 2004. Adv. Appl. Prob. 37: 1094-1115 and T. Stadler 2009. J. Theor. Biol. 261: 58-66. |
| btorph               | Sets whether the likelihood is for the branching times (0) or the phylogeny (1)   |
| soc                  | Sets whether stem or crown age should be used (1 or 2)  |
| allbp                | Sets whether a search should be done with various initial conditions, with tshift at each of the branching points (TRUE/FALSE)  |
| tol                  | Sets the tolerances in the optimization. Consists of:<br>reltolx = relative tolerance of parameter values in optimization<br>reltolf = relative tolerance of function value in optimization<br>abstolx = absolute tolerance of parameter values in optimization   |
| maxiter              | Sets the maximum number of iterations in the optimization   |
| changeloglikifnoconv | if TRUE the loglik will be set to -Inf if ML does not converge  |
| optimmethod          | Method used in optimization of the likelihood. Current default is 'subplex'. Alternative is 'simplex' (default of previous versions)  |
| methode              | The method used to solve the master equation, default is 'analytical' which uses matrix exponentiation; alternatively numerical ODE solvers can be used, such as 'lsoda' or 'ode45'. These were used in the package before version 3.1.   |

## Details

The output is a dataframe containing estimated parameters and maximum loglikelihood. The computed loglikelihood contains the factor  $q! \cdot m! / (q + m)!$  where  $q$  is the number of species in the phylogeny and  $m$  is the number of missing species, as explained in the supplementary material to Etienne et al. 2012.

## Value

|          |  |
|----------|--|
| lambda_1 | gives the maximum likelihood estimate of lambda before the shift |
| mu_1     | gives the maximum likelihood estimate of mu before the shift     |
| K_1      | gives the maximum likelihood estimate of K before the shift      |
| lambda_2 | gives the maximum likelihood estimate of lambda after the shift  |
| mu_2     | gives the maximum likelihood estimate of mu after the shift      |

|         |  |
|---------|--|
| K_2     | gives the maximum likelihood estimate of K after the shift                 |
| t_shift | gives the time of the shift  |
| 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 |

### Note

The optimization may get trapped in local optima. Try different starting values to search for the global optimum.

### Author(s)

Rampal S. Etienne & Bart Haegeman

### References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

### See Also

[dd\\_SR\\_loglik](#), [dd\\_ML](#), [dd\\_KI\\_ML](#),

### Examples

```
cat("This will estimate parameters for a sets of branching times brts without conditioning.\n")
cat("The tolerance of the optimization is set ridiculously high to make runtime fast.\n")
cat("In real applications, use the default or more stringent settings for tol.\n")
brts = 1:10
dd_SR_ML(brts = brts, initparsopt = c(0.4581, 1E-6, 17.69, 11.09, 8.9999), idparsopt = c(1:3,6,7),
        idparsfix = NULL, parsfix = NULL, idparsnoshift = c(4,5), cond = 0,
        tol = c(1E-1,1E-1,1E-1),optimmethod = 'simplex'
)
```

---

dd\_SR\_sim

*Function to simulate the diversity-dependent diversification process  
with a shift in one or more of the parameters*

---

### Description

Simulating the diversity-dependent diversification process with a parameter shift at a certain time

