Package 'SSP'

October 12, 2022

Title Simulated Sampling Procedure for Community Ecology

Version 1.0.1

Maintainer Edlin Guerra-Castro <edlinguerra@gmail.com>

Description Simulation-
based sampling protocol (SSP) is an R package design to estimate sampling effort in studies of
ecological communities based on the definition of pseudo-
multivariate standard error (MultSE) (Anderson & Santana-
Garcon, 2015) <doi:10.1111/ele.12385> and simulation
of ecological data. The theoretical background is described in Guerra-
Castro et al. (2020) <doi:10.1101/2020.03.19.996991>.

Depends R (>= 3.5.0)

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

Suggests knitr, rmarkdown, testthat, roxygen2

VignetteBuilder knitr

URL https://github.com/edlinguerra/SSP

BugReports https://github.com/edlinguerra/SSP/issues

Imports vegan, stats, sampling, ggplot2

NeedsCompilation no

Author Edlin Guerra-Castro [aut, cre],
Maite Mascaro [aut],
Nuno Simoes [aut],
Juan Cruz-Motta [aut],
Juan Cajas [aut]

Repository CRAN

Date/Publication 2020-03-28 14:10:06 UTC

R **topics documented:**

SSP-package   *SSP: Simulated Sampling Procedure for Community Ecology*

Description

SSP is an R package designed to estimate sampling effort in studies of ecological communities based on the definition of pseudo multivariate standard error (MultSE) (Anderson & Santana-Garcon 2015) and simulation of data (Guerra-Castro et al., 2020).

Details

The protocol in SSP consists in simulating several extensive data matrices that mimic some of the relevant ecological features of the community of interest using a pilot data set. For each simulated data, several sampling efforts are repeatedly executed and MultSE is calculated to each one. The mean value, 0.025 and 0.975 quantiles of MultSE for each sampling effort across all simulated data are then estimated and plotted. The mean values are standardized in relation to the lowest sampling effort (consequently, the worst precision), and an optimal sampling effort can be identified as that in which the increase in sample size do not improve the precision beyond a threshold value (e.g. 2.5%).

SSP includes seven functions: assempar for extrapolation of assemblage parameters using pilot data; simdata for simulation of several data sets based on extrapolated parameters; datquality for evaluation of plausibility of simulated data; sampsd for repeated estimations of MultSE for different sampling designs in simulated data sets; summary_ssp for summarizing the behavior of MultSE for each sampling design across all simulated data sets, ioptimum for identification of the optimal sampling effort, and plot_ssp to plot sampling effort vs MultSE of simulated data.

The SSP package is developed at GitHub (https://github.com/edlinguerra/SSP/).

Author(s)

The SSP development team is Edlin Guerra-Castro, Maite Mascaro, Nuno Simoes, Juan Cruz-Motta and Juan Cajas

References

-Anderson, M. J., & J. Santana-Garcon. (2015). Measures of precision for dissimilarity-based multivariate analysis of ecological communities. Ecology Letters 18:66-73.

-Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

Examples

```
###To speed up the simulation of these examples, the cases, sites and N were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar (data = micromollusk,
                   type= "P/A",
                   Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases= 3, N = 20, sites = 1)

#Sampling and estimation of MultSE for each sample size (few repetitions
#to speed up the example)

sam.mic<-sampsd(dat.sim = sim.mic,
             Par = par.mic,
             transformation = "P/A",
             method = "jaccard",
             n = 10,
             m = 1,
             k = 3)

#Summary of MultSE for each sampling effort
summ.mic<-summary_ssp(results = sam.mic, multi.site = FALSE)

#Cut-off points to identify optimal sampling effort
opt.mic<-ioptimum(xx = summ.mic, multi.site = FALSE)

#Plot
plot_ssp(xx = summ.mic, opt = opt.mic, multi.site = FALSE)

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar(data = sponges,
                type= "counts",
                Sest.method = "average")

#Simulation of 3 data sets, each one with 10 potential sampling units in 3 sites.
```

```
sim.spo<-simdata(par.spo, cases= 3, N = 10, sites = 3)

#Sampling and estimation of MultSE for each sampling design (few repetitions
#to speed up the example)

sam.spo<-sampsd(dat.sim = sim.spo,
                Par = par.spo,
                transformation = "square root",
                method = "bray",
                n = 10,
                m = 3,
                k = 3)

#Summary of MultSE for each sampling effort
summ.spo<-summary_ssp(results = sam.spo, multi.site = TRUE)

#Cut-off points to identify optimal sampling effort
opt.spo<-ioptimum(xx = summ.spo, multi.site = TRUE)

#Plot
plot_ssp(xx = summ.spo, opt = opt.spo, multi.site = TRUE)
```

---

assempar   *Estimation of Ecological Parameters of the Assemblage*

Description

The function extracts the main parameters of the pilot data using basic R functions as well as functions like [specpool](#) and [dispweight](#)

Usage

```
assempar(data, type, Sest.method)
```

Arguments

data	Data frame with species names (columns) and samples (rows) information. The first column should indicate the site to which the sample belongs, regardless of whether a single site has been sampled.
type	Nature of the data to be processed. It may be presence / absence ("P/A"), counts of individuals ("counts"), or coverage ("cover")
Sest.method	Method for estimating species richness. The function [specpool](#) is used for this. Available methods are the incidence-based Chao "chao", first order jackknife "jack1", second order jackknife "jack2" and Bootstrap "boot". By default, the "average" of the four estimates is used.

