Package 'Ostats'

October 12, 2022

Type Package

Title O-Stats, or Pairwise Community-Level Niche Overlap Statistics

Version 0.2.0

Description O-statistics, or overlap statistics, measure the degree of community-level trait overlap.
They are estimated by fitting nonparametric kernel density functions to each species' trait
distribution and calculating their areas of overlap. For instance, the median pairwise overlap
for a community is calculated by first determining the overlap of each species pair
in trait space, and then taking the median overlap of each species pair in a community.
This median overlap value is called the O-statistic (O for overlap).
The Ostats() function calculates separate univariate overlap statistics for each trait,
while the Ostats_multivariate() function calculates a single multivariate overlap statis-
tic for all traits.
O-statistics can be evaluated against null models to obtain standardized effect sizes.
'Ostats' is part of the collaborative Macrosystems Biodiversity Project ``Local- to continental-scale
drivers of biodiversity across the National Ecological Observatory Network (NEON).''
For more information on this project, see the Macrosystems Biodiversity Website
(<https://neon-biodiversity.github.io/>). Calculation of O-statistics is described in
Read et al. (2018) <doi:10.1111/ecog.03641>, and a teaching module for introducing the
underlying biological concepts at an undergraduate level is described in Grady et al.
(2018) <http://tiee.esa.org/vol/v14/issues/figure_sets/grady/abstract.html>.

License MIT + file LICENSE

Encoding UTF-8

URL https://neon-biodiversity.github.io/Ostats/

BugReports https://github.com/NEON-biodiversity/Ostats/issues

LazyData true

Depends R (>= 3.5.0)

Imports sfsmisc, matrixStats, circular, hypervolume, ggplot2,
gridExtra, viridis, grid, MASS

RoxygenNote 7.1.2

Suggests testthat (>= 2.1.0), knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Quentin D. Read [aut, cre],
Arya Yue [aut],
Isadora E. Fluck [aut],
Benjamin Baiser [aut],
John M. Grady [aut],
Phoebe L. Zarnetske [aut],
Sydne Record [aut]

Maintainer Quentin D. Read <quentin.read@usda.gov>

Repository CRAN

Date/Publication 2022-09-12 16:20:02 UTC

R topics documented:

---

ant_data   *Sample Circular Data*

Description

This is a sample dataset with discrete hourly data collection from a study in which ants were observed during 24-hour periods in multiple experimental chambers. Each row represents an observation of an individual ant, with its species identity and the time it was observed, rounded to the nearest hour. The experimental setup is described in Stuble et al. 2014.

Usage

    ant_data

Format

An object of class data.frame with 4837 rows and 3 columns.

References

Stuble, K.L., Patterson, C.M., Rodriguez-Cabal, M.A., Ribbons, R.R., Dunn, R.R. & Sanders, N.J. (2014). Ant-mediated seed dispersal in a warmed world. PeerJ, 2, e286. https://peerj.com/articles/286/

---

community_overlap	*Community Overlap Calculation*

Description

This function calculates the median or mean of pairwise overlaps between density estimates of trait distributions of all species within a community, which can be weighted by species abundances.

Usage

```
community_overlap(
  traits,
  sp,
  discrete = FALSE,
  circular = FALSE,
  normal = TRUE,
  output = "median",
  weight_type = "hmean",
  randomize_weights = FALSE,
  unique_values = NULL,
  raw = FALSE,
  circular_args = list(),
  density_args = list(),
  hypervolume_set_args = list()
)
```

Arguments

traits	a vector of trait measurements in the univariate case, or a matrix in the multi-variate case where each column is a trait.
sp	a vector with length equal to length(traits) that indicates the taxon of each individual.
discrete	whether trait data may take continuous or discrete values. Defaults to FALSE (all traits continuous). A single logical value or a logical vector with length equal to the number of columns in traits.
circular	whether trait data are circular (e.g., hours or angles). Defaults to FALSE (all traits non-circular). A single logical value or a logical vector with length equal to the number of columns in traits.
normal	if TRUE, the area under all density functions is normalized to 1, if FALSE, the area under all density functions is proportional to the number of observations in that group.

