

Package ‘BAT’

April 10, 2020

Title Biodiversity Assessment Tools

Version 2.0.1

Date 2020-04-10

Author Pedro Cardoso, Stefano Mammola, Francois Rigal, Jose Carlos Carvalho

Maintainer Pedro Cardoso <pedro.cardoso@helsinki.fi>

Description Includes algorithms to assess alpha and beta diversity in all their dimensions (taxon, phylogenetic and functional diversity), whether communities are completely sampled or not. It allows performing a number of analyses based on either species identities or phylogenetic/functional trees or functional kernel n-dimensional hypervolumes depicting species relationships.

Depends R (>= 3.0.0)

Imports graphics, hypervolume, nls2, raster, spatstat, stats, utils, vegan

Encoding UTF-8

License GPL-3

LazyData true

RoxygenNote 7.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2020-04-10 07:30:02 UTC

R topics documented:

accuracy	2
alpha	4
alpha.accum	5
alpha.estimate	7
arrabida	9
beta	10
beta.accum	11

beta.multi	13
contribution	14
dispersion	15
evenness	17
functree	18
gdm	18
geres	19
guadiana	20
iaor	20
kernel.alpha	21
kernel.beta	23
kernel.contribution	25
kernel.dispersion	27
kernel.evenness	28
kernel.originality	30
kernel.similarity	32
linnean	33
optim.alpha	34
optim.alpha.stats	35
optim.beta	36
optim.beta.stats	38
optim.spatial	39
originality	40
phylotree	41
raster.alpha	41
raster.beta	42
raster.dispersion	43
raster.evenness	45
sad	46
sar	47
sim.plot	48
sim.sad	49
sim.sample	50
sim.spatial	51
sim.tree	52
slope	53
uniqueness	54

Index	55
--------------	-----------

accuracy	<i>Scaled mean squared error of accumulation curves.</i>
----------	--

Description

Accuracy (scaled mean squared error) of accumulation curves compared with a known true diversity value (target).

Usage

```
accuracy(accum, target = -1)
```

Arguments

accum	A matrix resulting from the alpha.accum or beta.accum functions (sampling units x diversity values).
target	The true known diversity value, with which the curve will be compared. If not specified, default is the diversity observed with all sampling units.

Details

Among multiple measures of accuracy (Walther & Moore 2005) the SMSE presents several advantages, as it is (Cardoso et al. 2014): (i) scaled to true diversity, so that similar absolute differences are weighted according to how much they represent of the real value; (ii) scaled to the number of sampling units, so that values are independent of sample size; (iii) squared, so that small, mostly meaningless fluctuations around the true value are down-weighted; and (iv) independent of positive or negative deviation from the real value, as such differentiation is usually not necessary. For alpha diversity accuracy may also be weighted according to how good the data is predicted to be. The weight of each point in the curve is proportional to its sampling intensity (i.e. $n/Sobs$).

Value

Accuracy values (both raw and weighted) for all observed and estimated curves.

References

Cardoso, P., Rigal, F., Borges, P.A.V. & Carvalho, J.C. (2014) A new frontier in biodiversity inventory: a proposal for estimators of phylogenetic and functional diversity. *Methods in Ecology and Evolution*, in press.

Walther, B.A. & Moore, J.L. (2005) The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28, 815-829.

Examples

```
comm1 <- matrix(c(2,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,0,2,2), nrow = 4, ncol = 5, byrow = TRUE)
comm2 <- matrix(c(1,1,0,0,0,0,2,1,0,0,0,0,2,1,0,0,0,0,2,1), nrow = 4, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5)), method="euclidean", method="average")
acc.alpha = alpha.accum(comm1)
accuracy(acc.alpha)
accuracy(acc.alpha, 10)
acc.beta = beta.accum(comm1, comm2, tree)
accuracy(acc.beta)
accuracy(acc.beta, c(1,1,0))
```

alpha	<i>Alpha diversity (Taxon, Phylogenetic or Functional Diversity - TD, PD, FD).</i>
-------	--

Description

Observed richness with possible rarefaction, multiple sites simultaneously.

Usage

```
alpha(comm, tree, raref = 0, runs = 100)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data.
tree	An hclust or phylo object (used only for PD or FD).
raref	An integer specifying the number of individuals for rarefaction (individual based). If raref < 1 no rarefaction is made. If raref = 1 rarefaction is made by the minimum abundance among all sites. If raref > 1 rarefaction is made by the abundance indicated. If not specified, default is 0.
runs	Number of resampling runs for rarefaction. If not specified, default is 100.

Details

TD is equivalent to species richness. Calculations of PD and FD are based on Faith (1992) and Petchey & Gaston (2002, 2006), which measure PD and FD of a community as the total branch length of a tree linking all species represented in such community. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in comm must be the same as in tree. The rarefaction option is useful to compare communities with much different numbers of individuals sampled, which might bias diversity comparisons (Gotelli & Colwell 2001)

Value

A matrix of sites x diversity values (either "Obs" OR "Median, Min, LowerCL, UpperCL and Max").

References

- Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1-10.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4, 379-391.
- Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, 5, 402-411.
- Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters*, 9, 741-758.

Examples

```

comm <- matrix(c(0,0,1,1,0,0,2,1,0,0), nrow = 2, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
alpha(comm)
alpha(comm, raref = 0)
alpha(comm, tree, 2, 100)

```

alpha.accum

*Alpha diversity accumulation curves (observed and estimated).***Description**

Estimation of alpha diversity of a single site with accumulation of sampling units.

Usage

```

alpha.accum(
  comm,
  tree,
  func = "nonparametric",
  target = -2,
  runs = 100,
  prog = TRUE
)

```

Arguments

comm	A sampling units x species matrix, with either abundance or incidence data.
tree	An hclust or phylo object (used only for Phylogenetic (PD) or Functional (FD) Diversity, not for Taxon Diversity (TD)).
func	The class of estimators to be used: If func is partial match of "curve", TD, PD or FD are based on extrapolating the accumulation curve of observed diversity. If func is partial match of "nonparametric", TD, PD or FD are based on non-parametric estimators. If func is partial match of "completeness", PD or FD estimates are based on the completeness of TD (requires a tree to be used). If not specified, default is "nonparametric".
target	True diversity value to calculate the accuracy of curves (scaled mean squared error). If not specified do not calculate accuracy (default), -1 uses the total observed diversity as true diversity and any other value is the true known diversity.
runs	Number of random permutations to be made to the sampling order. If not specified, default is 100.
prog	Present a text progress bar in the R console.

Details

Observed diversity often is an underestimation of true diversity. Several approaches have been devised to estimate species richness (TD) from incomplete sampling. These include: (1) fitting asymptotic functions to randomised accumulation curves (Soberon & Llorente 1993; Flather 1996; Cardoso et al. in prep.) (2) the use of non-parametric estimators based on the incidence or abundance of rare species (Heltshe & Forrester 1983; Chao 1984, 1987; Colwell & Coddington 1994). A correction to non-parametric estimators has also been recently proposed, based on the proportion of singleton or unique species (species represented by a single individual or in a single sampling unit respectively; Lopez et al. 2012). Cardoso et al. (2014) have proposed a way of adapting these approaches to estimate PD and FD, also adding a third possible approach for these dimensions of diversity: (3) correct PD and FD values based on the completeness of TD, where completeness equals the proportion of estimated true diversity that was observed. Calculations of PD and FD are based on Faith (1992) and Petchey & Gaston (2002, 2006), which measure PD and FD of a community as the total branch length of a tree linking all species represented in such community. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in comm must be the same as in tree.

Value

A matrix of sampling units x diversity values (sampling units, individuals, observed and estimated diversity). The values provided by this function are:

Sampl - Number of sampling units;

Ind - Number of individuals;

Obs - Observed diversity;

S1 - Singletons;

S2 - Doubletons;

Q1 - Uniques;

Q2 - Duplicates;

Jack1ab - First order jackknife estimator for abundance data;

Jack1in - First order jackknife estimator for incidence data;

Jack2ab - Second order jackknife estimator for abundance data;

Jack2in - Second order jackknife estimator for incidence data;

Chao1 - Chao estimator for abundance data;

Chao2 - Chao estimator for incidence data;

Clench - Clench or Michaelis-Menten curve;

Exponential - Exponential curve;

Rational - Rational function;

Weibull - Weibull curve;

The P-corrected version of all non-parametric estimators is also provided.

Accuracy - if accuracy is to be calculated a list is returned instead, with the second element being the scaled mean squared error of each estimator.

References

- Cardoso, P., Rigal, F., Borges, P.A.V. & Carvalho, J.C. (2014) A new frontier in biodiversity inventory: a proposal for estimators of phylogenetic and functional diversity. *Methods in Ecology and Evolution*, in press.
- Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, 11, 265-270.
- Chao, A. (1987) Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43, 783-791.
- Colwell, R.K. & Coddington, J.A. (1994) Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. Roy. Soc. London B* 345, 101-118.
- Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1-10.
- Flather, C. (1996) Fitting species-accumulation functions and assessing regional land use impacts on avian diversity. *Journal of Biogeography*, 23, 155-168.
- Heltshe, J. & Forrester, N.E. (1983) Estimating species richness using the jackknife procedure. *Biometrics*, 39, 1-11.
- Lopez, L.C.S., Fracasso, M.P.A., Mesquita, D.O., Palma, A.R.T. & Riul, P. (2012) The relationship between percentage of singletons and sampling effort: a new approach to reduce the bias of richness estimates. *Ecological Indicators*, 14, 164-169.
- Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, 5, 402-411.
- Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters*, 9, 741-758.
- Soberon, M.J. & Llorente, J. (1993) The use of species accumulation functions for the prediction of species richness. *Conservation Biology*, 7, 480-488.

Examples

```
comm <- matrix(c(1,1,0,0,0,0,2,1,0,0,0,0,2,1,0,0,0,0,2,1), nrow = 4, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5)), method="euclidean"), method="average")
alpha.accum(comm)
alpha.accum(comm, func = "nonparametric")
alpha.accum(comm, tree, "completeness")
alpha.accum(comm, tree, "curve", runs = 1000)
alpha.accum(comm, target = -1)
```

alpha.estimate

Alpha diversity estimates.

Description

Estimation of alpha diversity of multiple sites simultaneously.

Usage