Details

The expected number of species in the assemblage is estimated using non-parametric methods (Gotelli et al. 2011). Due to the variability in the estimates of each approximation (Reese et al. 2014), we recommend using an average of these. The probability detection of each species is estimated among and within sites. The former is calculated as the frequency of occurrences of each species against the number of sites sampled, the second as the weighted average frequencies in sites where the species were present. Also, the degree of spatial aggregation of species (only for real counts of individuals), is identified with the index of dispersion D (Clarke et al. 2006). The corresponding properties of unseen species are approximated using the information of observed species. Specifically, the probabilities of detection are assumed to be equal to the rarest species in the pilot data. The mean and variance of the abundances are defined using random Poisson values with lambda as the overall mean of species abundances.

Value

Par   The function returns an object of class list, to be used by simdata

Note

Important: the first column should indicate the site ID of each sample (as character or numeric), even when a single site was sampled.

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx).

References

Clarke, K. R., Chapman, M. G., Somerfield, P. J., & Needham, H. R. (2006). Dispersion-based weighting of species counts in assemblage analyses. Journal of Experimental Marine Biology and Ecology, 320, 11-27.

Gotelli, N. J., & Colwell, R. K. (2011). Estimating species richness. Pages 39-54, in A. E. Magurran and B. J. McGill (editors). Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford, UK.

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

Reese, G. C., Wilson, K. R., & Flather, C. H. (2014). Performance of species richness estimators across assemblage types and survey parameters. Global Ecology and Biogeography, 23(5), 585-594.

See Also

dispweight, specpool, simdata

Examples

```
##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)
par.mic<-assempar(data = micromollusk,
                  type= "P/A",
                  Sest.method = "average")
par.mic


##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)
par.spo<-assempar(data = sponges,
                  type= "counts",
                  Sest.method = "average")
par.spo
```

---

datquality   *Diversity Metrics of Simulated and Original Data*

Description

The function estimates the average number of species, and the Simpson diversity index per sampling
unit, as well as the total multivariate dispersion of pilot data and simulated data

Usage

```
datquality(data, dat.sim, Par, transformation, method)
```

Arguments

data	Data frame with species names (columns) and samples (rows) information. The first column should indicate the site to which the sample belongs, regardless of whether a single site has been sampled or not
dat.sim	List of simulated data generated by simdata
Par	List of parameters generated by assempar
transformation	Mathematical function to reduce the weight of dominant species: 'square root', 'fourth root', 'Log (X+1)', 'P/A', 'none'
method	The appropriate distance/dissimilarity metric. The function [vegdist](#) is called for that purpose

Details

The quality of the simulated data sets is quantified through the statistical similarity with respect to
the pilot data using the following estimators: (i) average number of species per sampling unit, (ii)
diversity, defined as the average Simpson diversity index per sampling unit, and (iii) the multivariate dispersion (MVD), measured as the average dissimilarity from all sampling units to the main

centroid in the space of the dissimilarity measure used (Anderson 2006). For the simulated data, the overall mean and standard deviation for (i) and (ii) are presented. However, to assess the magnitude of variability in the simulated data, 0.95 quantiles of the MVD for all simulated data sets are also presented.

Value

divmetrics   A data frame that includes the mean and standard deviation of richness and diversity per sampling unit, and the MVD for original and 0.95 quantiles of MVD of simulated data.

Note

It is desirable that the simulated data would be similar to the data observed in terms of species richness and diversity per sampling unit.

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx).

References

Anderson, M.J. (2006) Distance-based tests for homogeneity of multivariate dispersions. Biometrics, 62, 245-253

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

See Also

vegdist, diversity

Examples

```
###To speed up the simulation of these examples, the cases, sites and n were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar (data = micromollusk,
                   type= "P/A",
                   Sest.method = "average")

#Simulation of 3 data sets, each one with 10 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases= 3, N = 10, sites = 1)

#Estimation of diversity metrics of original and simulated data
qua.mic<-datquality(data = micromollusk,
                    dat.sim = sim.mic,
```

```
                               Par = par.mic,
                               transformation = "none",
                               method = "jaccard" )
qua.mic

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar(data = sponges,
                  type= "counts",
                  Sest.method = "average")

#Simulation of 3 data sets, each one with 10 potential sampling units in 3 sites.
sim.spo<-simdata(par.spo, cases= 3, N = 10, sites = 3)

#Estimation of diversity metrics of original and simulated data
qua.spo<-datquality(data = sponges,
                    dat.sim = sim.spo,
                    Par = par.spo,
                    transformation = "square root",
                    method = "bray")
qua.spo
```

---

epibionts	*Epibionts on Caribbean mangrove roots*

Description

Data corresponds to epibenthic organisms in mangrove roots from Laguna de La Restinga National Park, Venezuela (Guerra-Castro et al. 2016).

Usage

```
data("epibionts")
```

Format

A data frame with 96 observations on the following 152 variables.