output	specifies whether median or mean is calculated.
weight_type	specifies weights to be used to calculate the median or mean.
randomize_weights	
	If TRUE, randomize weights given to pairwise overlaps within a community. This can be used to generate null models.
unique_values	Vector of all possible discrete values that `traits` can take. Only used if `discrete = TRUE` and `circular = TRUE`.
raw	If TRUE, also return the raw individual pairwise overlaps used to calculate the community-level statistic. Default is FALSE.
circular_args	optional list of additional arguments to pass to `circular`. Only used if `circular = TRUE` and `discrete = FALSE`.
density_args	list of additional arguments to be passed to `density` if univariate, or `hypervolume` if multivariate.
hypervolume_set_args	
	list of additional arguments to be passed to `hypervolume_set`. Used only in multivariate case.

Details

The function evaluates weighted mean or median of overlaps of density estimates of all species in a community taking complete cases with species abundances greater than 1 from the dataset. The default calculates the median of pairwise overlaps for the whole community using the harmonic means of abundances of the species pairs as weights, which minimizes the effect of outliners and rare species.If the argument `weight_type = "none"`, no weights are used for the calculation of mean/median. If `weight_type = "mean"`, arithmetic means of abundances are used as weights. To change the output to mean, specify the argument `output = "mean"`.

Value

The function returns the O-statistic for the community as a numeric value. If `raw = TRUE`, instead a list is returned, where the first element `value` is the numeric value, and the second element `raw` is a data frame with all the raw pairwise overlaps.

References

Read, Q. D. et al. Among-species overlap in rodent body size distributions predicts species richness along a temperature gradient. Ecography 41, 1718-1727 (2018).

Examples

```
# Keep only the relevant part of small mammal data
dat <- small_mammal_data[small_mammal_data$siteID %in% c('HARV','JORN'), ]
dat <- dat[!is.na(dat$weight), ]
dat$log_weight <- log10(dat$weight)

# Calculate median of pairwise overlaps for the community,weighted by harmonic means
# of abundances
community_overlap(traits = as.matrix(dat$log_weight),
```

```
    sp = factor(dat$taxonID))
```

| Ostats | *Calculate O-statistics (community-level pairwise niche overlap statistics)* |

Description

This is the primary function in the Ostats package. It calculates O-statistics by finding the trait density overlap among all pairs of species in each community and taking the mean or median. Next it optionally evaluates the O-statistics against a local null model. This is done separately for each trait.

Usage

```
Ostats(
  traits,
  plots,
  sp,
  discrete = FALSE,
  circular = FALSE,
  output = "median",
  weight_type = "hmean",
  run_null_model = TRUE,
  nperm = 99,
  nullqs = c(0.025, 0.975),
  shuffle_weights = FALSE,
  swap_means = FALSE,
  random_seed = NULL,
  unique_values = NULL,
  circular_args = list(),
  density_args = list(),
  verbose = FALSE
)
```

Arguments

traits	a numeric vector or matrix of trait measurements. The number of elements in the vector or number of rows in the matrix is the number of individuals, and the number of columns of the matrix is the number of traits.
plots	a factor with length equal to `nrow(traits)` that indicates the community each individual belongs to.
sp	a factor with length equal to `nrow(traits)` that indicates the taxon of each individual.

discrete   whether trait data may take continuous or discrete values. Defaults to FALSE (all
                  traits continuous). A single logical value or a logical vector with length equal to
                  the number of columns in traits. See details below.

circular   whether trait data are circular (e.g., hours or angles). Defaults to FALSE (all traits
                  non-circular). A single logical value or a logical vector with length equal to the
                  number of columns in traits. See details below.

output   specifies whether median or mean is calculated. Default "median".

weight_type   specifies weights to be used to calculate the median or mean. Default "hmean"
                  (harmonic mean), meaning each pair of species is weighted by the harmonic
                  mean of abundances.

run_null_model   whether to run a null model (if TRUE) and evaluate the O-statistics against it, or
                  simply return the raw O-statistics (if FALSE). Defaults to TRUE.

nperm   the number of null model permutations to generate. Defaults to 99.

nullqs   numeric vector of probabilities with values in [0,1] to set effect size quantiles.
                  Defaults to c(0.025, 0.975).