```
alpha.estimate(comm, tree, func = "nonparametric")
```

Arguments

comm	A sites x species matrix, with either abundances or number of incidences.
tree	An hclust or phylo object (used only for Phylogenetic (PD) or Functional (FD) Diversity, not for Taxon Diversity (TD)).
func	The class of estimators to be used: If func is partial match of "nonparametric", TD, PD or FD are based on non-parametric estimators. If func is partial match of "completeness", PD or FD estimates are based on the completeness of TD (requires a tree to be used). If not specified, default is "nonparametric".

Details

Observed diversity often is an underestimation of true diversity. Non-parametric estimators based on the incidence or abundance of rare species have been proposed to overcome the problem of undersampling (Heltsh & Forrester 1983; Chao 1984, 1987; Colwell & Coddington 1994). A correction to non-parametric estimators has also been recently proposed, based on the proportion (P) of singleton or unique species (species represented by a single individual or in a single sampling unit respectively; Lopez et al. 2012). Cardoso et al. (2014) have proposed a way of adapting non-parametric species richness estimators to PD and FD. They have also proposed correcting PD and FD values based on the completeness of TD, where completeness equals the proportion of estimated true diversity that was observed. Calculations of PD and FD are based on Faith (1992) and Petchey & Gaston (2002, 2006), which measure PD and FD of a community as the total branch length of a tree linking all species represented in such community. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in comm must be the same as in tree.

Value

A matrix of sites x diversity values (individuals, observed and estimated diversity). The values provided by this function are:

Ind - Number of individuals;

Obs - Observed diversity;

S1 - Singletons;

S2 - Doubletons;

Jack1ab - First order jackknife estimator for abundance data;

Jack2ab - Second order jackknife estimator for abundance data;

Chao1 - Chao estimator for abundance data.

The P-corrected version of all estimators is also provided.

References

- Cardoso, P., Rigal, F., Borges, P.A.V. & Carvalho, J.C. (2014) A new frontier in biodiversity inventory: a proposal for estimators of phylogenetic and functional diversity. *Methods in Ecology and Evolution*, in press.
- Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, 11, 265-270.
- Chao, A. (1987) Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43, 783-791.
- Colwell, R.K. & Coddington, J.A. (1994) Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. Roy. Soc. London B* 345, 101-118.
- Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1-10.
- Heltshe, J. & Forrester, N.E. (1983) Estimating species richness using the jackknife procedure. *Biometrics*, 39, 1-11.
- Lopez, L.C.S., Fracasso, M.P.A., Mesquita, D.O., Palma, A.R.T. & Riul, P. (2012) The relationship between percentage of singletons and sampling effort: a new approach to reduce the bias of richness estimates. *Ecological Indicators*, 14, 164-169.
- Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, 5, 402-411.
- Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters*, 9, 741-758.

Examples

```
comm <- matrix(c(1,1,0,0,0,0,2,1,0,0,0,0,2,1,0,0,0,0,2,1), nrow = 4, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
alpha.estimate(comm)
alpha.estimate(comm, tree)
alpha.estimate(comm, tree, func = "completeness")
```

arrabida

Sample data of spiders in Arrabida (Portugal)

Description

A dataset containing the abundance of 338 spider species in each of 320 sampling units. Details are described in: Cardoso, P., Gaspar, C., Pereira, L.C., Silva, I., Henriques, S.S., Silva, R.R. & Sousa, P. (2008) Assessing spider species richness and composition in Mediterranean cork oak forests. *Acta Oecologica*, 33: 114-127.

Usage

```
data(arrabida)
```

Format

A data frame with 320 sampling units (rows) and 338 species (variables).

beta	<i>Beta diversity (Taxon, Phylogenetic or Functional Diversity - TD, PD, FD).</i>
------	---

Description

Beta diversity with possible rarefaction, multiple sites simultaneously.

Usage

```
beta(comm, tree, func = "jaccard", abund = TRUE, raref = 0, runs = 100)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data.
tree	An hclust or phylo object (used only for PD or FD).
func	Partial match indicating whether the Jaccard or Soerensen family of beta diversity measures should be used. If not specified, default is Jaccard.
abund	A boolean (T/F) indicating whether abundance data should be used or converted to incidence before analysis.
raref	An integer specifying the number of individuals for rarefaction (individual based). If raref < 1 no rarefaction is made. If raref = 1 rarefaction is made by the minimum abundance among all sites. If raref > 1 rarefaction is made by the abundance indicated. If not specified, default is 0.
runs	Number of resampling runs for rarefaction. If not specified, default is 100.

Details

The beta diversity measures used here follow the partitioning framework independently developed by Podani & Schmera (2011) and Carvalho et al. (2012) and later expanded to PD and FD by Cardoso et al. (2014), where $B_{total} = B_{repl} + B_{rich}$. B_{total} = total beta diversity, reflecting both species replacement and loss/gain; B_{repl} = beta diversity explained by replacement of species alone; B_{rich} = beta diversity explained by species loss/gain (richness differences) alone. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in comm must be the same as in tree. The rarefaction option is useful to compare communities with much different numbers of individuals sampled, which might bias diversity comparisons (Gotelli & Colwell 2001)

Value

Three distance matrices between sites, one per each of the three beta diversity measures (either "Obs" OR "Median, Min, LowerCL, UpperCL and Max").

References

- Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *Journal of Biogeography*, 41, 749-761.
- Carvalho, J.C., Cardoso, P. & Gomes, P. (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. *Global Ecology and Biogeography*, 21, 760-771.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4, 379-391.
- Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data. *Oikos*, 120, 1625-1638.

Examples

```
comm <- matrix(c(2,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,0,2,2), nrow = 4, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
beta(comm)
beta(comm, func = "Soerensen")
beta(comm, tree)
beta(comm, raref = 1)
beta(comm, tree, "s", abund = FALSE, raref = 2)
```

beta.accum

Beta diversity accumulation curves.

Description

Beta diversity between two sites with accumulation of sampling units.

Usage

```
beta.accum(
  comm1,
  comm2,
  tree,
  func = "jaccard",
  abund = TRUE,
  runs = 100,
  prog = TRUE
)
```

Arguments

comm1 A sampling units x species matrix for the first site, with either abundance or incidence data.

comm2	A sampling units x species matrix for the second site, with either abundance or incidence data.
tree	An hclust or phylo object (used only for Phylogenetic (PD) or Functional (FD) Diversity, not for Taxon Diversity (TD)).
func	Partial match indicating whether the Jaccard or Soerensen family of beta diversity measures should be used. If not specified, default is jaccard.
abund	A boolean (T/F) indicating whether abundance data should be used or converted to incidence before analysis.
runs	Number of random permutations to be made to the sampling order. If not specified, default is 100.
prog	Present a text progress bar in the R console.

Details

As widely recognized for species richness, beta diversity is also biased when communities are undersampled. Beta diversity accumulation curves have been proposed by Cardoso et al. (2009) to test if beta diversity has approached an asymptote when comparing two undersampled sites. The beta diversity measures used here follow the partitioning framework independently developed by Podani & Schmera (2011) and Carvalho et al. (2012) and later expanded to PD and FD by Cardoso et al. (2014), where $B_{total} = B_{repl} + B_{rich}$. B_{total} = total beta diversity, reflecting both species replacement and loss/gain; B_{repl} = beta diversity explained by replacement of species alone; B_{rich} = beta diversity explained by species loss/gain (richness differences) alone. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in comm1 and comm2 must be the same as in tree. Also, the number of sampling units should be similar in both sites.

Value

Three matrices of sampling units x diversity values, one per each of the three beta diversity measures (sampling units, individuals and observed diversity).

References

- Cardoso, P., Borges, P.A.V. & Veech, J.A. (2009) Testing the performance of beta diversity measures based on incidence data: the robustness to undersampling. *Diversity and Distributions*, 15, 1081-1090.
- Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *Journal of Biogeography*, 41, 749-761.
- Carvalho, J.C., Cardoso, P. & Gomes, P. (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. *Global Ecology and Biogeography*, 21, 760-771.
- Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data. *Oikos*, 120, 1625-1638.

Examples

```

comm1 <- matrix(c(2,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,2,2), nrow = 4, byrow = TRUE)
comm2 <- matrix(c(1,1,0,0,0,0,2,1,0,0,0,0,2,1,0,0,0,0,2,1), nrow = 4, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
beta accum(comm1, comm2)
beta accum(comm1, comm2, func = "Soerensen")
beta accum(comm1, comm2, tree)
beta accum(comm1, comm2, abund = FALSE)
beta accum(comm1, comm2, tree, , FALSE)

```

beta.multi

*Beta diversity among multiple communities.***Description**

Beta diversity with possible rarefaction - multiple sites measure calculated as the average or variance of all pairwise values.

Usage

```
beta.multi(comm, tree, func = "jaccard", abund = TRUE, raref = 0, runs = 100)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data.
tree	An hclust or phylo object (used only for Phylogenetic (PD) or Functional (FD) Diversity, not for Taxon Diversity (TD)).
func	Indicates whether the Jaccard or Soerensen family of beta diversity measures should be used. If not specified, default is jaccard.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis.
raref	An integer specifying the number of individuals for rarefaction (individual based). If raref < 1 no rarefaction is made. If raref = 1 rarefaction is made by the minimum abundance among all sites. If raref > 1 rarefaction is made by the abundance indicated. If not specified, default is 0.
runs	Number of resampling runs for rarefaction. If not specified, default is 100.

Details

Beta diversity of multiple sites simultaneously is calculated as either the average or the variance among all pairwise comparisons (Legendre, 2014). The beta diversity measures used here follow the partitioning framework independently developed by Podani & Schmera (2011) and Carvalho et al. (2012) and later expanded to PD and FD by Cardoso et al. (2014), where $B_{total} = B_{repl} + B_{rich}$. B_{total} = total beta diversity, reflecting both species replacement and loss/gain; B_{repl} = beta diversity explained by replacement of species alone; B_{rich} = beta diversity explained by species loss/gain (richness differences) alone. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in comm must be the same as in tree.

Value

A matrix of beta measures x diversity values (average and variance).

References

Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *Journal of Biogeography*, 41, 749-761.

Carvalho, J.C., Cardoso, P. & Gomes, P. (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. *Global Ecology and Biogeography*, 21, 760-771.

Legendre, P. (2014) Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography*, in press.

Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data. *Oikos*, 120, 1625-1638.

Examples

```
comm <- matrix(c(2,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,2,2), nrow = 4, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
beta.multi(comm)
beta.multi(comm, func = "Soerensen")
beta.multi(comm, tree)
beta.multi(comm, raref = 1)
beta.multi(comm, tree, "s", FALSE, raref = 2)
```

contribution	<i>Contribution of species or individuals to total phylogenetic/functional diversity.</i>
--------------	---

Description

Contribution of each species or individual to the total PD or FD of a number of communities.

Usage

```
contribution(comm, tree, abund = TRUE, relative = TRUE)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data. If missing, the contribution of all species to the full tree is calculated.
tree	An hclust or phylo object representing a phylogenetic or functional tree.
abund	A boolean (T/F) indicating whether contribution should be weighted by abundance of each species.
relative	A boolean (T/F) indicating whether contribution should be relative to total PD or FD (proportional contribution per individual or species). If FALSE, the sum of contributions for each site is equal to total PD/FD, if TRUE it is 1.