sector a factor with levels E I M

site a numeric vector

Aaptos.sp a numeric vector

Acanthophora.spicifera a numeric vector

Acetabularia.crenulata a numeric vector

Aglaothamnion.sp a numeric vector

Amathia.sp a numeric vector

`Amorphinopsis.atlantica` a numeric vector

`Amphimedon.erina` a numeric vector

`Anemonia.sargassensis` a numeric vector

`Aplidium.accarense` a numeric vector

`Aplysilla.glacialis` a numeric vector

`Ascidia.curvata` a numeric vector

`Ascidia.sp` a numeric vector

`Ascidia.sydneiensis` a numeric vector

`Balanus.sp` a numeric vector

`Bartholomea.annulata` a numeric vector

`Biemna.caribea` a numeric vector

`Bostrychia.tenella` a numeric vector

`Botrylloides.nigrum` a numeric vector

`Botrylloides.sp.1` a numeric vector

`Botrylloides.sp.2` a numeric vector

`Brachidontes.exustus` a numeric vector

`Branchiomma.conspersum` a numeric vector

`Branchiomma.nigromaculatum` a numeric vector

`Bryopsis.sp` a numeric vector

`Bugula.neritina` a numeric vector

`Bugula.sp` a numeric vector

`Calliactis.tricolor` a numeric vector

`Callyspongia..Callyspongia..pallida` a numeric vector

`Carijoa.riisei` a numeric vector

`Caulerpa.racemosa` a numeric vector

`Caulerpa.racemosa.var.peltata` a numeric vector

`Caulerpa.sertularioides` a numeric vector

`Caulerpa.verticillata` a numeric vector

`Caulibugula.sp` a numeric vector

`Celleporaria.sp` a numeric vector

`Ceramium.diaphanum` a numeric vector

`Chaetomorpha.sp.1` a numeric vector

`Chaetomorpha.sp.2` a numeric vector

`Chalinula.molitba` a numeric vector

`Chelonaplysilla.erecta` a numeric vector

`Chondrilla.nucula` a numeric vector

`Chthamalus.sp` a numeric vector

`Clathria..Clathria..microchela` a numeric vector

`Clathria.sp` a numeric vector

`Clavelina.oblonga` a numeric vector

`Clavelina.picta` a numeric vector

`Complejo.Cliona.celata` a numeric vector

`Crassostrea.rhizophorae` a numeric vector

`Dictyota.sp` a numeric vector

`Didemnum.cineraceum` a numeric vector

`Didemnum.perlucidum` a numeric vector

`Didemnum.sp` a numeric vector

`Diplosoma.listerianum` a numeric vector

`Distaplia.bermudensis` a numeric vector

`Distaplia.stylifera` a numeric vector

`Dynamena.sp` a numeric vector

`Dysidea.etheria` a numeric vector

`Dysidea.sp` a numeric vector

`Ecteinascidia.sp` a numeric vector

`Ecteinascidia.styeloides` a numeric vector

`Ecteinascidia.turbinata` a numeric vector

`Eudistoma.olivaceum` a numeric vector

`Eusynstyela.tincta` a numeric vector

`Exaiptasia.pallida` a numeric vector

`Ficopomatus.sp` a numeric vector

`Geodia.papyracea` a numeric vector

`Halichondria..Halichondria..magniconulosa` a numeric vector

`Halichondria..Halichondria..melanadocia` a numeric vector

`Haliclona..Halichoclona..magnifica` a numeric vector

`Haliclona..Reniera..implexiformis` a numeric vector

`Haliclona..Reniera..manglaris` a numeric vector

`Haliclona..Reniera..ruetzleri` a numeric vector

`Haliclona..Reniera..tubifera` a numeric vector

`Haliclona..Rhizoniera..curacaoensis` a numeric vector

`Haliclona..Soestella..caerulea` a numeric vector

`Haliclona..Soestella..smithae` a numeric vector

`Haliclona..Soestella..twincayensis` a numeric vector

`Halimeda.sp` a numeric vector

`Halisarca.sp` a numeric vector

`Halopteris.sp` a numeric vector

`Herdmania.pallida` a numeric vector

`Hippopodina.feegeensis` a numeric vector

`Hydroides.sp` a numeric vector

`Hyrtios.proteus` a numeric vector

`Iotrochota.birotulata` a numeric vector

`Ircinia.felix` a numeric vector

`Ircinia.sp` a numeric vector

`Isognomon.alatus` a numeric vector

`Kirchenpaueria.sp` a numeric vector

`Lissoclinum.sp` a numeric vector

`Lissodendoryx..Lissodendoryx..isodictyalis` a numeric vector

`Lithophyllum.pustulatum` a numeric vector

`Microcosmus.exasperatus` a numeric vector

`Molgula.occidentalis` a numeric vector

`Murrayella.periclados` a numeric vector

`Mycale..Aegogropila..carmigropila` a numeric vector

`Mycale..Aegogropila..citrina` a numeric vector

`Mycale..Carmia..magnirhaphidifera` a numeric vector

`Mycale..Carmia..microsigmatosa` a numeric vector

`Mycale..Mycale..laevis` a numeric vector

`Mycale..Zygomycale..angulosa` a numeric vector

`Mycale.sp` a numeric vector

`Nemalecium.sp` a numeric vector

`Notaulax.nudicollis` a numeric vector

`Obelia.sp` a numeric vector

`Oceanapia.nodosa` a numeric vector

`Padina.sp` a numeric vector

`Perna.viridis` a numeric vector

`Perophora.viridis` a numeric vector

`Phaeophyceae` a numeric vector

`Phallusia.nigra` a numeric vector

`Phyllangia.americana` a numeric vector

`Pinctada.imbricata` a numeric vector

`Plakortis.angulospiculatus` a numeric vector

`Polyclinum.constellatum` a numeric vector

`Polysiphonia.sp.1` a numeric vector

`Polysiphonia.sp.3` a numeric vector

`Polysiphonia.subtilissima` a numeric vector

`Pteria.colymbus` a numeric vector

`Pyura.sp..1` a numeric vector

`Pyura.sp..2` a numeric vector

`Pyura.vittata` a numeric vector

`Rhizoclonium.sp` a numeric vector

`Rhodosoma.turcicum` a numeric vector

`Sabella.sp` a numeric vector

`Sabellastarte.magnifica` a numeric vector

`Schizoporella.pungens` a numeric vector

`Scopalina.ruetzleri` a numeric vector

`Scopalina.sp` a numeric vector

`Scrupocellaria.sp` a numeric vector

`Sphacelaria.rigidula` a numeric vector

`Spongia..Spongia..pertusa` a numeric vector

`Spongia..Spongia..tubulifera` a numeric vector

`Sporolithon.episporum` a numeric vector

`Spyridia.hypnoides` a numeric vector

`Styela.canopus` a numeric vector

`Styela.sp.1` a numeric vector

`Styela.sp.2` a numeric vector

`Suberites.aurantiacus` a numeric vector

`Symplegma.brakenhielmi` a numeric vector

