rdrr.io Find an R package R language docs Run R in your browser R Notebooks packages, doc text, code...

fit_env: Maximum likelihood fit of the environmental birth-death model

In RPANDA: Phylogenetic ANalyses of DiversificAtion

Description Usage Arguments Details Value Note Author(s) References See Also Examples

View source: R/fit_env.R

Description

Fits the environmental birth-death model with potentially missing extant species to a phylogeny, by maximum likelihood. Notations follow Morlon et al. PNAS 2011 and Condamine et al. ELE 2013.

Usage

```
fit_env(phylo, env_data, tot_time, f.lamb, f.mu, lamb_par, mu_par, df= NULL, f = 1,
       meth = "Nelder-Mead", cst.lamb = FALSE, cst.mu = FALSE,
       expo.lamb = FALSE, expo.mu = FALSE, fix.mu = FALSE,
       dt=0, cond = "crown")
```

Arguments

phylo	an object of type 'phylo' (see ape documentation)
env_data	environmental data, given as a data frame with two columns. The first column is time, the second column is the environmental data (temperature for instance).
tot_time	the age of the phylogeny (crown age, or stem age if known). If working with crown ages, tot_time is given by max(node.age(phylo)$ages).
f.lamb	a function specifying the hypothesized functional form of the variation of the speciation rate λ with time and the environmental variable. Any functional form may be used. This function has three arguments: the first argument is time; the second argument is the environmental variable; the third arguement is a numeric vector of the parameters controlling the time and environmental variation (to be estimated).
f.mu	a function specifying the hypothesized functional form of the variation of the extinction rate μ with time and the environmental variable. Any functional form may be used. This function has three arguments: the first argument is time; the second argument is the environmental variable; the second argument is a numeric vector of the parameters controlling the time and environmental variation (to be estimated).
lamb_par	a numeric vector of initial values for the parameters of f.lamb to be estimated (these values are used by the optimization algorithm). The length of this vector is used to compute the total number of parameters in the model, so to fit a model with constant speciation rate (for example), lamb_par should be a vector of length 1. Otherwise aic values will be wrong.
mu_par	a numeric vector of initial values for the parameters of f.mu to be estimated (these values are used by the optimization algorithm). The length of this vector is used to compute the total number of parameters in the model, so to fit a model without extinction (for example), mu_par should be empty (vector of length 0). Otherwise aic values will be wrong.
df	the degree of freedom to use to define the spline. As a default, smooth.spline(env_data[,1], env_data for details.

Improve this page

`meth`	optimization to use to maximize the likelihood function, see *optim* for more details.
`cst.lamb`	logical: should be set to TRUE only if f.lamb is constant (i.e. does not depend on time or the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
`cst.mu`	logical: should be set to TRUE only if f.mu is constant (i.e. does not depend on time or the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
`expo.lamb`	logical: should be set to TRUE only if f.lamb is an exponential function of time (and does not depend on the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
`expo.mu`	logical: should be set to TRUE only if f.mu is an exponential function of time (and does not depend on the environmental variable) to use analytical instead of numerical computation in order to reduce computation time.
`fix.mu`	logical: if set to TRUE, the extinction rate μ is fixed and will not be optimized.
`dt`	the default value is 0. In this case, integrals in the likelihood are computed using R "integrate" function, which can be quite slow. If a positive dt is given as argument, integrals are computed using a piece-wise contant approximation, and dt represents the length of the intervals on which functions are assumed to be constant. We found that 1e-3 generally provides a good trade-off between precision and computation time.
`cond`	conditioning to use to fit the model: • FALSE: no conditioning (not recommended); • "stem": conditioning on the survival of the stem lineage (use when the stem age is known, in this case tot_time should be the stem age); • "crown" (default): conditioning on a speciation event at the crown age and survival of the 2 daugther lineages (use when the stem age is not known, in this case tot_time should be the crown age).

Details

The lengths of lamb_par and mu_par are used to compute the total number of parameters in the model, so to fit a model with constant speciation rate (for example), lamb_par should be a vector of length 1. Otherwise aic values will be wrong. In the f.lamb and f.mu functions, time runs from the present to the past.

Value

a list with the following components

`model`	the name of the fitted model
`LH`	the maximum log-likelihood value
`aicc`	the second order Akaike's Information Criterion
`lamb_par`	a numeric vector of estimated f.lamb parameters, in the same order as defined in f.lamb
`mu_par`	a numeric vector of estimated f.mu parameters, in the same order as defined in f.mu (if fix.mu is FALSE)

Note

The speed of convergence of the fit might depend on the degree of freedom chosen to define the spline.

Author(s)

H Morlon and F Condamine

Morlon, H., Parsons, T.L. and Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record *Proc Nat Acad Sci* 108: 16327-16332

Condamine, F.L., Rolland, J., and Morlon, H. (2013) Macroevolutionary perspectives to environmental change, Eco Lett 16: 72-85

Morlon, H. (2014) Phylogenetic approaches for studying diversification, Eco Lett, 17:508-525

See Also

`plot_fit_env` , `fit_bd` , `likelihood_bd`

Examples

```
data(Cetacea)
tot_time<-max(node.age(Cetacea)$ages)
data(InfTemp)
dof<-smooth.spline(InfTemp[,1], InfTemp[,2])$df

# Fits a model with lambda varying as an exponential function of temperature
# and mu fixed to 0 (no extinction).  Here t stands for time and x for temperature.
f.lamb <-function(t,x,y){y[1] * exp(y[2] * x)}
f.mu<-function(t,x,y){0}
lamb_par<-c(0.10, 0.01)
mu_par<-c()
#result_exp <- fit_env(Cetacea,InfTemp,tot_time,f.lamb,f.mu,lamb_par,mu_par,
#                      f=87/89,fix.mu=TRUE,df=dof,dt=1e-3)
```

Run this example Create a free Jupyter Notebook Embed on your website

RPANDA documentation built on May 29, 2017, 3:34 p.m.

Related to fit_env in RPANDA...

fit_env

RPANDA index Package overview
README.md

R Package Documentation
rdrr.io home
R language documentation
Run R code online
Create free R Jupyter Notebooks

Browse R Packages
CRAN packages
Bioconductor packages
R-Forge packages
GitHub packages

We want your feedback!
Note that we can't provide technical support on individual packages. You should contact the package authors for that.
Tweet to @rdrrHQ
GitHub issue tracker
ian@mutexlabs.com
Personal blog