Details

Contribution is equivalent to the evolutionary distinctiveness index (ED) of Isaac et al. (2007) if done by species and to the abundance weighted evolutionary distinctiveness (AED) of Cadotte et al. (2010) if done by individual.

Value

A matrix of sites x species values (or values per species if no comm is given).

References

Isaac, N.J.B., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E.M. (2007) Mammals on the EDGE: conservation priorities based on threat and phylogeny. PLoS One, 2: e296.

Cadotte, M.W., Davies, T.J., Regetz, J., Kembel, S.W., Cleland, E. & Oakley, T.H. (2010) Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. Ecology Letters, 13: 96-105.

Examples

```
comm <- matrix(c(1,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,1,1,1), nrow = 4, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
contribution(tree = tree)
contribution(comm, tree)
contribution(comm, tree, FALSE)
contribution(comm, tree, abund = FALSE, relative = FALSE)
```

 dispersion

Phylogenetic/functional dispersion of species or individuals.

Description

Average dissimilarity between any two species or individuals randomly chosen in a community.

Usage

```
dispersion(
  comm,
  tree,
  distance,
  func = "originality",
  abund = TRUE,
  relative = TRUE
)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data. If missing, the dispersion using the full tree or distance matrix is calculated.
tree	An hclust or phylo object representing a phylogenetic or functional tree. One of tree or distance must be provided.
distance	A dist object representing the phylogenetic or functional distance between species.
func	Calculate dispersion using originality (default), uniqueness or contribution.
abund	A boolean (T/F) indicating whether dispersion should be calculated using individuals (T) or species (F).
relative	A boolean (T/F) indicating whether dispersion should be relative to the maximum distance between any two species in the tree or distance matrix.

Details

If abundance data is used and a tree is given, dispersion is the quadratic entropy of Rao (1982). If abundance data is not used but a tree is given, dispersion is the phylogenetic dispersion measure of Webb et al. (2002).

Value

A vector of values per site (or a single value if no comm is given).

References

Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, 21: 24-43.

Webb, C.O., Ackerly, D.D., McPeck, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33: 475-505.

Examples

```
comm <- matrix(c(1,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,1,1,1), nrow = 4, byrow = TRUE)
distance <- dist(c(1:5), method="euclidean")
tree <- hclust(distance, method="average")
dispersion(tree = tree)
dispersion(distance = distance)
dispersion(comm, tree)
dispersion(comm, tree, abund = FALSE)
dispersion(comm, tree, abund = FALSE, relative = FALSE)
```

evenness *Phylogenetic/functional evenness of species or individuals.*

Description

Regularity of distance and abundance between two species in a community.

Usage

```
evenness(
  comm,
  tree,
  distance,
  method = "expected",
  func = "camargo",
  abund = TRUE
)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data. If missing, the evenness using the full tree or distance matrix is calculated.
tree	An hclust or phylo object representing a phylogenetic or functional tree. One of tree or distance must be provided.
distance	A dist or matrix object representing the phylogenetic or functional distance between species.
method	Calculate dispersion using "expected" values (default) or values based on "contribution" of species to the tree.
func	Calculate dispersion using "Camargo" (default) or "Bulla" index.
abund	A boolean (T/F) indicating whether evenness should be calculated using abundance data.

Details

Evenness is calculated based on the index of Camargo (1993) or Bulla (1994) using the values of both edge lengths in the tree and their abundance.

If no tree or distance is provided the result is the original index.

Value

A vector of values per site (or a single value if no comm is given).

References

Bulla, L. (1994) An index of evenness and its associated diversity measure. *Oikos*, 70: 167-171.
 Camargo, J.A. (1993) Must dominance increase with the number of subordinate species in competitive interactions? *Journal of Theoretical Biology*, 161: 537-542.

Examples

```

comm <- matrix(c(1,2,0,0,0,1,1,0,0,0,0,2,2,0,0,1,1,1,1,100), nrow = 4, byrow = TRUE)
distance <- dist(c(1:5), method = "euclidean")
tree <- hclust(distance, method = "average")
evenness(comm)
evenness(tree = tree, func = "bulla")
evenness(comm, tree)
evenness(comm, tree, method = "contribution")
evenness(comm, tree, abund = FALSE)

```

 functree

Functional tree for 338 species of spiders

Description

A dataset representing the functional tree for 338 species of spiders captured in Portugal. For each species were recorded: average size, type of web, type of hunting, stenophagy, vertical stratification in vegetation and circadian activity. Details are described in: Cardoso, P., Pekar, S., Jocque, R. & Coddington, J.A. (2011) Global patterns of guild composition and functional diversity of spiders. PLoS One, 6: e21710.

Usage

```
data(functree)
```

Format

An hclust object with 338 species.

 gdm

General dynamic model of oceanic island biogeography (GDM).

Description

Fits and compares several of the most supported models for the GDM (using TD, PD or FD).

Usage

```
gdm(comm, tree, area, time)
```

Arguments

comm	Either a vector with the diversity values per island, or an island x species matrix.
tree	An hclust or phylo object (used only to fit the PD or FD GDM, requires comm to be a sites x species matrix).
area	A vector with the area of islands.
time	A vector with the age of islands. If not given, the species-area relationship is returned instead.

Details

The general dynamic model of oceanic island biogeography was proposed to account for diversity patterns within and across oceanic archipelagos as a function of area and age of the islands (Whittaker et al. 2008). Several different equations have been found to describe the GDM, extending the different SAR models with the addition of a polynomial term using island age and its square (TT2), depicting the island ontogeny. The first to be proposed was an extension of the exponential model (Whittaker et al. 2008), the power model extensions following shortly after (Fattorini 2009; Steinbauer et al. 2013), as was the linear model (Cardoso et al. *subm.*). The relationships for PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric).

Value

A matrix with the different model parameters and explanatory power.

References

- Cardoso, P., Borges, P.A.V., Carvalho, J.C., Rigal, F., Gabriel, R., Cascalho, J. & Correia, L. (*subm.*) Automated discovery of relationships, models and principles in ecology. Pre-print available from bioRxiv doi: <http://dx.doi.org/10.1101/027839>
- Fattorini, S. (2009) On the general dynamic model of oceanic island biogeography. *Journal of Biogeography*, 36: 1100-1110.
- Steinbauer, M.J, Klara, D., Field, R., Reineking, B. & Beierkuhnlein, C. (2013) Re-evaluating the general dynamic theory of oceanic island biogeography. *Frontiers of Biogeography*, 5: 185-194.
- Whittaker, R.J., Triantis, K.A. & Ladle, R.J. (2008) A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, 35: 977-994.

Examples

```
div <- c(1,3,5,8,10)
comm <- matrix(c(2,0,0,0,3,1,0,0,2,4,5,0,1,3,2,5,1,1,1,1), nrow = 5, ncol = 4, byrow = TRUE)
tree <- hclust(dist(c(1:4), method="euclidean"), method="average")
area <- c(10,40,80,160,160)
time <- c(1,2,3,4,5)
gdm(div,,area,time)
gdm(comm,tree,area,time)
gdm(div,,area)
```

geres

Sample data of spiders in Geres (Portugal)

Description

A dataset containing the abundance of 338 spider species in each of 320 sampling units. Details are described in: Cardoso, P., Scharff, N., Gaspar, C., Henriques, S.S., Carvalho, R., Castro, P.H., Schmidt, J.B., Silva, I., Szuts, T., Castro, A. & Crespo, L.C. (2008) Rapid biodiversity assessment of spiders (Araneae) using semi-quantitative sampling: a case study in a Mediterranean forest. *Insect Conservation and Diversity*, 1: 71-84.

Usage

```
data(geres)
```

Format

A data frame with 320 sampling units (rows) and 338 species (variables).

guadiana	<i>Sample data of spiders in Guadiana (Portugal)</i>
----------	--

Description

A dataset containing the abundance of 338 spider species in each of 320 sampling units. Details are described in: Cardoso, P., Henriques, S.S., Gaspar, C., Crespo, L.C., Carvalho, R., Schmidt, J.B., Sousa, P. & Szuts, T. (2009) Species richness and composition assessment of spiders in a Mediterranean scrubland. *Journal of Insect Conservation*, 13: 45-55.

Usage

```
data(guadiana)
```

Format

A data frame with 192 sampling units (rows) and 338 species (variables).

iaor	<i>Interspecific abundance-occupancy relationship (IAOR).</i>
------	---

Description

Fits and compares several of the most supported models for the IAOR.

Usage

```
iaor(comm)
```

Arguments

comm A sites x species matrix with abundance values.

Details

Locally abundant species tend to be widespread while locally rare species tend to be narrowly distributed. That is, for a given species assemblage, there is a positive interspecific abundance-occupancy relationship (Brown 1984). This function compares some of the most commonly used and theoretically or empirically supported models (Nachman 1981; He & Gaston 2000; Cardoso et al. *subm.*).

Value

A matrix with the different model parameters and explanatory power.

References

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *American Naturalist*, 124: 255-279.