shuffle_weights

                  If TRUE, shuffle weights given to pairwise overlaps within a community when
                  generating null models.

swap_means   If TRUE, swap means of body sizes within a community when generating null
                  models.

random_seed   User may supply a random seed to enable reproducibility of null model output.
                  A warning is issued, and a random seed is generated based on the local time, if
                  the user does not supply a seed.

unique_values   Vector of all possible discrete values that traits can take. Only used if discrete
                  = TRUE and circular = TRUE.

circular_args   optional list of additional arguments to pass to circular. Only used if circular
                  = TRUE and discrete = FALSE. Note that continuous circular data must be pro-
                  vided in radian units.

density_args   additional arguments to pass to density, such as bw, n, or adjust. If none are
                  provided, default values are used.

verbose   If TRUE, progress messages are displayed. Defaults to FALSE.

Details

This function calculates overlap statistics and optionally evaluates them against a local null model.
By default, it calculates the median of pairwise overlaps, weighted by harmonic mean of species
abundaces of the species pairs in each community. Two results are produced, one normalizing
the area under all density functions to 1, the other making the area under all density functions
proportional to the number of observations in that group.

If discrete = FALSE, continuous kernel density functions are estimated for each species at each
community, if TRUE, discrete functions (histograms) are estimated.

If circular = TRUE and discrete = FALSE, the function circular is used to convert each column
of traits to an object of class circular. Unless additional arguments about input data type are
specified, it is assumed that the circular input data are in radian units (0 to 2*pi).

If `circular = TRUE` and `discrete = TRUE`, data will be interpreted as discrete values on a circular scale. For example, data might be integer values representing hours and ranging from 0 to 23.

If `run_null_model` is TRUE, the O-statistics are evaluated relative to a null model. When both `shuffle_weights` and `swap_means` are FALSE, null communities are generated by randomly assigning a taxon that is present in the community to each individual. If `shuffle_weights` is TRUE, species abundances are also randomly assigned to each species to weight the O-statistic for each null community. If `swap_means` is TRUE, instead of sampling individuals randomly, species means are sampled randomly among species, keeping the deviation of each individual from its species mean the same. After the null communities are generated, O-stats are calculated for each null community to compare with the observed O-stat.

Effect size statistics are calculated by z-transforming the O-statistics using the mean and standard deviation of the null distribution.

Value

The function returns a list containing four objects:

overlaps_norm   a matrix showing the O-statistic for each trait and each community, with the area under all density functions normalized to 1.

overlaps_unnorm

a matrix showing O-stats calculated with the area under all density functions proportional to the number of observations in that group.

overlaps_norm_ses

List of matrices of effect size statistics against a null model with the area under all density functions normalized to 1. `ses` contains the effect sizes (z-scores), `ses_lower` contains the effect size lower critical values for significance at the level determined by `nullqs`, and `ses_upper` contains the upper critical values. `raw_lower` and `raw_upper` are the critical values in raw units ranging from 0 to 1.

overlaps_unnorm_ses

List of matrices of effect size statistics against a null model with the area under all density functions proportional to the number of observations in that group. Elements are as in `overlaps_norm_ses`.

Author(s)

Quentin D. Read, John M. Grady, Arya Y. Yue, Isadora Fluck E., Ben Baiser, Angela Strecker, Phoebe L. Zarnetske, and Sydne Record

References

Read, Q. D. et al. Among-species overlap in rodent body size distributions predicts species richness along a temperature gradient. Ecography 41, 1718-1727 (2018).

See Also

`Ostats_multivariate` for multidimensional overlap.

`Ostats_plot` for plotting community overlap for each community.

Examples

```
# overlap statistics for body weights of small mammals in NEON sites

# Keep only the relevant part of data
dat <- small_mammal_data[small_mammal_data$siteID %in% c('HARV','JORN'), ]
dat <- dat[!is.na(dat$weight), ]
dat$log_weight <- log10(dat$weight)


#Run O-stats on the data with only a few null model iterations
Ostats_example <- Ostats(traits = as.matrix(dat[,'log_weight']),
                    sp = factor(dat$taxonID),
                    plots = factor(dat$siteID),
                    nperm = 10)
```

Ostats_multivariate   *Calculate multivariate O-statistics*

Description

This function calculates a single O-statistic across multiple traits by estimating hypervolumes for each species in multivariate trait space that contain the trait values for all the individuals in each species. It uses a stochastic estimation procedure implemented in the R package hypervolume. The community-level O-statistics estimated from the hypervolume overlaps can be evaluated against a null model.