### Usage

```
dd_SR_sim(pars, age, ddmodel = 1)
```

**Arguments**

|                      |   |
|----------------------|---|
| <code>pars</code>    | <p>Vector of parameters:</p> <p><code>pars[1]</code> corresponds to <code>lambda1</code> (speciation rate before the rate shift)<br/> <code>pars[2]</code> corresponds to <code>mu1</code> (extinction rate before the rate shift)<br/> <code>pars[3]</code> corresponds to <code>K1</code> (clade-level carrying capacity before the rate shift)<br/> <code>pars[4]</code> corresponds to <code>lambda2</code> (speciation rate after the rate shift)<br/> <code>pars[5]</code> corresponds to <code>mu2</code> (extinction rate after the rate shift)<br/> <code>pars[6]</code> corresponds to <code>K2</code> (clade-level carrying capacity after the rate shift)<br/> <code>pars[7]</code> corresponds to the time of shift</p>  |
| <code>age</code>     | Sets the crown age for the simulation   |
| <code>ddmodel</code> | <p>Sets the model of diversity-dependence:</p> <p><code>ddmodel == 1</code> : linear dependence in speciation rate with parameter <code>K</code> (= diversity where speciation = extinction)<br/> <code>ddmodel == 1.3</code> : linear dependence in speciation rate with parameter <code>K'</code> (= diversity where speciation = 0)<br/> <code>ddmodel == 2</code> : exponential dependence in speciation rate with parameter <code>K</code> (= diversity where speciation = extinction)<br/> <code>ddmodel == 2.1</code> : variant of exponential dependence in speciation rate with offset at infinity<br/> <code>ddmodel == 2.2</code> : 1/n dependence in speciation rate<br/> <code>ddmodel == 2.3</code> : exponential dependence in speciation rate with parameter <code>x</code> (= exponent)<br/> <code>ddmodel == 3</code> : linear dependence in extinction rate<br/> <code>ddmodel == 4</code> : exponential dependence in extinction rate<br/> <code>ddmodel == 4.1</code> : variant of exponential dependence in extinction rate with offset at infinity<br/> <code>ddmodel == 4.2</code> : 1/n dependence in extinction rate with offset at infinity<br/> <code>ddmodel == 5</code> : linear dependence in speciation and extinction rate</p> |

**Value**

|                  |  |
|------------------|--|
| <code>out</code> | <p>A list with the following three elements: The first element is the tree of extant species in phylo format<br/> The second element is the tree of all species, including extinct species, in phylo format<br/> The third element is a matrix of all species where</p> <ul style="list-style-type: none"> <li>- the first column is the time at which a species is born</li> <li>- the second column is the label of the parent of the species; positive and negative values only indicate whether the species belongs to the left or right crown lineage</li> <li>- the third column is the label of the daughter species itself; positive and negative values only indicate whether the species belongs to the left or right crown lineage</li> <li>- the fourth column is the time of extinction of the species</li> </ul> <p>If the fourth element equals -1, then the species is still extant.</p> |
|------------------|--|

**Author(s)**

Rampal S. Etienne

## References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

## Examples

```
dd_SR_sim(c(0.2,0.1,20,0.2,0.1,40,5),10)
```

---

|         |   |
|---------|---|
| L2phylo | <i>Function to convert a table with speciation and extinction events to a phylogeny</i> |
|---------|---|

---

## Description

Converting a table with speciation and extinction events to a phylogeny

## Usage

```
L2phylo(L, dropextinct = T)
```

## Arguments

|             |   |
|-------------|---|
| L           | Matrix of events as produced by dd_sim: <ul style="list-style-type: none"> <li>- the first column is the time at which a species is born in Mya</li> <li>- the second column is the label of the parent of the species; positive and negative values indicate whether the species belongs to the left or right crown lineage</li> <li>- the third column is the label of the daughter species itself; positive and negative values indicate whether the species belongs to the left or right crown lineage</li> <li>- the fourth column is the time of extinction of the species; if the fourth element equals -1, then the species is still extant.</li> </ul> |
| dropextinct | Sets whether the phylogeny should drop species that are extinct at the present  |

## Value

|     |                               |
|-----|-------------------------------|
| phy | A phylogeny of the phylo type |
|-----|-------------------------------|

## Author(s)

Rampal S. Etienne

## References

- Etienne, R.S. et al. 2012, Proc. Roy. Soc. B 279: 1300-1309, doi: 10.1098/rspb.2011.1439
- Etienne, R.S. & B. Haegeman 2012. Am. Nat. 180: E75-E89, doi: 10.1086/667574

## Examples

```
sim = dd_sim(c(0.2,0.1,20),10)
phy = L2phylo(sim$L)
plot(phy)
```

---

|           |   |
|-----------|---|
| optimizer | <i>Carries out optimization (finding a minimum)</i> |
|-----------|---|

---

**Description**

A wrapper to use several optimization routines, currently only 'simplex' (a method adopted from Matlab, or 'subplex', from the R package subplex). The function is called from several packages by the same author.

