A Quick Introduction to iNEXT.3D via Examples

Anne Chao and Kai-Hsiang Hu

2024-02-06

`iNEXT.3D` (INterpolation and EXTrapolation for three dimensions of biodiversity) is a sequel to iNEXT (Hsieh et al., 2016). Here the three dimensions (3D) of diversity include taxonomic diversity (TD), phylogenetic diversity (PD) and functional diversity (FD). An online version "iNEXT.3D Online" (https://chao.shinyapps.io/iNEXT_3D/) is also available for users without an R background.

A unified framework based on Hill numbers (for TD) and their generalizations (Hill-Chao numbers, for PD and FD) is adopted to quantify 3D. In this framework, TD quantifies the effective number of species, PD quantifies the effective total branch length, mean-PD (PD divided by tree depth) quantifies the effective number of lineages, and FD quantifies the effective number of virtual functional groups (or functional "species"). Thus, TD, mean-PD, and FD are all in the same units of species/lineage equivalents and can be meaningfully compared; see Chao et al. (2014) for the basic standardization theory for TD, and Chao et al. (2021) for a review of the unified theory for 3D.

For each of the three dimensions of biodiversity, `iNEXT.3D` features two statistical analyses (non-asymptotic and asymptotic):

1. A non-asymptotic approach based on interpolation and extrapolation for 3D diversity (i.e., Hill-Chao numbers)

`iNEXT.3D` computes the estimated 3D diversity for standardized samples with a common sample size or sample completeness. This approach aims to compare diversity estimates for equally-large (with a common sample size) or equally-complete (with a common sample coverage) samples; it is based on the seamless rarefaction and extrapolation (R/E) sampling curves of Hill-Chao numbers for q = 0, 1 and 2. For each dimension of biodiversity, `iNEXT.3D` offers three types of R/E sampling curves:

- Sample-size-based (or size-based) R/E sampling curves: This type of sampling curve plots the diversity estimates with respect to sample size.
- Coverage-based R/E sampling curves: This type of sampling curve plots the diversity estimates with respect to sample coverage.
- Sample completeness curve: This curve depicts how sample coverage varies with sample size. The sample completeness curve provides a bridge between the size- and coverage-based R/E sampling curves.

2. An asymptotic approach to infer asymptotic 3D diversity (i.e., Hill-Chao numbers)

`iNEXT.3D` computes the estimated asymptotic 3D diversity and also plots 3D diversity profiles (q-profiles) for q between 0 and 2, in comparison with the observed diversity. Typically, the asymptotic estimates for $q \geq \geq 1$ are reliable, but for q < 1 (especially for q = 0, species richness), the asymptotic estimates represent only lower bounds. `iNEXT.3D` also features a time-profile (which depicts the observed and asymptotic estimate of PD or mean PD with respect to reference times), and a tau-profile (which depicts the observed and asymptotic estimate of FD with respect to threshold level tau).

How to cite

If you publish your work based on results from `iNEXT.3D` package, you should make references to the following methodology paper and the package:

- Chao, A., Henderson, P. A., Chiu, C.-H., Moyes, F., Hu, K-H., Dornelas, M and. Magurran, A. E. (2021). Measuring temporal change in alpha diversity: a framework integrating taxonomic, phylogenetic and functional diversity and the iNEXT.3D standardization. Methods in Ecology and Evolution, 12, 1926-1940.
- Chao, A. and Hu, K.-H. (2023). The iNEXT.3D package: interpolation and extrapolation for three dimensions of biodiversity. R package available from CRAN.

SOFTWARE NEEDED TO RUN iNEXT.3D IN R

- Required: R
- Suggested: RStudio IDE

HOW TO RUN iNEXT.3D:

The `iNEXT.3D` package can be downloaded from CRAN or Anne Chao's iNEXT.3D_github using the commands below. For a first-time installation, some additional packages must be installed and loaded; see package manual.

```
## install iNEXT.3D package from CRAN
install.packages("iNEXT.3D")
```

```
## or install the latest version from github
install.packages('devtools')
library(devtools)
install_github('AnneChao/iNEXT.3D')

## import packages
library(iNEXT.3D)
```

There are six main functions in this package:

Two functions for non-asymptotic analysis with graphical displays:

- **iNEXT3D** computes standardized 3D diversity estimates of order q = 0, 1 and 2 for rarefied and extrapolated samples at specified sample coverage values and sample sizes.
- **ggiNEXT3D** visualizes the output from the function `iNEXT3D`.

Two functions for point estimation and basic data information

- **estimate3D** computes 3D diversity of order q = 0, 1 and 2 with a particular set of user-specified level of sample sizes or sample coverage values.
- **DataInfo3D** provides basic data information based on the observed data.