`Symplegma.rubra` a numeric vector

`Synnotum.circinatum` a numeric vector

`Tedania..Tedania..ignis` a numeric vector

`Terpios.manglaris` a numeric vector

`Tethya.actinia` a numeric vector

`Tethya.sp` a numeric vector

`Trididemnum.orbiculatum` a numeric vector

`Ulva.sp` a numeric vector

`Viatrix.globulifera` a numeric vector

`Zoobotryon.verticillatum` a numeric vector

Details

Data consists of the coverage (by point-intercept) of 110 taxa identified in 240 mangrove roots, sampled under a hierarchically nested spatial design that included four random sites within each of three sectors of the lagoon system corresponding to a strong environmental gradient: external (E), intermediate (M), and internal (I). The abundance of epibenthic organisms of 8 roots were described within each site, producing a total of 32 roots in each sector. This spatial protocol was repeated five times over a period of 14 months. For demonstrative purpose, data from the 4th sampling period was randomly chosen as data for this package.

Source

https://doi.org/10.3354/meps11693

References

Guerra-Castro, E. J., J. E. Conde, and J. J. Cruz-Motta. (2016). Scales of spatial variation in tropical benthic assemblages and their ecological relevance: epibionts on Caribbean mangrove roots as a model system. Marine Ecology Progress Series 548:97-110.

Examples

```
data(epibionts)
str(epibionts)
```

---

ioptimum   *Identification of the Optimal Sampling Effort*

Description

The function estimates the sampling effort in which the rate of change for each additional sampling unit can be considered optimal.

Usage

```
ioptimum(xx, multi.site = TRUE, c1 = 10, c2 = 5, c3 = 2.5)
```

Arguments

xx	A data frame generated by summary_ssp
multi.site	Logical argument indicating if several sites were simulated
c1	First cut. By default 10% improvement for each sample with respect to the highest MultSE.
c2	Second cut. By default 5% improvement for each sample with respect to the highest MultSE.
c3	Third cut. By default 2.5% improvement for each sample with respect to the highest MultSE.

Details

Sampling efforts between the minimum (i.e. 2) and c1, can be considered the necessary efforts to improve the precision. The number of samples between c1 and c2 reflects the sub-optimal sampling efforts. The number of samples between c2 and c3 indicate the optimal sampling effort. A cost / benefit criterion (e.g. Underwood, 1990) can be used to set the final sample size within this range. The sampling effort beyond c3 would imply a marginal improvement of the MultSE for each increase in sample size, which would result in an unnecessary sampling effort due to redundancy. The relationship between MultSe, sampling effort, and optimal sampling can be visualized with plot_ssp

Value

sample.cut   A vector or matrix with the sampling size for each cut point

Note

The cuts that define the sampling effort as necessary, sub-optimal, optimal or redundant are arbitrary and can be modified according to each research problem. In particular, it is possible that c3 as 2.5% is not generated because this would be achieved with a sample size larger than the maximum simulated. In this case, the maximum effort generated with sampsd will be returned with a warning message.

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx).

References

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

Underwood, A. J. (1990). Experiments in ecology and management: Their logics, functions and interpretations. Australian Journal of Ecology, 15, 365-389.

See Also

sampsd

Examples

```
###To speed up the simulation of these examples, the cases, sites and N were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar (data = micromollusk,
                   type= "P/A",
                   Sest.method = "average")
```

```
#Simulation of 3 data sets, each one with 20 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases= 3, N = 20, sites = 1)

#Sampling and estimation of MultSE for each sample size (few repetitions
#to speed up the example)

sam.mic<-sampsd(dat.sim = sim.mic,
                Par = par.mic,
                transformation = "P/A",
                method = "jaccard",
                n = 10,
                m = 1,
                k = 3)

#Summary of MultSE for each sampling effort
summ.mic<-summary_ssp(results = sam.mic, multi.site = FALSE)

#Cut-off points to identify optimal sampling effort
opt.mic<-ioptimum(xx = summ.mic, multi.site = FALSE)

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar(data = sponges,
                  type= "counts",
                  Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units in 3 sites.
sim.spo<-simdata(par.spo, cases= 3, N = 10, sites = 3)

#Sampling and estimation of MultSE for each sampling design (few repetitions
#to speed up the example)

sam.spo<-sampsd(dat.sim = sim.spo,
                Par = par.spo,
                transformation = "square root",
                method = "bray",
                n = 10,
                m = 3,
                k = 3)

#Summary of MultSE for each sampling effort
summ.spo<-summary_ssp(results = sam.spo, multi.site = TRUE)

#Cut-off points to identify optimal sampling effort
opt.spo<-ioptimum(xx = summ.spo, multi.site = TRUE)
```

micromollusk	*Micromollusks of marine shallow sandy bottoms around Cayo Nuevo, Gulf of Mexico, Mexico*

Description

Presence/absence of 68 species registered in six cores of 4 cm diameter and 10 cm depth taken in sandy bottoms around Cayo Nuevo, Gulf of Mexico, Mexico

Usage

```
data("micromollusk")
```

Format

A data frame with 6 observations on the following 69 variables.