**Usage**

```
optimizer(optimmethod = 'simplex', optimpars = c(1E-4, 1E-4, 1E-6, 1000), fun, trparsopt, ...)
```

**Arguments**

|             |  |
|-------------|--|
| optimmethod | The method to use for optimization, either 'simplex' or 'subplex'  |
| optimpars   | Parameters of the optimization: relative tolerance in function arguments, relative tolerance in function value, absolute tolerance in function arguments, and maximum number of iterations |
| fun         | Function to be optimized   |
| trparsopt   | Initial guess of the parameters to be optimized  |
| ...         | Any other arguments of the function to be optimized, or settings of the optimization routine   |

**Value**

|     |   |
|-----|---|
| out | A list containing optimal function arguments (par, the optimal function value (fvalues) and whether the optimization converged (conv) |
|-----|---|

.

**Author(s)**

Rampal S. Etienne

**Examples**

```
cat("No examples")
```

---

|        |                                      |
|--------|--------------------------------------|
| roundn | <i>Rounds up in the usual manner</i> |
|--------|--------------------------------------|

---

**Description**

The standard round function in R rounds x.5 to the nearest even integer. This is odd behavior that is corrected in roundn

**Usage**

```
roundn(x, digits = 0)
```

**Arguments**

|        |  |
|--------|--|
| x      | Number to be rounded                     |
| digits | Sets the number of decimals in rounding. |

**Value**

|   |          |
|---|----------|
| n | A number |
|---|----------|

**Author(s)**

Rampal S. Etienne

**Examples**

```
round(2.5)
roundn(2.5)
round(3.5)
roundn(3.5)
round(2.65,digits = 1)
roundn(2.65,digits = 1)
round(2.75,digits = 1)
roundn(2.75,digits = 1)
```

---

sample2

*Takes samples in the usual manner*

---

**Description**

The standard sample function in R samples from n numbers when x = n. This is unwanted behavior when the size of the vector to sample from changes dynamically. This is corrected in sample2

**Usage**

```
sample2(x,size,replace = FALSE,prob = NULL)
```

**Arguments**

|         |   |
|---------|---|
| x       | A vector of one or more elements  |
| size    | A non-negative integer giving the number of items to choose.                            |
| replace | Should sampling be with replacement?  |
| prob    | A vector of probability weights for obtaining the elements of the vector being sampled. |

**Value**

|     |   |
|-----|---|
| sam | A vector of length size that is sampled from x. |
|-----|---|

**Author(s)**

Rampal S. Etienne



**Examples**

```
sample(x = 10,size = 5,replace = TRUE)  
sample2(x = 10,size = 5,replace = TRUE)
```

# Index

## \*Topic **models**

- bd\_loglik, [6](#)
- bd\_ML, [7](#)
- dd\_KI\_loglik, [10](#)
- dd\_KI\_ML, [12](#)
- dd\_KI\_sim, [15](#)
- dd\_loglik, [16](#)
- dd\_LR, [18](#)
- dd\_ML, [21](#)
- dd\_MS\_loglik, [24](#)
- dd\_MS\_ML, [26](#)
- dd\_MS\_sim, [29](#)
- dd\_sim, [30](#)
- dd\_SR\_loglik, [32](#)
- dd\_SR\_ML, [33](#)
- dd\_SR\_sim, [36](#)
- DDD-package, [2](#)
- L2phylo, [38](#)
- optimizer, [39](#)
- roundn, [39](#)
- sample2, [40](#)

bd\_loglik, [6](#), [9](#)

bd\_ML, [7](#), [7](#)

dd\_KI\_loglik, [10](#), [14](#), [18](#), [25](#), [33](#)

dd\_KI\_ML, [11](#), [12](#), [24](#), [28](#), [36](#)

dd\_KI\_sim, [15](#)

dd\_loglik, [11](#), [16](#), [21](#), [24](#), [25](#), [33](#)

dd\_LR, [18](#)

dd\_ML, [14](#), [18](#), [21](#), [21](#), [28](#), [36](#)

dd\_MS\_loglik, [24](#), [28](#)

dd\_MS\_ML, [25](#), [26](#)

dd\_MS\_sim, [29](#)

dd\_sim, [30](#)

dd\_SR\_loglik, [11](#), [18](#), [25](#), [32](#), [36](#)

dd\_SR\_ML, [14](#), [24](#), [28](#), [33](#), [33](#)

dd\_SR\_sim, [36](#)

DDD (DDD-package), [2](#)

DDD-package, [2](#)

L2phylo, [38](#)

optimizer, [39](#)

roundn, [39](#)

sample2, [40](#)