Usage

```
Ostats_multivariate(
  traits,
  plots,
  sp,
  output = "median",
  weight_type = "hmean",
  run_null_model = TRUE,
  nperm = 99,
  nullqs = c(0.025, 0.975),
  shuffle_weights = FALSE,
  swap_means = FALSE,
  random_seed = NULL,
  hypervolume_args = list(),
  hypervolume_set_args = list(),
  verbose = FALSE
)
```

Arguments

traits	matrix of trait measurements. The number of rows in the matrix is the number of individuals, and the number of columns of the matrix is the number of traits.
plots	a factor with length equal to nrow(traits) that indicates the community each individual belongs to.
sp	a factor with length equal to nrow(traits) that indicates the taxon of each individual.
output	specifies whether median or mean is calculated.
weight_type	specifies weights to be used to calculate the median or mean.
run_null_model	whether to run a null model (if TRUE) and evaluate the O-statistics against it, or simply return the raw O-statistics (if FALSE). Defaults to TRUE.
nperm	the number of null model permutations to generate. Defaults to 99.
nullqs	numeric vector of probabilities with values in [0,1] to set effect size quantiles. Defaults to c(0.025, 0.975).
shuffle_weights	
	If TRUE, shuffle weights given to pairwise overlaps within a community when generating null models.
swap_means	If TRUE, swap means of body sizes within a community when generating null models.
random_seed	User may supply a random seed to enable reproducibility of null model output. A warning is issued, and a random seed is generated based on the local time, if the user does not supply a seed.
hypervolume_args	
	additional arguments to pass to [hypervolume](#), such as method. If none are provided, default values are used. By default, method = "gaussian".
hypervolume_set_args	
	additional arguments to pass to [hypervolume_set](#), such as num.points.max. If none are provided, default values are used.
verbose	If TRUE, progress messages are displayed. Defaults to FALSE.

Details

This function calculates multivariate O-statistics and optionally evaluates them against a local null model. By default, it calculates the median of pairwise hypervolume overlaps, weighted by harmonic mean of species abundaces of the species pairs in each community. Two results are produced, one normalizing the area under all density functions to 1, the other making the area under all density functions proportional to the number of observations in that group.

Functions from the R package hypervolume by Blonder and colleagues (Blonder 2018) are used internally. The function [hypervolume](#) stochastically estimates hypervolumes for each species in each community, and the function [hypervolume_set](#) finds the proportional overlap between each pair of hypervolumes.

If run_null_model is TRUE, the O-statistics are evaluated relative to a null model. When both shuffle_weights and swap_means are FALSE, null communities are generated by randomly assigning a taxon that is present in the community to each individual. If shuffle_weights is TRUE,

species abundances are also randomly assigned to each species to weight the O-statistic for each null community. If swap_means is TRUE, instead of sampling individuals randomly, species means are sampled randomly among species, keeping the deviation of each individual from its species mean the same. After the null communities are generated, O-stats are calculated for each null community to compare with the observed O-stat.

Effect size statistics are calculated by z-transforming the O-statistics using the mean and standard deviation of the null distribution.

Value

The function returns a list containing four objects:

overlaps_norm   a matrix showing the O-statistic for each community, with the area under all density functions normalized to 1.

overlaps_unnorm

        a matrix showing O-stats calculated with the area under all density functions proportional to the number of observations in that group.

overlaps_norm_ses

        List of matrices of effect size statistics against a null model with the area under all density functions normalized to 1. ses contains the effect sizes (z-scores), ses_lower contains the effect size lower critical values for significance at the level determined by nullqs, and ses_upper contains the upper critical values. raw_lower and raw_upper are the critical values in raw units ranging from 0 to 1.

overlaps_unnorm_ses

        List of matrices of effect size statistics against a null model with the area under all density functions proportional to the number of observations in that group. Elements are as in overlaps_norm_ses.