site a numeric vector

Leptochiton.sp. a numeric vector

Ischnochiton..Ischnochiton..erythronotus a numeric vector

Arcidae.sp. a numeric vector

Arca.imbricata a numeric vector

Barbatia.domingensis a numeric vector

Bentharca.sp. a numeric vector

Arcopsis.adamsi a numeric vector

Crenella.sp. a numeric vector

Anomia.sp.. a numeric vector

Carditopsis.smithii a numeric vector

Lucinidae.. a numeric vector

Chama.sinuosa a numeric vector

Chama.sp. a numeric vector

Galeommatidae.sp. a numeric vector

Chione.elevata a numeric vector

Semele.bellastriata a numeric vector

Gastropoda.sp..1.. a numeric vector

Gastropoda.sp..2.. a numeric vector

Gastropoda.sp..3.. a numeric vector

Diodora.minuta a numeric vector

Diodora.sp... a numeric vector

Scissurella.redferni a numeric vector

Synaptocochlea.picta a numeric vector

Lodderena.ornata a numeric vector

Cerithium.sp... a numeric vector

Sansonia.tuberculata a numeric vector

Iniforis.turristhomae a numeric vector

Metaxia.rugulosa a numeric vector

Cerithiopsis.cf..iuxtafuniculata a numeric vector

Cerithiopsis.sp. a numeric vector

Vermetidae.incertae.sedis.irregularis a numeric vector

Dendropoma.corrodens a numeric vector

Vermetid.sp..C a numeric vector

Petaloconchus.mcgintyi a numeric vector

Thylacodes.sp. a numeric vector

Alvania.auberiana a numeric vector

Alvania.colombiana a numeric vector

Alvania.sp. a numeric vector

Simulamerelina.caribaea a numeric vector

Schwartziella.fischeri a numeric vector

Zebina.browniana a numeric vector

Zebina.sp. a numeric vector

Caecum.circumvolutum a numeric vector

Caecum.donmoorei a numeric vector

Caecum.floridanum a numeric vector

Caecum.johnsoni a numeric vector

Caecum.pulchellum a numeric vector

Caecum.textile a numeric vector

Caecum.sp..B a numeric vector

Meioceras.nitidum a numeric vector

Cochliolepis.striata a numeric vector

Parviturboides.interruptus a numeric vector

Vitrinella.sp. a numeric vector

Gibberula.lavalleeana a numeric vector

Prunum.apicinum a numeric vector

Volvarina.avena a numeric vector

Astyris.lunata a numeric vector

Phrontis.albus a numeric vector

Phrontis.sp. a numeric vector

Trachypollia.sp... a numeric vector

Turridae.sp..1 a numeric vector

Turridae.sp..2.. a numeric vector

Turridae.sp..3.. a numeric vector

Ammonicera.lineofuscata a numeric vector

Ammonicera.minortalis a numeric vector

Rissoella.galba a numeric vector

Pyramidellidae.sp. a numeric vector

Pseudoscilla.babylonia a numeric vector

Details

Cayo Nuevo is a small reef cay located 240 km off the North-Western coast of Yucatan. Data correspond to a study about the biodiversity of marine benthic reef habitats off the Yucatan shelf (Ortigosa, Suarez-Mozo, Barrera et al. 2018).

Source

https://doi.org/10.3897/zookeys.779.24562

References

Ortigosa, D., Suarez-Mozo, N. Y., Barrera, N. C., & Simoes, N. (2018). First survey of Interstitial molluscs from Cayo Nuevo, Campeche Bank, Gulf of Mexico. Zookeys, 779. doi:10.3897/zookeys.779.24562

Examples

```
data(micromollusk)
```

---

pilot   *Epibionts on Caribbean mangrove roots: pilot data*

Description

Data corresponds to a pilot study abput epibenthic organisms in mangrove roots from Laguna de La Restinga National Park, Venezuela (Guerra-Castro et al. 2011).

Usage

```
data("pilot")
```

Format

A data frame with 180 observations on the following 118 variables.