Cardoso, P., Borges, P.A.V., Carvalho, J.C., Rigal, F., Gabriel, R., Mammola, S., Cascalho, J. & Correia, L. (subm.) Automated discovery of relationships, models and principles in ecology. Preprint available from bioRxiv doi: <http://dx.doi.org/10.1101/027839>

He, F.L. & Gaston, K.J. (2000) Estimating species abundance from occurrence. *American Naturalist*, 156: 553-559.

Nachman, G. (1981) A mathematical model of the functional relationship between density and spatial distribution of a population. *Journal of Animal Ecology*, 50: 453-460.

Examples

```
comm <- matrix(c(4,3,2,1,5,4,3,2,3,2,1,0,6,3,0,0,0,0,0), nrow = 5, ncol = 4, byrow = TRUE)
iaor(comm)
```

kernel.alpha

Alpha diversity using kernel density hypervolumes.

Description

Estimation of functional richness of one or multiple sites, based on n-dimensional hypervolumes.

Usage

```
kernel.alpha(
  comm,
  trait,
  method = "gaussian",
  abund = FALSE,
  return.hv = FALSE,
  ...
)
```

Arguments

comm	A 'Hypervolume' object or a 'HypervolumesList' object (one for each species or community) constructed with the hypervolume R package. Alternatively, a sites x species matrix, with incidence or abundance data about the species in the community.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.

method	Method for constructing the 'Hypervolume' object. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume R package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a 'Hypervolume' or 'HypervolumeList' is provided as input data.
return.hv	Boolean indicating whether the 'Hypervolume' objects used to calculate alpha diversity should be returned (default is FALSE).
...	further arguments to be passed for other methods in hypervolume package.

Details

Estimates the functional richness (alpha FD) of one or more communities using kernel density hypervolumes, as implemented in Blonder et al. (2014, 2018). Functional richness is expressed as the total volume of the n-dimensional hypervolume (Mammola & Cardoso, 2020), as returned by the function `hypervolume::get_volume`. Note that the hypervolume is dimensionless, and that only hypervolumes with the same number of dimensions can be compared in terms of functional richness. Given that the density and positions of stochastic points in the hypervolume are probabilistic, the functional richness of the trait space will intimately depend on the quality of input hypervolumes (details in Mammola & Cardoso, 2020). If abundance data of species in the community are provided as input data (`abund = TRUE`), each species trait is weighted by replicating it by the abundance in the estimation of the hypervolume.

Value

A vector of alpha diversity values for each site. If `return.hv` is set to `TRUE`, the function also returns the list of hypervolumes used to compute alpha diversity.

References

- Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23: 595-609.
- Blonder, B., Morrow, C.B., Maitner, B., Harris, D.J., Lamanna, C., Violle, C., ... & Kerkhoff, A.J. (2018) New approaches for delineating n-dimensional hypervolumes. *Methods in Ecology and Evolution*, 9: 305-319.
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>

Examples

```
comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
rownames(comm) <- c("Community_1", "Community_2", "Community_3")
colnames(comm) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")

trait <- cbind(c(2.2,4.4,6.1,8.3),c(0.5,1,0.5,0.4),c(0.7,1.2,0.5,0.4))
```

```

rownames(trait) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")
colnames(trait) <- c("Trait_1", "Trait_2", "Trait_3")

#Example with community and trait matrices as input data
#kernel.alpha(comm = comm, trait = trait, method = "box", return.hv = FALSE)

#Example with community and trait matrices as input data and abundance data
#kernel.alpha(comm = comm, trait = trait, method = "box", abund = TRUE, return.hv = FALSE)

#Example with hypervolume as input data
#kernel.alpha(comm = hypervolume_box(trait[comm[1,]==1,], name="Community_1"))

#Example with hypervolumeList as input data
#hv1 <- hypervolume_box(trait[comm[1,]==1,], name="Community_1")
#hv2 <- hypervolume_box(trait[comm[2,]==1,], name="Community_2")
#hv3 <- hypervolume_box(trait[comm[3,]==1,], name="Community_3")
#hvlist <- hypervolume_join(hv1, hv2, hv3)
#kernel.alpha(hvlist)

```

kernel.beta

Beta diversity partitioning using kernel density hypervolumes.

Description

Pairwise beta diversity partitioning into replacement and net difference in amplitude components of n-dimensional hypervolumes.

Usage

```

kernel.beta(
  comm,
  trait,
  method = "gaussian",
  func = "jaccard",
  abund = FALSE,
  return.hv = FALSE,
  ...
)

```

Arguments

comm	A 'HypervolumeList' object (one 'Hypervolume' object for each species or community) constructed with the hypervolume R package. Alternatively, a sites x species matrix, with incidence or abundance data about the species in the community.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.

method	Method for constructing the 'Hypervolume' objects. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume R package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.
func	Partial match indicating whether the Jaccard or Soerensen family of beta diversity measures should be used. If not specified, default is Jaccard.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a HypervolumeList is provided as input data.
return.hv	Boolean indicating whether the hypervolume objects used to calculate beta diversity should be returned (default is FALSE).
...	further arguments to be passed for other methods in hypervolume package.

Details

Computes a pairwise decomposition of the overall differentiation among kernel hypervolumes into two components: the replacement (shifts) of space between hypervolumes and net differences between the amount of space enclosed by each hypervolume. The beta diversity measures used here follow the FD partitioning framework developed by Carvalho & Cardoso (2018), where $B_{total} = B_{replacement} + B_{richness}$. Beta diversity ranges from 0 (when hypervolumes are identical) to 1 (when hypervolumes are fully dissimilar). See Carvalho & Cardoso (2018) and Mammola & Cardoso (2020) for the full formulas of beta diversity used here. If abundance data of species in the community are provided as input data (`abund = TRUE`), each species trait is weighted by replicating it by the abundance in the estimation of the hypervolume.

Value

Three pairwise distance matrices, one per each of the three beta diversity components. If `return.hv` is set to `TRUE`, the function also returns the list of hypervolumes used to compute the distance matrices.

References

- Carvalho, J.C. & Cardoso, P. (2018) Decomposing the causes for niche differentiation between species using hypervolumes. *bioRxiv*, <https://doi.org/10.1101/485920>
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>

Examples

```
comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
rownames(comm) <- c("Community_1", "Community_2", "Community_3")
colnames(comm) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")

trait <- cbind(c(2.2,4.4,6.1,8.3),c(0.5,1,0.5,0.4),c(0.7,1.2,0.5,0.4))
rownames(trait) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")
colnames(trait) <- c("Trait_1", "Trait_2", "Trait_3")
```



```

#Example with community and trait matrices as input data:
#kernel.beta(comm = comm, trait = trait, return.hv = TRUE)

#Example with community and trait matrices as input data and abundance data
#kernel.beta(comm = comm, trait = trait, abund = TRUE, return.hv = FALSE)

#Example with hypervolumeList as input data:
#hv1 <- hypervolume_box(trait[comm[1,]==1,],name="Community_1")
#hv2 <- hypervolume_box(trait[comm[2,]==1,],name="Community_2")
#hv3 <- hypervolume_box(trait[comm[3,]==1,],name="Community_3")
#hvlist = hypervolume_join(hv1, hv2, hv3)
#kernel.beta(hvlist)

```

kernel.contribution	<i>Contribution of each observation (species or individuals) to the n-dimensional hypervolume representing a given species or community.</i>
---------------------	--

Description

Contribution of each species or individual to the total volume of one or more kernel hypervolumes.

Usage

```
kernel.contribution(comm, trait, method = "gaussian", abund = FALSE, ...)
```

Arguments

comm	A 'Hypervolume' object constructed with the hypervolume R package. Alternatively, a sites x species matrix, with incidence or abundance data about the species in the community. Note that the use of 'HypervolumeList' object is not implemented for this function yet.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.
method	Method for constructing the 'Hypervolume' object. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume R package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a Hypervolume is provided as input data.
...	further arguments to be passed for other methods in hypervolume package.

Details

The contribution of each observation (species or individual) to the total volume of a kernel hypervolume, calculated as the difference in volume between the total hypervolume and a second hypervolume lacking this specific observation (i.e., leave-one-out approach; Mammola & Cardoso, 2020). Contribution is a measure of functional rarity (sensu Violle et al., 2017; Carmona et al., 2017) that allows to map the contribution of each observation to the richness components of FD (Mammola & Cardoso, 2020). Note that the contribution of a species or individual can be negative, if the removal of an observation increases the total volume (see Figure 2d in Mammola & Cardoso 2020). This might happen, although not always, in cases when the presence of a given species decreases the average distance between all the species in the community, i.e., when a given species is close to the "average" species of that community, making that community less diverse in some sense (Mammola & Cardoso, 2020). By definition, this does not happen in the case of functional dendrograms (BAT::contribution). For large sample sizes, computation time can be high (use method = 'box' for a quicker estimation). If abundance data are provided (abund = TRUE), the contribution of each observation is divided by its abundance value, thus representing the contribution of each individual.

Value

A matrix with the contribution values of each species or individual for each site.

References

- Carmona, C.P., de Bello, F., Sasaki, T., Uchida, K. & Partel, M. (2017) Towards a common toolbox for rarity: A response to Violle et al. *Trends in Ecology and Evolution*, 32(12): 889-891.
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>
- Violle, C., Thuiller, W., Mouquet, N., Munoz, F., Kraft, N.J.B., Cadotte, M.W., ... & Mouillot, D. (2017) Functional rarity: The ecology of outliers. *Trends in Ecology and Evolution*, 32: 356-367.

Examples

```
comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
rownames(comm) <- c("Community_1", "Community_2", "Community_3")
colnames(comm) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")

trait <- cbind(c(2.2,4.4,6.1,8.3), c(0.5,1,0.5,0.4), c(0.7,1.2,0.5,0.4))
rownames(trait) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")
colnames(trait) <- c("Trait_1", "Trait_2", "Trait_3")

#Example with community and trait matrices as input data
#kernel.contribution(comm = comm, trait = trait, method = "gaussian")

#Example with hypervolume as input data
#kernel.contribution(hypervolume_box(trait))
```

kernel.dispersion	<i>Functional dispersion of a n-dimensional hypervolume representing a given community.</i>
-------------------	---

Description

Average distance to centroid or dissimilarity between random points within the boundaries of the kernel density hypervolume.