References

Blonder, B. Hypervolume concepts in niche- and trait-based ecology. Ecography 41, 1441–1455 (2018). https://doi.org/10.1111/ecog.03187

See Also

[Ostats](#) for univariate data.

[Ostats_multivariate_plot](#) for plotting multivariate trait overlap in each community.

Examples

```
## Not run:
# overlap statistic between populations of pitcher plants at two different sites

# Select two sites and three dimensions and scale data
site_index <- pitcher_traits[, 'site_id'] %in% c('FLK', 'MYR')
dat <- as.matrix(pitcher_traits[site_index,
                          c('rosette_diameter_1', 'pitcher_width', 'mouth_diameter')])
dat <- scale(dat, center = TRUE, scale = TRUE)
```

```
# Here a low number is used for num.points.max for example purposes
Ostats_multi_example <- Ostats_multivariate(traits = dat,
                                            plots = factor(rep(1, nrow(dat))),
                                       sp = factor(pitcher_traits$site_id[site_index]),
                                            random_seed = 111,
                                            run_null_model = FALSE,
                                            hypervolume_args = list(method = 'box'),
                                       hypervolume_set_args = list(num.points.max = 100)
)

## End(Not run)
```

Ostats_multivariate_plot

Plot multivariate community overlap

Description

This function plots the overlap of traits among species for each community in multivariate space, showing projections of trait hypervolumes into two-dimensional space for all pairs of traits.

Usage

```
Ostats_multivariate_plot(
  plots,
  sp,
  traits,
  overlap_dat = NULL,
  use_plots = NULL,
  colorvalues = NULL,
  plot_points = TRUE,
  contour_level,
  axis_expansion = 0.01,
  contour_buffer_factor = 0.25,
  panel_height = 3,
  panel_width = 3,
  units = "cm",
  hypervolume_args = list()
)
```

Arguments

plots	Site identity: a vector of names of each community.
sp	Taxon identity: a vector of species or taxa names.
traits	A matrix or data frame with rows representing individuals and columns representing traits.

overlap_dat   Optional: an object containing the output of `Ostats_multivariate`. If pro-
                  vided, overlap statistics will be displayed in the plot panels.

use_plots   a vector of sites to plot. If NULL, the function will plot all the sites.

colorvalues   Vector of color values for the density polygons. Defaults to a viridis palette if
                  none provided.

plot_points   whether to plot individual data points in addition to the hypervolume slices.
                  Default is TRUE.

contour_level   level at which to plot contour lines. If not provided by the user, a message is
                  issued stating the default plotting level is 0.01.

axis_expansion   multiplicative expansion factor by which to expand the x and y axes around the
                  hypervolume contours before plotting. Default is 0.01.

contour_buffer_factor

                  multiplicative expansion factor by which to expand the x and y axes in all di-
                  rections, relative to the range of the axis, before calculating the hypervolume
                  contours for plotting. If this is not set to a sufficiently large value, the contour
                  lines of the hypervolumes will be cut off. Default value is 0.25 (25% expansion
                  of the axis limits in all directions).

panel_height   height of the individual plot panels, in units given by `units`. Default is 3 cm.

panel_width   height of the individual plot panels, in units given by `units`. Default is 3 cm.

units   units for panel height and width. Default is centimeters.

hypervolume_args

                  additional arguments to pass to `hypervolume`, such as `method`. If none are pro-
                  vided, default values are used.

Details

Some of the code for generating contour lines is modified from `plot.HypervolumeList`.

Value

Two-dimensional projections of species trait hypervolumes for each pair of traits, plotted together
for each community to show how they overlap each other. The overlap value obtained as output
from `Ostats_multivariate` is labelled on each community graph, if provided by the user.

The class of the returned object is `Ostats_plot_object`. Calling `print` on this object will invoke
a method to draw the plot using `grid.draw`.

If more than one community is provided, a list of objects of class `Ostats_plot_object` will be
returned.

See Also

`Ostats_multivariate` for generating multivariate O-statistics

Ostats_plot   *Plot community trait overlap*

Description

This function plots the overlap of traits among species for each community. If there are multiple traits, each trait is plotted separately in one-dimensional space.