Sector a factor with levels E I M

Site a numeric vector

sp1 a numeric vector

sp2 a numeric vector

sp3 a numeric vector

sp4 a numeric vector

sp5 a numeric vector

sp6 a numeric vector

sp7 a numeric vector

sp8 a numeric vector

sp9 a numeric vector

sp10 a numeric vector

sp11 a numeric vector

sp12 a numeric vector

sp13 a numeric vector

sp14 a numeric vector

sp15 a numeric vector

sp16 a numeric vector

sp17 a numeric vector

sp18 a numeric vector

sp19 a numeric vector

sp20 a numeric vector

sp21 a numeric vector

sp22 a numeric vector

sp23 a numeric vector

sp24 a numeric vector

sp25 a numeric vector

sp26 a numeric vector

sp27 a numeric vector

sp28 a numeric vector

sp29 a numeric vector

sp30 a numeric vector

sp31 a numeric vector

sp32 a numeric vector

sp33 a numeric vector

sp34 a numeric vector

sp35 a numeric vector

sp36 a numeric vector

sp37 a numeric vector

sp38 a numeric vector

sp39 a numeric vector

sp40 a numeric vector

sp41 a numeric vector

sp42 a numeric vector

sp43 a numeric vector

sp44 a numeric vector

sp45 a numeric vector

sp46 a numeric vector

sp47 a numeric vector

sp48 a numeric vector

sp49 a numeric vector

sp50 a numeric vector

sp51 a numeric vector

sp52 a numeric vector

sp53 a numeric vector

sp54 a numeric vector

sp55 a numeric vector

sp56 a numeric vector

sp57 a numeric vector

sp58 a numeric vector

sp59 a numeric vector

sp60 a numeric vector

sp61 a numeric vector

sp62 a numeric vector

sp63 a numeric vector

sp64 a numeric vector

sp65 a numeric vector

sp66 a numeric vector

sp67 a numeric vector

sp68 a numeric vector

sp69 a numeric vector

sp70 a numeric vector

sp71 a numeric vector

sp72 a numeric vector

sp73 a numeric vector

sp74 a numeric vector

sp75 a numeric vector

sp76 a numeric vector

sp77 a numeric vector

sp78 a numeric vector

sp79 a numeric vector

sp80 a numeric vector

sp81 a numeric vector

sp82 a numeric vector

sp83 a numeric vector

sp84 a numeric vector

sp85 a numeric vector

sp86 a numeric vector

sp87 a numeric vector

sp88 a numeric vector

sp89 a numeric vector

sp90 a numeric vector

sp91 a numeric vector

sp92 a numeric vector

sp93 a numeric vector

sp94 a numeric vector

sp95 a numeric vector

sp96 a numeric vector

sp97 a numeric vector

sp98 a numeric vector

sp99 a numeric vector

sp100 a numeric vector

sp101 a numeric vector

sp102 a numeric vector

sp103 a numeric vector

sp104 a numeric vector

sp105 a numeric vector

sp106 a numeric vector

sp107 a numeric vector

sp108 a numeric vector

sp109 a numeric vector

sp110 a numeric vector

sp111 a numeric vector

sp112 a numeric vector

sp113 a numeric vector

sp114 a numeric vector

sp115 a numeric vector

sp116 a numeric vector

Details

Data consists of the coverage (by point-intercept) of 116 taxa identified in 180 mangrove roots, sampled under a hierarchically nested spatial design that included six random sites within each of three sectors of the lagoon system corresponding to a strong environmental gradient: external (E), intermediate (M), and internal (I). The abundance of epibenthic organisms of 10 roots were described within each site, producing a total of 60 roots in each sector. The analysis of these pilot data defined the sampling design used by Guerra-Castro et al. (2016).

Source

https://www.interciencia.net/wp-content/uploads/2018/01/923-GUERRA-8.pdf

References

Guerra-Castro, E., J. J. Cruz-Motta, and J. E. Conde. 2011. Cuantificación de la diversidad de especies incrustantes asociadas a las raíces de Rhizophora mangle L. en el Parque Nacional Laguna de La Restinga. Interciencia 36:923-930.

Guerra-Castro, E. J., J. E. Conde, and J. J. Cruz-Motta. (2016). Scales of spatial variation in tropical benthic assemblages and their ecological relevance: epibionts on Caribbean mangrove roots as a model system. Marine Ecology Progress Series 548:97-110.

Examples

```
data(pilot)
str(pilot)
```

---

plot_ssp   *SSP plot*

Description

Plotting MultSE and sampling effort relationships of simulated data

Usage

```
plot_ssp(xx, opt, multi.site)
```

Arguments

xx	A data frame generated by summary_ssp
opt	A vector or data matrix generated by ioptimum
multi.site	Logical argument indicating whether several sites were simulated

Details

This function allows to visualize the behavior of the MultSE as sampling effort increases. When the simulation involves two sampling scales, a graph for samples and one for sites will be generated. Above the MultSE~Sampling effort projection, two shaded areas are drawn, highlighting: sub-optimal improvement (light grey), and optimal improvement (dark gray). Both reflect the sampling effort that improves the precision at acceptable (light gray) or desirable levels (dark gray), but beyond the later, any gain could be considered unnecessary. In addition, for each sampling effort, the relativized improvement (in relation to the MultSE estimated with the lower sampling effort) is presented cumulatively (as percentages). This is very useful because it indicates exactly how much the precision is improved for each sampling effort.The plot is generated using ggplot2.

Value

A ggplot2 object

Note

This is an exploratory plot that can be edited using ggplot2 functions.

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx)

References

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer.

See Also

[ggplot2](#)

Examples

```
###To speed up the simulation of these examples, the cases, sites and N were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar (data = micromollusk,
                   type= "P/A",
                   Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases= 3, N = 20, sites = 1)

#Sampling and estimation of MultSE for each sample size (few repetitions
#to speed up the example)

sam.mic<-sampsd(dat.sim = sim.mic,
                Par = par.mic,
                transformation = "P/A",
                method = "jaccard",
                n = 10,
                m = 1,
                k = 3)

#Summary of MultSE for each sampling effort
summ.mic<-summary_ssp(results = sam.mic, multi.site = FALSE)

#Cut-off points to identify optimal sampling effort
opt.mic<-ioptimum(xx = summ.mic, multi.site = FALSE)

#Plot
plot_ssp(xx = summ.mic, opt = opt.mic, multi.site = FALSE)

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar(data = sponges,
                  type= "counts",
                  Sest.method = "average")

#Simulation of 3 data sets, each one with 10 potential sampling units in 3 sites.
sim.spo<-simdata(par.spo, cases= 3, N = 10, sites = 3)

#Sampling and estimation of MultSE for each sampling design (few repetitions
#to speed up the example)
```

```
sam.spo<-sampsd(dat.sim = sim.spo,
                Par = par.spo,
                transformation = "square root",
                method = "bray",
                n = 10,
                m = 3,
                k = 3)

#Summary of MultSE for each sampling effort
summ.spo<-summary_ssp(results = sam.spo, multi.site = TRUE)

#Cut-off points to identify optimal sampling effort
opt.spo<-ioptimum(xx = summ.spo, multi.site = TRUE)

#Plot
plot_ssp(xx = summ.spo, opt = opt.spo, multi.site = TRUE)
```

sampsd	*Sampling Simulated Data and Estimates of Multivariate Standard Errors*

Description

Each set of simulated data is sampled many times for each sampling effort, from 2 replicates to those defined as an argument in the function. Then, distance-based multivariate standard errors are estimated using pseudo-variance (for single site evaluation) or Mean Squares Estimates in a linear model (for multisite evaluation).