Usage

```
kernel.dispersion(
  comm,
  trait,
  method = "gaussian",
  func = "divergence",
  abund = FALSE,
  frac = 0.1,
  ...
)
```

Arguments

comm	A 'Hypervolume' object or a 'HypervolumeList' object (one for each species or community) constructed with the hypervolume R package. Alternatively, a sites x species matrix, with incidence or abundance data about the species in the community.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.
method	Method for constructing the 'Hypervolume'. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume R package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.
func	Function for calculating dispersion. One of 'divergence' or 'dissimilarity'.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a Hypervolume or HypervolumeList is provided as input data.
frac	A value between 0.01 and 1, indicating the fraction of random points to be used. Default is 0.1.
...	further arguments to be passed for other methods in hypervolume package.

Details

This function calculates either the average distance between stochastic points within the kernel density hypervolume and the centroid of these points (divergence; Laliberte & Legendre, 2010; see also Carmona et al., 2019) or the average distance between all points (dissimilarity, see also function `BAT::dispersion`). The number of stochastic points is controlled by the 'frac' parameter (increase this number for less deviation in the estimation). If abundance data of species in the community are provided as input data (`abund = TRUE`), each species trait is weighted by replicating it by the abundance in the estimation of the hypervolume.

Value

A vector of dispersion values for each site.

References

- Carmona, C.P., de Bello, F., Mason, N.W.H. & Leps, J. (2019) Trait probability density (TPD): measuring functional diversity across scales based on TPD with R. *Ecology* 100: e02876.
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>
- Laliberte, E. & Legendre, P. (2010) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* 91: 299-305.

Examples

```
comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
rownames(comm) = c("Community_1", "Community_2", "Community_3")
colnames(comm) = c("Sp_1", "Sp_2", "Sp_3", "Sp_4")

trait <- cbind(c(2.2,4.4,6.1,8.3), c(0.5,1,0.5,0.4), c(0.7,1.2,0.5,0.4))
rownames(trait) = c("Sp_1", "Sp_2", "Sp_3", "Sp_4")
colnames(trait) = c("Trait_1", "Trait_2", "Trait_3")

#Example with community and trait matrices as input data
#kernel.dispersion(comm = comm, trait = trait)

#Example with hypervolume as input data and the dissimilarity method
#kernel.dispersion(hypervolume_gaussian(trait), func = 'dissimilarity')
```

kernel.evenness	<i>Functional evenness of a n-dimensional hypervolume representing a given community.</i>
-----------------	---

Description

Functional evenness of a community, measuring the regularity of stochastic points distribution within the total functional space.

Usage

```
kernel.evenness(comm, trait, method = "gaussian", abund = FALSE, ...)
```

Arguments

comm	A 'Hypervolume' object or a 'HypervolumeList' object (one for each species or community) constructed with the hypervolume R package. Alternatively, a sites x species matrix, with incidence or abundance data about the species in the community.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.
method	Method for constructing the 'Hypervolume' object. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume R package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a Hypervolume or HypervolumeList is provided as input data.
...	further arguments to be passed for other methods in hypervolume package.

Details

This function measures the functional evenness (Mason et al., 2005) of a n-dimensional hypervolume, namely the regularity of stochastic points distribution within the total trait space (Mammola & Cardoso, 2020). Evenness is calculated as the overlap between the observed hypervolume and a theoretical hypervolume where traits and abundances are evenly distributed within the range of their values (Carmona et al., 2016, 2019). If abundance data of species in the community are provided as input data (abund = TRUE), each species trait is weighted by replicating it by the abundance in the estimation of the hypervolume.

Value

A vector of evenness values for each site.

References

- Carmona, C.P., de Bello, F., Mason, N.W.H. & Leps, J. (2016) Traits without borders: integrating functional diversity across scales. *Trends in Ecology and Evolution* 31: 382-394.
- Carmona, C.P., de Bello, F., Mason, N.W.H. & Leps, J. (2019) Trait probability density (TPD): measuring functional diversity across scales based on TPD with R. *Ecology* 100: e02876.
- Mason, N.W.H., Mouillot, D., Lee, W.G. & Wilson, J.B. (2005) Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos* 111: 112-118.
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>

Examples

```

comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
rownames(comm) <- c("Community_1","Community_2","Community_3")
colnames(comm) <- c("Sp_1","Sp_2","Sp_3","Sp_4")
trait <- cbind(c(2.2,4.4,6.1,8.3),c(0.5,1,0.5,0.4),c(0.7,1.2,0.5,0.4))
rownames(trait) <- c("Sp_1","Sp_2","Sp_3","Sp_4")
colnames(trait) <- c("Trait_1","Trait_2","Trait_3")

#Example with community and trait matrices as input data
#kernel.evenness(comm = comm, trait = trait)

#Example with hypervolume as input data
#kernel.evenness(hypervolume_gaussian(trait))

```

kernel.originality	<i>Functional originality of observations (species or individuals) in a n-dimensional hypervolume representing a given species or community.</i>
--------------------	--

Description

Average dissimilarity between a species or individual and a sample of random points within the boundaries of the n-dimensional hypervolume.

Usage

```

kernel.originality(
  comm,
  trait,
  method = "gaussian",
  abund = FALSE,
  frac = 0.1,
  relative = FALSE,
  ...
)

```

Arguments

comm	A 'Hypervolume' object constructed with the hypervolume R package or a sites x species matrix, with incidence or abundance data about the species in the community. Note that the use of 'HypervolumeList' object is not implemented for this function yet.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.
method	Method for constructing the 'Hypervolume' object. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume R package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.

abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a Hypervolume is provided as input data.
frac	A value between 0.01 and 1, indicating the fraction of random points to be used in the estimation of originality. Default is 0.1.
relative	A boolean (T/F) indicating whether originality should be relative to the most original species.
...	further arguments to be passed for other methods in hypervolume package.

Details

A measure of the originality (sensu Pavoine et al., 2005) of each observation (species or individuals) used to construct the n-dimensional hypervolume. In a probabilistic hypervolume, originality is calculated as the average distance between each observation to a sample of stochastic points within the boundaries of the n-dimensional hypervolume (Mammola & Cardoso, 2020). Originality is a measure of functional rarity (sensu Violle et al., 2017; Carmona et al., 2017) that allows to map the contribution of each observation to the divergence components of FD (Mammola & Cardoso, 2020). The number of sample points to be used in the estimation of the originality is controlled by the frac parameter. Increase frac for less deviation in the estimation, but mind that computation time also increases. For large sample sizes, computation time can be very high (use method = 'box' for a quicker estimation). If abundance data of species in the community are provided as input data (abund = TRUE), each species trait is weighted by replicating it by the abundance in the estimation of the hypervolume.

Value

A matrix with the originality values of each species or individual in each site.

References

- Carmona, C.P., de Bello, F., Sasaki, T., Uchida, K. & Partel, M. (2017) Towards a common toolbox for rarity: A response to Violle et al. *Trends in Ecology and Evolution*, 32: 889-891.
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>
- Pavoine, S., Ollier, S. & Dufour, A.-B. (2005) Is the originality of a species measurable? *Ecology Letters*, 8: 579-586.
- Violle, C., Thuiller, W., Mouquet, N., Munoz, F., Kraft, N.J.B., Cadotte, M.W., ... & Mouillot, D. (2017) Functional rarity: the ecology of outliers. *Trends in Ecology and Evolution*, 32: 356-367.

Examples

```
comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
rownames(comm) <- c("Community_1", "Community_2", "Community_3")
colnames(comm) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")

trait <- cbind(c(2.2,4.4,6.1,8.3),c(0.5,1,0.5,0.4),c(0.7,1.2,0.5,0.4))
rownames(trait) <- c("Sp_1", "Sp_2", "Sp_3", "Sp_4")
colnames(trait) <- c("Trait_1", "Trait_2", "Trait_3")
```

```
#Example with community and trait matrices as input data
#kernel.originality(comm = comm, trait = trait, method='gaussian', abund = TRUE, frac = 0.01)

#Example with hypervolume as input data
#kernel.originality(comm = hypervolume_gaussian(trait))
```

kernel.similarity *Pairwise similarity among n-dimensional hypervolumes.*

Description

Calculate pairwise distance metrics (centroid and minimum distance) and similarity indices (Intersection, Jaccard, Soerensen-Dice) among n-dimensional hypervolumes.

Usage

```
kernel.similarity(
  comm,
  trait,
  method = "gaussian",
  abund = FALSE,
  return.hv = FALSE,
  ...
)
```

Arguments

comm	A 'HypervolumeList' object (one hypervolume for each species or community) constructed with the hypervolume R package. Alternatively, a sites x species matrix, with incidence or abundance data about the species in the community.
trait	A matrix of traits for each species in comm (a species for each row and traits as columns). Must be provided only if 'comm' is a sites x species matrix.
method	Method for constructing the 'HypervolumeList' object. One of "box" (box kernel density estimation), "gaussian" (Gaussian kernel density estimation), or "svm" (one-class support vector machine). See respective functions of the hypervolume package for details. Must be provided only if 'comm' is a sites x species matrix. Default is 'gaussian'.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis. If not specified, default is FALSE. Ignored if a 'HypervolumeList' is provided as input data.
return.hv	Boolean indicating whether the hypervolume objects used to calculate pairwise similarity should be returned (default is FALSE).
...	further arguments to be passed for other methods in hypervolume package.

Details

Computes a pairwise comparison between kernel density hypervolumes of multiple species or communities, based on the distance and similarity metrics implemented in hypervolume R package (Blonder et al., 2014, 2018). See Mammola (2019) for a description of the different indices, and a comparison between their performance. Note that computation time largely depends on the number of 'Hypervolume' objects in the list, and scales almost exponentially with the number of hypervolume axes. If abundance data of species in the community are provided as input data (abund = TRUE), each species trait is weighted by replicating it by the abundance in the estimation of the hypervolume.

Value

Five pairwise distance matrices, one per each of the distance and similarity indices (in order: distance between centroids, minimum distance, Jaccard overlap, Soerensen-Dice overlap, and Intersection among hypervolumes).