Usage

```
Ostats_plot(
  plots,
  sp,
  traits,
  overlap_dat = NULL,
  use_plots = NULL,
  n_col = 1,
  scale = "fixed",
  colorvalues = NULL,
  alpha = 0.5,
  adjust = 2,
  bin_width = 1,
  limits_x = NULL,
  legend = FALSE,
  name_x = "trait value",
  name_y = "probability density",
  normalize = TRUE,
  means = FALSE,
  circular = FALSE,
  discrete = FALSE,
  circular_args = list()
)
```

Arguments

plots	Site identity: a vector of names of each community.
sp	Taxon identity: a vector of species or taxa names.
traits	A vector of trait measurements for each individual, or a matrix or data frame with rows representing individuals and columns representing traits.
overlap_dat	Optional: an object containing the output of `Ostats`. If provided, overlap statistics will be displayed in the plot panels.
use_plots	a vector of sites to plot. If NULL, the function will plot all the sites.
n_col	Number of columns for layout of individual panels. Default is 1.
scale	If you want the scale of x, y or both x and y axis to be independent, set the argument to "free_x", "free_y" or "free" respectively. Default = "fixed" which uses the same scale across all sites. See `facet_grid`.

colorvalues	Vector of color values for the density polygons. Defaults to a viridis palette if none provided.
alpha	defines the transparency level for the density polygons. Default is 0.5.
adjust	the bandwidth adjustment of the density polygons. Default is 2. See [density](). Only used if discrete = FALSE.
bin_width	the width of each bin of the histograms. Default is 1. Only used if discrete = TRUE.
limits_x	Vector of length 2, with multiplicative factor to apply to the minimum and maximum values of each trait to expand the limits of the x axis. Default is c(0.5, 1.5), or 0.5 times the minimum and 1.5 times the maximum value of each trait, for continuous traits. For discrete traits the default is c(1, 1) or no expansion of limits.
legend	Whether to include a legend. Defaults to FALSE.
name_x	x-axis label. Default is 'trait value'
name_y	y-axis label. Default is 'probability density'
normalize	if TRUE, areas of density plots are normalized to be equal across taxa; if FALSE, areas will be proportional to abundance. Default is TRUE.
means	if TRUE, trait means for each species are plotted in an additional plot column next to the traits distribution plots for each site. Default is FALSE.
circular	if TRUE, plots density plots or histograms using polar coordinates, and estimates density using method for objects of class circular. Default is FALSE.
discrete	if TRUE, plots histograms at discrete trait values instead of smooth kernel density plots. Default is FALSE.
circular_args	optional list of additional arguments to pass to [circular](). Only used if circular = TRUE and discrete = FALSE. If no arguments are provided, default arguments to [circular]() are used.

Value

Density plots of species trait distributions plotted together for each community to show how they overlap each other. Each community is plotted on a separate panel within a multipanel figure. The overlap value obtained as output from [Ostats]() is labelled on each community graph, if provided by the user.

If trait values are discrete rather than continuous, histograms are plotted instead of kernel density plots.

If trait values are circular, a circular kernel density estimate for each species is plotted on a polar coordinate plot. If trait values are both circular and discrete, a "sunburst" plot is returned.

The class of the returned object is Ostats_plot_object. Calling print on this object will draw the plot using [grid.draw]().

If more than one trait is provided, a list of objects of class Ostats_plot_object will be returned.

See Also

[Ostats]() to Calculate O-statistics (community-level pairwise niche overlap statistics)

Examples

```
# set the arguments:
plots <- small_mammal_data$siteID
sp <- small_mammal_data$taxonID
traits <- log10(small_mammal_data$weight)

# to plot only selected sites:
use_plots <- c('BART','KONZ','JORN')


Ostats_plot(plots = plots, sp = sp, traits = traits,
            overlap_dat = small_mammal_Ostats,
            use_plots = use_plots, means = TRUE)
```

---

pitcher_traits   *Sample Multivariate Trait Data (Pitcher Plants)*

Description

These data represent a subset of trait information collected on carnivorous pitchers of the Northern Pitcher plant, *Sarracenia purpurea*, across the species' native North American latitudinal range 30.197°N-53.722°N (Freedman et al. 2021). We provide data from 5 sites out of the 36 sites sampled.