Usage

```
sampsd(dat.sim, Par, transformation, method, n, m, k)
```

Arguments

dat.sim	A list of data sets generated by [simdata](#)
Par	A list of parameters estimated by [assempar](#)
transformation	Mathematical function to reduce the weight of very dominant species: 'square root', 'fourth root', 'Log (X+1)', 'P/A', 'none'
method	The appropriate distance/dissimilarity metric (e.g. Gower, Bray–Curtis, Jaccard, etc). The function [vegdist](#) is called for that purpose.
n	Maximum number of samples to take at each site. Can be equal or less than N
m	Maximum number of sites to sample at each data set. Can be equal or less than sites
k	Number of repetitions of each sampling effort (samples and sites) for each data set

Details

If several virtual sites have been generated, subsets of sites of size 2 to m are sampled, followed by the selection of sampling units (from 2 to n) using inclusion probabilities and self-weighted two-stage sampling (Tille, 2006). Each combination of sampling effort (number of sample units and sites), are repeated several times (e.g. k = 100) for all simulated matrices. If simulated data correspond to a single site, sampling without replacement is performed several times (e.g. k = 100) for each sample size (from 2 to n) within each simulated matrix. This approach is computationally intensive, especially when k is high (> 10). Keep this in mind as it will affect the time to get results. For each sample, suitable pre-treatments are applied and distance/similarity matrices constructed using the appropriate coefficient. When simulations are done for a single site, the MultSE is estimated as $\sqrt{(V/n)}$, being V the pseudo variance measured at each sample of size n (Anderson & Santana-Garcon, 2015). When several sites were generated, MultSE are estimated using the residual mean squares and the sites mean squares from a PERMANOVA model (Anderson & Santana-Garcon, 2015).

Value

mse.results   A matrix including all estimated MultSE for each simulated data, combination
                 of sample replicates and sites for each k repetition. This matrix will be used by
                 summary_ssp

Note

For quick exploratory analyzes, keep the number of repetitions small. Once you have explored the behavior of the MultSE, you can repeat the process keeping k-values large (e.g. 100). This process will take some time and it will depend on the power of your computer.

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx).

References

Anderson, M.J. & Santana-Garcon, J. (2015) Measures of precision for dissimilarity-based multivariate analysis of ecological communities. Ecology Letters, 18, 66-73

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

Tillé, Y. (2006). Sampling algorithms. Springer, New York, NY.

See Also

assempar, simdata, summary_ssp, vegdist

Examples

```
###To speed up the simulation of these examples, the cases, sites and n were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar (data = micromollusk,
                    type= "P/A",
                    Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases = 3, N = 20, sites = 1)

#Sampling and estimation of MultSE for each sample size (few repetitions to speed up the example)
sam.mic<-sampsd(dat.sim = sim.mic,
                Par = par.mic,
                transformation = "P/A",
                method = "jaccard",
                n = 10,
                m = 1,
                k = 3)

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar(data = sponges,
                   type= "counts",
                   Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units in 3 sites.
sim.spo<-simdata(par.spo, cases = 3, N = 20, sites = 3)

#Sampling and estimation of MultSE for each sampling design (few
#repetitions to speed up the example)

sam.spo<-sampsd(dat.sim = sim.spo,
                Par = par.spo,
                transformation = "square root",
                method = "bray",
                n = 10,
                m = 3,
                k = 3)
```

---

simdata   *Simulation of Data Sets*

Description

The function simulates data sets (as many as requested) using estimated parameters from the list generated by assempar. The function returns an object of class list that includes all the simulated data to be used by datquality and sampsd.

Usage

```
simdata(Par, cases, N, sites)
```

Arguments

Par	A list of parameters estimated by assempar
cases	Number of data sets to be simulated
N	Total number of samples to be simulated in each site
sites	Total number of sites to be simulated in each data set

Details

The presence/absence of each species at each site are simulated with Bernoulli trials and probability of success equals to the empirical frequency of occurrence of each species among sites in the pilot data. For sites with the presence of a particular species, Bernoulli trials are used (with a probability of success equal to the estimated empirical frequency within the sites where it appears), to simulate the distribution of the species at that site. Once created, the P/A matrices are converted to matrices of abundances replacing presences by random values from an adequate statistical distribution and parameters equal to those estimated in the pilot data. Simulations of counts of individuals are generated using Poisson or negative binomial distributions, depending on the degree of aggregation of each species in the pilot data (McArdle & Anderson 2004; Anderson & Walsh 2013). Simulations of continuous variables (i.e. coverage, biomass), are generated using the log-normal distribution. The simulation procedure is repeated to generate as many simulated data matrices as needed.

Value

simulated.data	The function returns an object of class List, that includes all simulated data. This object will be used by sampsd and datquality

Note

This approach is not free from assumptions. Simulations do not consider any environmental constraint, neither co-occurrence structure of species. It is assumed that potential differences in species composition/abundance among samples and sites are mainly due to spatial aggregation of species, as estimated from the pilot data. Hence, any ecological property of the assemblage that was not captured by the pilot data, will not be reflected in the simulated data. Associations among species can be modeled using copulas, as suggested by Anderson et al (2019), which could be included in an upcoming version of SSP.

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx).

References

Anderson, M. J., & Walsh, D. C. I. (2013). PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? Ecological Monographs, 83(4), 557-574.

Anderson, M. J., P. de Valpine, A. Punnett, & Miller, A. E. (2019). A pathway for multivariate analysis of ecological communities using copulas. Ecology and Evolution 9:3276-3294.

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.2003.2019.996991.

McArdle, B. H., & Anderson, M. J. (2004). Variance heterogeneity, transformations, and models of species abundance: a cautionary tale. Canadian Journal of Fisheries and Aquatic Sciences, 61, 1294-1302.