References

- Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23: 595-609.
- Blonder, B., Morrow, C.B., Maitner, B., Harris, D.J., Lamanna, C., Violle, C., ... & Kerkhoff, A.J. (2018) New approaches for delineating n-dimensional hypervolumes. *Methods in Ecology and Evolution*, 9: 305-319.
- Mammola, S. (2019) Assessing similarity of n-dimensional hypervolumes: Which metric to use?. *Journal of Biogeography*, 46: 2012-2023.
- Mammola, S. & Cardoso, P. (2020) Functional diversity metrics using kernel density n-dimensional hypervolumes. *bioRxiv*, <https://doi.org/10.1101/2020.01.25.919373>

Examples

```
comm <- rbind(c(0,3,2,1), c(1,5,6,2), c(0,0,2,1))
trait <- cbind(c(2.2,4.4,6.1,8.3),c(0.5,1,0.5,0.4),c(0.7,1.2,0.5,0.4))

#example with community and trait matrices as input data:
#kernel.similarity(comm = comm, trait = trait)

#'#example with a list of hypervolume as input data:
#A = hypervolume_box(trait[,1:2], name = "Community_1")
#B = hypervolume_box(trait[,2:3], name = "Community_2")
#kernel.similarity(hypervolume_join(A,B))
```

linnean

Create Linnean tree.

Description

Creates a Linnean tree from taxonomic hierarchy.

Usage

```
linnean(taxa, distance = NULL)
```

Arguments

taxa	A taxonomic matrix with columns ordered according to linnean hierarchy starting with the highest.
distance	A vector with distances between levels starting with the highest. If not provided distances will be evenly distributed from 1 to 0.

Value

An hclust with all species.

Examples

```
family <- c("Nemesiidae", "Nemesiidae", "Zodariidae", "Zodariidae")
genus <- c("Iberesia", "Nemesia", "Zodarion", "Zodarion")
species <- c("Imachadoi", "Nungoliant", "Zatlanticum", "Zlusitanicum")
taxa <- cbind(family, genus, species)
par(mfrow = c(1, 2))
plot(linnean(taxa))
plot(linnean(taxa, c(2, 0.5, 0.3)))
```

optim.alpha

Optimization of alpha diversity sampling protocols.

Description

Optimization of alpha diversity sampling protocols when different methods and multiple samples per method are available.

Usage

```
optim.alpha(comm, tree, methods, base, runs = 1000, prog = TRUE)
```

Arguments

comm	A samples x species x sites array, with either abundance or incidence data.
tree	An hclust or phylo object (used only to optimize PD or FD sampling).
methods	A vector specifying the method of each sample (length must be equal to nrow(comm))
base	A vector defining a base protocol from which to build upon (complementarity analysis) (length must be equal to number of methods).
runs	Number of random permutations to be made to the sample order. Default is 1000.
prog	Present a text progress bar in the R console.

Details

Often a combination of methods allows sampling maximum plot diversity with minimum effort, as it allows sampling different sub-communities, contrary to using single methods. Cardoso (2009) proposed a way to optimize the number of samples per method when the target is to maximize sampled alpha diversity. It is applied here for TD, PD and FD, and for one or multiple sites simultaneously. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric).

Value

A matrix of samples x methods (values being optimum number of samples per method). The last column is the average alpha diversity value, rescaled to 0-1 if made for several sites, where 1 is the true diversity of each site.

References

Cardoso, P. (2009) Standardization and optimization of arthropod inventories - the case of Iberian spiders. *Biodiversity and Conservation*, 18, 3949-3962.

Examples

```
comm1 <- matrix(c(1,1,0,2,4,0,0,1,2,0,0,3), nrow = 4, ncol = 3, byrow = TRUE)
comm2 <- matrix(c(2,2,0,3,1,0,0,0,5,0,0,2), nrow = 4, ncol = 3, byrow = TRUE)
comm <- array(c(comm1, comm2), c(4,3,2))
colnames(comm) <- c("Sp1", "Sp2", "Sp3")
methods <- c("Met1", "Met2", "Met2", "Met3")
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
optim.alpha(comm, methods)
optim.alpha(comm, tree, methods)
optim.alpha(comm, methods = methods, base = c(0,0,1), runs = 100)
```

optim.alpha.stats *Efficiency statistics for alpha-sampling.*

Description

Average alpha diversity observed with a given number of samples per method.

Usage

```
optim.alpha.stats(comm, tree, methods, samples, runs = 1000)
```

Arguments

comm	A samples x species x sites array, with either abundance or incidence data.
tree	An hclust or phylo object (used only to optimize PD or FD sampling).
methods	A vector specifying the method of each sample (length must be equal to nrow(comm))

samples	A vector defining the number of samples per method to be evaluated (length must be equal to number of methods).
runs	Number of random permutations to be made to the sample order. Default is 1000.

Details

Different combinations of samples per method allow sampling different sub-communities. This function allows knowing the average TD, PD or FD values for a given combination, for one or multiple sites simultaneously. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric).

Value

A single average alpha diversity value. Rescaled to 0-1 if made for several sites, where 1 is the true diversity of each site.

Examples

```
comm1 <- matrix(c(1,1,0,2,4,0,0,1,2,0,0,3), nrow = 4, ncol = 3, byrow = TRUE)
comm2 <- matrix(c(2,2,0,3,1,0,0,0,5,0,0,2), nrow = 4, ncol = 3, byrow = TRUE)
comm <- array(c(comm1, comm2), c(4,3,2))
colnames(comm) <- c("Sp1", "Sp2", "Sp3")
methods <- c("Met1", "Met2", "Met2", "Met3")
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
optim.alpha.stats(comm, methods, c(1,1,1))
optim.alpha.stats(comm, tree, methods = methods, samples = c(0,0,1), runs = 100)
```

optim.beta

Optimization of beta diversity sampling protocols.

Description

Optimization of beta diversity sampling protocols when different methods and multiple samples per method are available.

Usage

```
optim.beta(comm, tree, methods, base, abund = TRUE, runs = 1000, prog = TRUE)
```

Arguments

comm	A samples x species x sites array, with either abundance or incidence data.
tree	An hclust or phylo object (used only to optimize PD or FD sampling).
methods	A vector specifying the method of each sample (length must be equal to nrow(comm))
base	Allows defining a base mandatory protocol from which to build upon (completeness analysis). It should be a vector with length = number of methods.

abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis.
runs	Number of random permutations to be made to the sample order. Default is 1000.
prog	Present a text progress bar in the R console.

Details

Often, comparing differences between sites or the same site along time (i.e. measure beta diversity) it is not necessary to sample exhaustively. A minimum combination of samples targeting different sub-communities (that may behave differently) may be enough to perceive such differences, for example, for monitoring purposes. Cardoso et al. (in prep.) introduce and differentiate the concepts of alpha-sampling and beta-sampling. While alpha-sampling optimization implies maximizing local diversity sampled (Cardoso 2009), beta-sampling optimization implies minimizing differences in beta diversity values between partially and completely sampled communities. This function uses as beta diversity measures the Btotal, Brepl and Brich partitioning framework (Carvalho et al. 2012) and respective generalizations to PD and FD (Cardoso et al. 2014). PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric).

Value

A matrix of samples x methods (values being optimum number of samples per method). The last column is the average absolute difference from real beta.

References

- Cardoso, P. (2009) Standardization and optimization of arthropod inventories - the case of Iberian spiders. *Biodiversity and Conservation*, 18, 3949-3962.
- Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *Journal of Biogeography*, 41, 749-761.
- Cardoso, P., et al. (in prep.) Optimal inventorying and monitoring of taxon, phylogenetic and functional diversity.
- Carvalho, J.C., Cardoso, P. & Gomes, P. (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. *Global Ecology and Biogeography*, 21, 760-771.

Examples

```
comm1 <- matrix(c(1,1,0,2,4,0,0,1,2,0,0,3), nrow = 4, ncol = 3, byrow = TRUE)
comm2 <- matrix(c(2,2,0,3,1,0,0,0,5,0,0,2), nrow = 4, ncol = 3, byrow = TRUE)
comm3 <- matrix(c(2,0,0,3,1,0,0,0,5,0,0,2), nrow = 4, ncol = 3, byrow = TRUE)
comm <- array(c(comm1, comm2, comm3), c(4,3,3))
colnames(comm) <- c("sp1", "sp2", "sp3")
methods <- c("Met1", "Met2", "Met2", "Met3")
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
optim.beta(comm, methods = methods, runs = 100)
optim.beta(comm, tree, methods = methods, abund = FALSE, base = c(0,0,1), runs = 100)
```

optim.beta.stats *Efficiency statistics for beta-sampling.*

Description

Average absolute difference between sampled and real beta diversity when using a given number of samples per method.

Usage

```
optim.beta.stats(comm, tree, methods, samples, abund = TRUE, runs = 1000)
```

Arguments

comm	A samples x species x sites array, with either abundance or incidence data.
tree	An hclust or phylo object (used only to optimize PD or FD sampling).
methods	A vector specifying the method of each sample (length must be equal to nrow(comm))
samples	The combination of samples per method we want to test. It should be a vector with length = number of methods.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis.
runs	Number of random permutations to be made to the sample order. Default is 1000.

Details

Different combinations of samples per method allow sampling different sub-communities. This function allows knowing the average absolute difference between sampled and real beta diversity for a given combination, for one or multiple sites simultaneously. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric).

Value

A single average absolute beta diversity difference value.

Examples

```
comm1 <- matrix(c(1,1,0,2,4,0,0,1,2,0,0,3), nrow = 4, ncol = 3, byrow = TRUE)
comm2 <- matrix(c(2,2,0,3,1,0,0,0,5,0,0,2), nrow = 4, ncol = 3, byrow = TRUE)
comm3 <- matrix(c(2,0,0,3,1,0,0,0,5,0,0,2), nrow = 4, ncol = 3, byrow = TRUE)
comm <- array(c(comm1, comm2, comm3), c(4,3,3))
colnames(comm) <- c("sp1", "sp2", "sp3")
methods <- c("Met1", "Met2", "Met2", "Met3")
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
optim.beta.stats(comm, methods, c(1,1,1))
optim.beta.stats(comm, tree, methods = methods, samples = c(0,0,1), runs = 100)
```

optim.spatial	<i>Optimization of spatial sampling.</i>
---------------	--

Description

Optimization of sampling site distribution in space based on environmental (or other) variables.