Usage

```
pitcher_traits
```

Format

An object of class data.frame with 120 rows and 9 columns.

Details

Pitcher traits measured include:

- orthogonal rosette diameters: the length of the widest part of the basal rosette and the longest length of the basal rosette that is perpendicular to the length of the widest part of the basal rosette
- pitcher length: for the randomly selected pitcher selected for measurement this represents the length from the base of the pitcher (near the rosette of the individual plant) up to the topmost part of the pitcher
- pitcher width: length of the widest part of the pitcher randomly selected for measurement
- keel width: length of the widest part of the keel (the lobe that runs from the pitcher opening to the base on the front side of the pitcher)
- mouth diameter: the length of the widest part of the pitcher opening
- lip thickness: the length of the thickness of the pitcher at its opening

See Ellison and Gotelli (2002) for a diagram illustrating traits measured.

References

Ellison, A.M., and N.J. Gotelli. 2002. Nitrogen availability alters the expression of carnivory in the Northern Pitcher Plant, *Sarracenia purpurea*. *Proceedings of the National Academy of Science* 99(7):4409-12.

Freedman, Z.B., A. McGrew, B. Baiser, M. Besson, D. Gravel, T. Poisot, S. Record, L.B. Trotta, and N.J. Gotelli. 2021. Environment-host-microbial interactions shape the *Sarracenia purpurea* microbiome at the continental scale. *Ecology*, 102(5):e03308.

---

reg_pool   *Sample Data for Ostats_regional*

Description

This is a sample dataset with small mammal data from the NEON sites Bartlett (BART) and Harvard Forest (HARV) for all available years. The two sites belong to the same domain, the Northeast domain. Each row represents an individual small mammal that was trapped and weighed, with its species identity and body mass in grams.

Usage

```
reg_pool
```

Format

An object of class data.frame with 7905 rows and 6 columns.

References

Read, Q.D., J.M. Grady, P.L. Zarnetske, S. Record, B. Baiser, J. Belmaker, M.-N. Tuanmu, A. Strecker, L. Beaudrot, and K.M. Thibault. 2018. Among-species overlap in rodent body size distributions predicts species richness along a temperature gradient. *Ecography* 41(10):1718-1727. https://doi.org/10.1111/ecog.03641

---

small_mammal_data   *Sample Continuous Trait Data (NEON Small Mammals 2015)*

Description

This is a sample dataset with small mammal data from all available NEON sites in 2015. Each row represents an individual small mammal that was trapped and weighed, with its species identity, body weight in grams, sex, and life stage provided. Site name, site code, and sampling plot ID are provided. This dataset is adapted from one published by Read et al. (2018).

Usage

```
small_mammal_data
```

Format

An object of class data.frame with 7692 rows and 7 columns.

References

NEON (National Ecological Observatory Network). Small mammal box trapping (DP1.10072.001). https://data.neonscience.org (accessed July 13, 2021)

Read, Q.D., J.M. Grady, P.L. Zarnetske, S. Record, B. Baiser, J. Belmaker, M.-N. Tuanmu, A. Strecker, L. Beaudrot, and K.M. Thibault. 2018. Among-species overlap in rodent body size distributions predicts species richness along a temperature gradient. *Ecography* 41(10):1718-1727. https://doi.org/10.1111/ecog.03641

---

small_mammal_Ostats   *Pre-calculated Overlap Statistics for NEON Small Mammals 2015*

Description

This is the pre-calculated output of the Ostats() function run on the small_mammal_data dataset, after log-transforming the body masses.

Usage

```
small_mammal_Ostats
```

Format

An object of class list of length 4.

References

Read, Q.D., J.M. Grady, P.L. Zarnetske, S. Record, B. Baiser, J. Belmaker, M.-N. Tuanmu, A. Strecker, L. Beaudrot, and K.M. Thibault. 2018. Among-species overlap in rodent body size distributions predicts species richness along a temperature gradient. *Ecography* 41(10):1718-1727. https://doi.org/10.1111/ecog.03641

Index