See Also

sampsd, datquality

Examples

```
###To speed up the simulation of these examples, the cases, sites and N were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar(data = micromollusk,
                  type= "P/A",
                  Sest.method = "average")

#Simulation of 3 data sets, each one with 10 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases = 3, N = 10, sites = 1)

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar (data = sponges,
                   type= "counts",
                   Sest.method = "average")

#Simulation of 3 data sets, each one with 10 potential sampling units in 3 sites.
sim.spo<-simdata(par.spo, cases = 3, N = 10, sites = 3)
```

sponges	*Sponges in Alacranes Reef National Park (ARNP), Gulf of Mexico, Mexico*

Description

Counts of 41 species of sponges in 36 transects of 20 m * 1 m across 8 sites around ARNP

Usage

```
data("sponges")
```

Format

A data frame with 36 observations on the following 42 variables.

site Factor w/ 6 levels

Agelas.clathrodes a numeric vector

Agelas.dispar a numeric vector

Agelas.tubulata a numeric vector

Agelas.wiedenmayeri a numeric vector

Aiolocroia.crassa a numeric vector

Amphimedon.copressa a numeric vector

Aplysina.archeri a numeric vector

Aplysina.cauliformis a numeric vector

Aplysina.fistularis a numeric vector

Aplysina.fulva a numeric vector

Aplysina.insularis a numeric vector

Aplysina.lacunosa a numeric vector

Callyspongia.plicifera a numeric vector

Callyspongia.vaginalis a numeric vector

Callispongia.fallax a numeric vector

Callispongia.armigera a numeric vector

Cliona.delitrix a numeric vector

Cliona.varians a numeric vector

Cribochalina.vascolum a numeric vector

Dragmacidon.sp. a numeric vector

Dysidea.variabilis a numeric vector

Ectyoplasia.ferox a numeric vector

Geodia.neptuni a numeric vector

```
Hymeniacidon.caerulea a numeric vector
Iotrochota.birotulata a numeric vector
Igernella.notabilis a numeric vector
Ircinia.felix a numeric vector
Ircinia.strobilina a numeric vector
Monanchora.arbuscula a numeric vector
Mycale.laxissima a numeric vector
Mycale.laevis a numeric vector
Nipahtes.amorpha a numeric vector
Niphates.erecta a numeric vector
Niphathes.digitalis a numeric vector
Phorbas.amaranthus a numeric vector
Scopalina.rutzleri a numeric vector
Svenezea.flava a numeric vector
Spirastrella.coccinea a numeric vector
Verongula.reswigui a numeric vector
Verongula.rigida a numeric vector
Xestospongia.muta a numeric vector
```

Details

This data corresponds to a pilot study about sponge biodiversity in reef habitats in the Yucatán shelf (Ugalde et al., 2015)

Source

https://biotaxa.org/Zootaxa/article/view/zootaxa.3911.2.1

References

Ugalde, D., Gomez, P., & Simoes, N. (2015). Marine sponges (Porifera: Demospongiae) from the Gulf of Mexico, new records and redescription of Erylus trisphaerus (de Laubenfels, 1953). Zootaxa, 3911(2), 151-183.

Examples

```
data(sponges)
str(sponges)
```

---

summary_ssp   *Summary of MultSE for Each Sampling Effort in Simulated Data Sets*

Description

For each simulated data set, averages of MultSE are estimated for each sampling size. Then an overall mean, as well as lower and upper intervals of means for each sample size are tabulated. A relativization to the maximum is applied to the average MultSE and a numerical derivative, using a forward finite difference, of the resulting quantity is obtained.

Usage

```
summary_ssp(results, multi.site)
```

Arguments

results   A matrix generated by sampsd
multi.site   Logical argument indicating whether several sites were simulated

Details

For each set of simulated data, the average of the MultSE in each sampling effort is estimated (Anderson & Santana-Garcon 2015). Then, an overall mean, lower and upper quantiles of means are tabulated for each sampling effort among all simulated data. In order to have a general and comparable criteria to evaluate the rate of change of the average MultSE with respect to the sampling effort, a relativization to the maximum MultSE (obtained with the lower sampling effort) is calculated; then, a standard forward finite derivation is calculated.

Value

mse.results   A data frame including the summary of multivariate standard error for each sampling effort.

Note

This data frame can then be used to plot MultSE with respect to the sampling effort

Author(s)

Edlin Guerra-Castro (edlinguerra@gmail.com), Juan Carlos Cajas, Juan Jose Cruz-Motta, Nuno Simoes and Maite Mascaro (mmm@ciencias.unam.mx).

References

Anderson, M.J. & Santana-Garcon, J. (2015) Measures of precision for dissimilarity-based multivariate analysis of ecological communities. Ecology Letters, 18, 66-73

Guerra-Castro, E. J., J. C. Cajas, F. N. Dias Marques Simoes, J. J. Cruz-Motta, and M. Mascaro. (2020). SSP: An R package to estimate sampling effort in studies of ecological communities. bioRxiv:2020.003.2019.996991.

See Also

sampsd, ioptimum

Examples

```
###To speed up the simulation of these examples, the cases, sites and n were set small.

##Single site: micromollusk from Cayo Nuevo (Yucatan, Mexico)
data(micromollusk)

#Estimation of parameters of pilot data
par.mic<-assempar (data = micromollusk,
                   type= "P/A",
                   Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units from a single site
sim.mic<-simdata(par.mic, cases= 3, N = 10, sites = 1)

#Sampling and estimation of MultSE for each sample size (few repetitions
#to speed up the example)

sam.mic<-sampsd(dat.sim = sim.mic,
                Par = par.mic,
                transformation = "P/A",
                method = "jaccard",
                n = 10,
                m = 1,
                k = 3)

#Summary of MultSE for each sampling effort
summ.mic<-summary_ssp(results = sam.mic, multi.site = FALSE)

##Multiple sites: Sponges from Alacranes National Park (Yucatan, Mexico).
data(sponges)

#Estimation of parameters of pilot data
par.spo<-assempar(data = sponges,
                  type = "counts",
                  Sest.method = "average")

#Simulation of 3 data sets, each one with 20 potential sampling units in 3 sites.
sim.spo<-simdata(par.spo, cases= 3, N = 20, sites = 3)

#Sampling and estimation of MultSE for each sampling design (few repetitions
#to speed up the example)

sam.spo<-sampsd(dat.sim = sim.spo,
                Par = par.spo,
                transformation = "square root",
                method = "bray",
                n = 10,
                m = 3,
```

```
                   k = 3)

#Summary of MultSE for each sampling effort
summ.spo<-summary_ssp(results = sam.spo, multi.site = TRUE)
```

Index