Usage

```
optim.spatial(layers, n, latlong = TRUE, clusterMap = TRUE)
```

Arguments

layers	A Raster* object (typically a multi-layer type: RasterStack or RasterBrick).
n	The number of intended sampling sites (clusters).
latlong	Boolean indicating whether latitude and longitude should be taken into account when clustering.
clusterMap	Boolean indicating whether to build a new raster with clusters.

Details

Optimizing the selection of sampling sites often requires maximizing the environmental diversity covered by them. One possible solution to this problem, here adopted, is performing a k-means clustering using environmental data and choosing the sites closest to the multidimensional environmental centroid of each cluster for sampling (Jimenez-Valverde & Lobo 2004)

Value

Either a matrix of cells x clusters (also indicating distance to centroid, longitude and latitude of each cell) or a list with such matrix plus the clusterMap.

References

Jimenez-Valverde, A., & Lobo, J. M. (2004) Un metodo sencillo para seleccionar puntos de muestreo con el objetivo de inventariar taxones hiperdiversos: el caso practico de las familias Araneidae y Thomisidae (Araneae) en la comunidad de Madrid, Espana. *Ecologia*, 18: 297-305.

originality	<i>Phylogenetic/functional originality of species or individuals.</i>
-------------	---

Description

Average dissimilarity between a species or individual and all others in a community.

Usage

```
originality(comm, tree, distance, abund = TRUE, relative = TRUE)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data. If missing, the originality using the full tree or distance matrix is calculated.
tree	An hclust or phylo object representing a phylogenetic or functional tree. One of tree or distance must be provided.
distance	A dist object representing the phylogenetic or functional distance between species.
abund	A boolean (T/F) indicating whether originality should be calculated per individual (T) or species (F).
relative	A boolean (T/F) indicating whether originality should be relative to the maximum distance between any two species in the tree or distance matrix.

Details

This is the originality measure of Pavoine et al. (2005) without replacement.

Value

A matrix of sites x species values.

References

Pavoine, S., Ollier, S. & Dufour, A.-B. (2005) Is the originality of a species measurable? Ecology Letters, 8: 579-586.

Examples

```
comm <- matrix(c(1,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,1,1,1), nrow = 4, byrow = TRUE)
distance <- dist(c(1:5), method="euclidean")
tree <- hclust(distance, method="average")
originality(tree = tree)
originality(distance = distance)
originality(comm, tree)
originality(comm, tree, abund = FALSE)
originality(comm, tree, abund = FALSE, relative = FALSE)
```


phylotree

Taxonomic tree for 338 species of spiders (surrogate for phylogeny)

Description

A dataset representing an approximation to the phylogenetic tree for 338 species of spiders captured in Portugal. The tree is based on the linnean hierarchy, with different suborders separated by 1 unit, families by 0.75, genera by 0.5 and species by 0.25.

Usage

```
data(phylotree)
```

Format

An hclust object with 338 species.

raster.alpha

Maps of alpha diversity (Taxon, Phylogenetic or Functional Diversity - TD, PD, FD).

Description

Observed alpha diversity using rasters of species distributions (presence/absence).

Usage

```
raster.alpha(layers, tree)
```

Arguments

layers	A Raster* object of species distributions (typically a multi-layer type: RasterStack or RasterBrick).
tree	An hclust or phylo object (used only for PD or FD).

Details

TD is equivalent to species richness. Calculations of PD and FD are based on Faith (1992) and Petchey & Gaston (2002, 2006), which measure PD and FD of a community as the total branch length of a tree linking all species represented in such community. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in layers must be the same as in tree.

Value

A raster object representing richness in space.

References

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1-10.

Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, 5, 402-411.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters*, 9, 741-758.

Examples

```
sp1 <- raster::raster(matrix(c(NA,1,1,1,1,0,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
sp2 <- raster::raster(matrix(c(0,0,0,0,1,1,1,1,1), nrow = 3, ncol = 3, byrow = TRUE))
sp3 <- raster::raster(matrix(c(0,0,0,0,1,1,1,0,0), nrow = 3, ncol = 3, byrow = TRUE))
spp <- raster::stack(sp1, sp2, sp3)
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
raster.alpha(spp)
raster.alpha(spp, tree)
```

raster.beta	<i>Maps of beta diversity (Taxon, Phylogenetic or Functional Diversity - TD, PD, FD).</i>
-------------	---

Description

Observed beta diversity using rasters of species distributions (presence/absence or abundance).

Usage

```
raster.beta(layers, tree, func = "jaccard", neighbour = 8, abund = FALSE)
```

Arguments

layers	A Raster* object of species distributions (typically a multi-layer type: RasterStack or RasterBrick).
tree	An hclust or phylo object (used only for PD or FD).
func	Partial match indicating whether the Jaccard or Soerensen family of beta diversity measures should be used. If not specified, default is Jaccard.
neighbour	Either 8 (default) or 4 cells considered to calculate beta diversity of each focal cell.
abund	A boolean (T/F) indicating whether abundance data should be used (TRUE) or converted to incidence (FALSE) before analysis.

Details

The beta diversity measures used here follow the partitioning framework independently developed by Podani & Schmera (2011) and Carvalho et al. (2012) and later expanded to PD and FD by Cardoso et al. (2014), where $B_{total} = B_{repl} + B_{rich}$. B_{total} = total beta diversity, reflecting both species replacement and loss/gain; B_{repl} = beta diversity explained by replacement of species alone; B_{rich} = beta diversity explained by species loss/gain (richness differences) alone. PD and FD are calculated based on a tree (hclust or phylo object, no need to be ultrametric). The path to the root of the tree is always included in calculations of PD and FD. The number and order of species in layers must be the same as in tree.

Value

A raster.stack object representing B_{total} , B_{repl} and B_{rich} in space.

References

- Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *Journal of Biogeography*, 41, 749-761.
- Carvalho, J.C., Cardoso, P. & Gomes, P. (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. *Global Ecology and Biogeography*, 21, 760-771.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4, 379-391.
- Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data. *Oikos*, 120, 1625-1638.

Examples

```
sp1 <- raster::raster(matrix(c(NA,1,1,1,1,0,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
sp2 <- raster::raster(matrix(c(0,0,0,0,1,1,1,1,1), nrow = 3, ncol = 3, byrow = TRUE))
sp3 <- raster::raster(matrix(c(0,0,0,1,1,1,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
spp <- raster::stack(sp1, sp2, sp3)
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
raster.beta(spp)
raster.beta(spp, tree)
```

raster.dispersion *Maps of phylogenetic/functional dispersion of species or individuals.*

Description

Average dissimilarity between any two species or individuals randomly chosen in a community using rasters of species distributions (presence/absence or abundance).

Usage

```
raster.dispersion(
  layers,
  tree,
  distance,
  func = "originality",
  abund = FALSE,
  relative = FALSE
)
```

Arguments

layers	A Raster* object of species distributions (typically a multi-layer type: RasterStack or RasterBrick).
tree	An hclust or phylo object representing a phylogenetic or functional tree. One of tree or distance must be provided.
distance	A dist object representing the phylogenetic or functional distance between species.
func	Calculate dispersion using originality (default), uniqueness or contribution.
abund	A boolean (T/F) indicating whether dispersion should be calculated using individuals (T) or species (F).
relative	A boolean (T/F) indicating whether dispersion should be relative to the maximum distance between any two species in the tree or distance matrix.

Details

If abundance data is used and a tree is given, dispersion is the quadratic entropy of Rao (1982). If abundance data is not used but a tree is given, dispersion is the phylogenetic dispersion measure of Webb et al. (2002).

Value

A raster object representing dispersion in space.

References

- Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, 21: 24-43.
- Webb, C.O., Ackerly, D.D., McPeck, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33: 475-505.

Examples

```
sp1 <- raster::raster(matrix(c(NA,1,1,1,1,0,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
sp2 <- raster::raster(matrix(c(0,0,0,0,1,1,1,1,1), nrow = 3, ncol = 3, byrow = TRUE))
sp3 <- raster::raster(matrix(c(0,0,0,1,1,1,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
spp <- raster::stack(sp1, sp2, sp3)
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
raster.dispersion(spp, tree)
```

raster.evenness *Maps of phylogenetic/functional evenness of species or individuals.*

Description

Regularity of distance and abundance between any two species in a community using rasters of species distributions (presence/absence or abundance).

Usage

```
raster.evenness(
  layers,
  tree,
  distance,
  method = "expected",
  func = "camargo",
  abund = TRUE
)
```

Arguments

layers	A Raster* object of species distributions (typically a multi-layer type: RasterStack or RasterBrick).
tree	An hclust or phylo object representing a phylogenetic or functional tree. One of tree or distance must be provided.
distance	A dist object representing the phylogenetic or functional distance between species.
method	Calculate dispersion using "expected" values (default) or values based on "contribution" of species to the tree.
func	Calculate dispersion using "Camargo" (default) or "Bulla" index.
abund	A boolean (T/F) indicating whether evenness should be calculated using abundance data.

Details

Evenness is calculated based on the index of Bulla (1994) using the values of both edge lengths in the tree and their abundance.

If no tree or distance is provided the result is the original index of Bulla with correction.

Value

A raster object representing evenness in space.

References

Bulla, L. (1994) An index of evenness and its associated diversity measure. *Oikos*, 70: 167-171.

Examples

```

sp1 <- raster::raster(matrix(c(NA,1,1,1,1,0,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
sp2 <- raster::raster(matrix(c(0,0,0,0,1,1,1,1,1), nrow = 3, ncol = 3, byrow = TRUE))
sp3 <- raster::raster(matrix(c(0,0,0,1,1,1,0,0,0), nrow = 3, ncol = 3, byrow = TRUE))
spp <- raster::stack(sp1, sp2, sp3)
tree <- hclust(dist(c(1:3), method="euclidean"), method="average")
raster.evenness(spp, tree)

```

sad

*Species-abundance distribution (SAD).***Description**

Fits the SAD to community abundance data with possible rarefaction.

Usage

```
sad(comm, raref = 0, runs = 100)
```

Arguments

comm	Either a vector with the abundance per species, or a sites x species matrix.
raref	An integer specifying the number of individuals for rarefaction (individual based). If raref < 1 no rarefaction is made. If raref = 1 rarefaction is made by the minimum abundance among all sites. If raref > 1 rarefaction is made by the abundance indicated. If not specified, default is 0.
runs	Number of resampling runs for rarefaction. If not specified, default is 100.

Details

Classes defined as n = 1, 2-3, 4-7, 8-15, Rarefaction allows comparison of sites with different total abundances.

Value

A vector or matrix with the different values per class per community.

Examples

```

comm1 <- c(20,1,3,100,30)
comm2 <- c(1,2,12,0,45)
comm <- rbind(comm1, comm2)
sad(comm1)
sad(comm)
sad(comm, raref = 1)

```


Examples

```
div <- c(1,2,3,4,4)
comm <- matrix(c(2,0,0,0,3,1,0,0,2,4,5,0,1,3,2,5,1,1,1,1), nrow = 5, ncol = 4, byrow = TRUE)
tree <- hclust(dist(c(1:4), method="euclidean"), method="average")
area <- c(10,40,80,160,160)
sar(div, ,area)
sar(comm, ,area)
sar(comm, tree,area)
```

sim.plot

Plots of simulated species spatial distributions.

Description

Plots individuals from artificial communities with given SAD and spatial clustering.

Usage

```
sim.plot(comm, sad = FALSE, s = 0)
```

Arguments

comm	artificial community data from function sim.spatial.
sad	boolean indicating if the SAD plot should also be shown. Default is FALSE.
s	number of species to plot simultaneously. Default is the number of species in comm.

Details

Function useful for visualizing the results of sim.spatial.

Examples

```
comm <- sim.spatial(1000, 24)
sim.plot(comm)
sim.plot(comm, sad = TRUE)
sim.plot(comm, s = 9)
```

sim.sad	<i>Simulation of species abundance distributions (SAD).</i>
---------	---

Description

Creates artificial communities following given SADs.

Usage

```
sim.sad(n, s, sad = "lognormal", sd = 1)
```

Arguments

n	total number of individuals.
s	number of species.
sad	The SAD distribution type (lognormal, uniform, broken stick or geometric). Default is lognormal.
sd	The standard deviation of lognormal distributions. Default is 1.

Details

Species Abundance Distributions may take a number of forms. A lognormal SAD probably is the most supported by empirical data, but we include other common types useful for testing multiple algorithms including several of the functions in BAT.

Value

A matrix of species x abundance per species.

Examples

```
comm1 <- sim.sad(10000, 100)
comm2 <- sim.sad(10000, 100, sd = 2)
comm3 <- sim.sad(10000, 100, sad = "uniform")
par(mfrow=c(1,3))
hist(log(comm1$Freq))
hist(log(comm2$Freq))
hist(log(comm3$Freq))
```

sim.sample	<i>Simulation of sampling from artificial communities.</i>
------------	--

Description

Simulates a sampling process from artificial communities.

Usage

```
sim.sample(comm, cells = 100, samples = 0)
```

Arguments

comm	simulated community data from function sim.spatial.
cells	number of cells to divide the simulated space into. Default is 100.
samples	number of samples (cells) to randomly extract. Default is the number of cells (the entire community).

Details

The space will be divided in both dimensions by $\sqrt{\text{cells}}$.

Function useful for simulating sampling processes from the results of sim.spatial.

May be used as direct input to other functions (e.g. alpha, alpha.accum, beta, beta.accum) to test the behavior of multiple descriptors and estimators.

Value

A matrix of samples x species (values are abundance per species per sample).

Examples

```
comm <- sim.spatial(1000, 10)
sim.sample(comm)
sim.sample(comm, cells = 10, samples = 5)
```

sim.spatial	<i>Simulation of species spatial distributions.</i>
-------------	---

Description

Creates artificial communities with given SAD and spatial clustering.

Usage

```
sim.spatial(  
  n,  
  s,  
  sad = "lognormal",  
  sd = 1,  
  distribution = "aggregated",  
  clust = 1  
)
```

Arguments

n	total number of individuals.
s	number of species.
sad	The SAD distribution type (lognormal, uniform, broken stick or geometric). Default is lognormal.
sd	The standard deviation of lognormal distributions. Default is 1.
distribution	The spatial distribution of individual species populations (aggregated, random, uniform or gradient). Default is aggregated.
clust	The clustering parameter if distribution is either aggregated or gradient (higher values create more clustered populations). Default is 1.

Details

The spatial distribution of individuals of given species may take a number of forms. Competitive exclusion may cause overdispersion, specific habitat needs or cooperation may cause aggregation and environmental gradients may cause abundance gradients.

Value

A matrix of individuals x (species, x coords and y coords).

Examples

```
par(mfrow = c(3,3))  
comm = sim.spatial(100, 9, distribution = "uniform")  
for(i in 1:9){  
  sp <- comm[comm[,1] == paste("Sp", i, sep = ""), ]
```

```
plot(sp$x, sp$y, main = paste("Sp", i), xlim = c(0,1), ylim = c(0,1))
}
comm = sim.spatial(1000, 9, sad = "lognormal", sd = 0.5, distribution = "aggregated", clust = 2)
for(i in 1:9){
  sp <- comm[comm[,1] == paste("Sp", i, sep=""), ]
  plot(sp$x, sp$y, main = paste("Sp", i), xlim = c(0,1), ylim = c(0,1))
}
```

sim.tree

Simulation of phylogenetic or functional tree.

Description

Simulates a random tree.

Usage

```
sim.tree(s, m = 100)
```

Arguments

s	number of species.
m	a structural parameter defining the average difference between species. Default is 100. Lower numbers create trees dominated by increasingly similar species, higher numbers by increasingly dissimilar species.

Details

A very simple tree based on random genes/traits.

Value

An hclust object.

Examples

```
tree <- sim.tree(10)
plot(as.dendrogram(tree))
tree <- sim.tree(100,10)
plot(as.dendrogram(tree))
tree <- sim.tree(100,1000)
plot(as.dendrogram(tree))
```

slope	<i>Slope of accumulation curves.</i>
-------	--------------------------------------

Description

This is similar to the first derivative of the curves at each of its points.

Usage

```
slope(accum)
```

Arguments

accum	A matrix resulting from the alpha.accum or beta.accum functions (sampling units x diversity values).
-------	--

Details

Slope is the expected gain in diversity when sampling a new individual. The slope of an accumulation curve, of either observed or estimated diversity, allows verifying if the asymptote has been reached (Cardoso et al. 2011). This is an indication of either the completeness of the inventory (low final slopes of the observed curve indicate high completeness) or reliability of the estimators (stability of the slope around a value of 0 along the curve indicates reliability).

Value

A matrix of sampling units x slope values.

References

Cardoso, P., Pekar, S., Jocque, R. & Coddington, J.A. (2011) Global patterns of guild composition and functional diversity of spiders. PLoS One, 6, e21710.

Examples

```
comm1 <- matrix(c(2,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,2,2), nrow = 4, ncol = 5, byrow = TRUE)
comm2 <- matrix(c(1,1,0,0,0,0,2,1,0,0,0,0,2,1,0,0,0,0,2,1), nrow = 4, ncol = 5, byrow = TRUE)
tree <- hclust(dist(c(1:5), method="euclidean"), method="average")
acc.alpha = alpha.accum(comm1)
slope(acc.alpha)
acc.beta = beta.accum(comm1, comm2, tree)
slope(acc.beta)
```

uniqueness	<i>Phylogenetic/functional uniqueness of species.</i>
------------	---

Description

Dissimilarity between each species and the single closest in a community.

Usage

```
uniqueness(comm, tree, distance, relative = TRUE)
```

Arguments

comm	A sites x species matrix, with either abundance or incidence data. If missing, the uniqueness using the full tree or distance matrix is calculated.
tree	An hclust or phylo object representing a phylogenetic or functional tree. One of tree or distance must be provided.
distance	A dist object representing the phylogenetic or functional distance between species.
relative	A boolean (T/F) indicating whether uniqueness should be relative to the maximum distance between any two species in the tree or distance matrix.

Details

This is equivalent to the originality measure of Mouillot et al. (2013).

Value

A matrix of sites x species values.

References

Mouillot, D., Graham, N.A., Vileger, S., Mason, N.W. & Bellwood, D.R. (2013) A functional approach reveals community responses to disturbances. *Trends in Ecology and Evolution*, 28: 167-177.

Examples

```
comm <- matrix(c(1,2,0,0,0,1,1,0,0,0,0,2,2,0,0,0,0,1,1,1), nrow = 4, byrow = TRUE)
distance <- dist(c(1:5), method="euclidean")
tree <- hclust(distance, method="average")
uniqueness(tree = tree)
uniqueness(distance = distance)
uniqueness(comm, tree)
```

Index

*Topic **datasets**

- arrabida, [9](#)
 - functree, [18](#)
 - geres, [19](#)
 - gadiana, [20](#)
 - phylotree, [41](#)
- accuracy, [2](#)
- alpha, [4](#)
- alpha.accum, [5](#)
- alpha.estimate, [7](#)
- arrabida, [9](#)
- beta, [10](#)
- beta.accum, [11](#)
- beta.multi, [13](#)
- contribution, [14](#)
- dispersion, [15](#)
- evenness, [17](#)
- functree, [18](#)
- gdm, [18](#)
- geres, [19](#)
- gadiana, [20](#)
- iaor, [20](#)
- kernel.alpha, [21](#)
- kernel.beta, [23](#)
- kernel.contribution, [25](#)
- kernel.dispersion, [27](#)
- kernel.evenness, [28](#)
- kernel.originality, [30](#)
- kernel.similarity, [32](#)
- linnean, [33](#)
- optim.alpha, [34](#)
- optim.alpha.stats, [35](#)
- optim.beta, [36](#)
- optim.beta.stats, [38](#)
- optim.spatial, [39](#)
- originality, [40](#)
- phylotree, [41](#)
- raster.alpha, [41](#)
- raster.beta, [42](#)
- raster.dispersion, [43](#)
- raster.evenness, [45](#)
- sad, [46](#)
- sar, [47](#)
- sim.plot, [48](#)
- sim.sad, [49](#)
- sim.sample, [50](#)
- sim.spatial, [51](#)
- sim.tree, [52](#)
- slope, [53](#)
- uniqueness, [54](#)