Two functions for asymptotic analysis with graphical displays:

- **ObsAsy3D** computes observed and asymptotic diversity of order q between 0 and 2 (in increments of 0.2) for 3D diversity; it also computes observed and asymptotic PD for specified reference times, and observed and asymptotic FD for specified threshold levels.
- **ggObsAsy3D** visualizes the output from the function `ObsAsy3D`.

DATA INPUT FORMAT

Species abundance/incidence data format

Although species identities/names are not required to assess TD or compare TD across individual assemblages (as in the `iNEXT` package), they are required for PD and FD. Thus, for `iNEXT.3D` package, information on species identity (or any unique identification code) and assemblage affiliation is required. Two types of species abundance/incidence data are supported:

1. Individual-based abundance data (`datatype = "abundance"`): When there are multiple assemblages, in addition to the assemblage/site names (as column names) and the species names (as row names), species abundance data (reference sample) can be input as a species (in rows) by assemblage (in columns) matrix/data.frame or a list of species abundance vectors. In the special case that there is only one assemblage, all data should be read in one column.
2. Sampling-unit-based incidence data: Incidence-raw data (`datatype = "incidence_raw"`): for each assemblage, input data for a reference sample consist of a species-by-sampling-unit matrix, in addition to the sampling-unit names (as column names) and the species names (as row names). When there are N assemblages, input data consist of N lists of matrices, and each matrix is a species-by-sampling-unit matrix. Each element in the incidence raw matrix is 1 for a detection, and 0 for a non-detection. Input a matrix which combines data for all assemblages is allowed, but the argument `nT` in the function `iNEXT3D` must be specified so that the number of sampling units in each assemblage is specified.

For example, the dataset `Brazil_rainforest_abun_data` included in the `iNEXT.3D` package consists of species sample abundances of two assemblages/habitats: “Edge” and “Interior”. Run the following code to view the first 15 rows of the abundance data.

```
data("Brazil_rainforest_abun_data")
Brazil_rainforest_abun_data
```

```
                         Edge Interior
Carpotroche_brasiliensis   11       21
Astronium_concinnum       110       11
Astronium_graveolens       36        7
Spondias_macrocarpa        12        1
Spondias_venulosa           2        0
Tapirira_guianensis         7        1
Thyrsodium_spruceanum      11       11
Anaxagorea_silvatica        1       13
Annona_acutiflora           1        1
Annona_cacans               0        2
Annona_dolabripetala        3        3
Annona_sp                   0        1
Duguetia_chrysocarpa        1        1
Ephedranthus_sp1            1        0
Ephedranthus_sp2            0        1
```

We use data (`Fish_incidence_data`) collected from two time periods, namely `"2013-2015"` and `"2016-2018"`, as an example. Each time period is designated as an assemblage. The purpose was to compare 3D diversity of the two time periods. In each time period, species incidence/occurrence was recorded in 36 sampling units in each assemblage; each sampling unit represents a sampling date. Thus, there are 36 columns in each time period. Run the following code to view the first 6 rows and 6 columns for each matrix.

```
data("Fish_incidence_data")
Fish_incidence_data
```

```
$`2013-2015`
                   17/01/2013 18/02/2013 19/03/2013 17/04/2013 16/05/2013 14/06/2013
Agonus_cataphractus          0          1          1          1          0          0
Alosa_fallax                 0          0          0          0          0          0
Ammodytes_tobianus           0          0          0          0          0          0
Anguilla_anguilla            0          1          1          0          0          0
Aphia_minuta                 0          0          0          0          1          1
Arnoglossus_laterna          0          0          0          0          0          0

$`2016-2018`
                   18/01/2016 15/02/2016 16/03/2016 14/04/2016 12/05/2016 10/06/2016
Agonus_cataphractus          1          1          1          1          1          0
Alosa_fallax                 0          0          0          0          0          0
Ammodytes_tobianus           0          0          0          0          0          0
Anguilla_anguilla            0          0          0          0          0          0
Aphia_minuta                 0          0          0          0          1          0
Arnoglossus_laterna          0          0          0          0          0          0
```

Phylogenetic tree format for PD

To perform PD analysis, the phylogenetic tree (in Newick format) spanned by species observed in the pooled data is required. For the dataset `Fish_incidence_data`, the phylogenetic tree for all observed species (including species in both time periods) is stored in the file `fish_phylo_tree`; for the dataset `Brazil_rainforest_abun_data`, the phylogenetic tree for all observed species (including species in both Edge and Interior habitats) is stored in the file `Brazil_rainforest_phylo_tree`. A partial list of the tip labels and node labels are shown below.

```
data("Brazil_rainforest_phylo_tree")
Brazil_rainforest_phylo_tree

Phylogenetic tree with 425 tips and 205 internal nodes.

Tip labels:
  Carpotroche_brasiliensis, Casearia_ulmifolia, Casearia_sp4, Casearia_sylvestris,
        Casearia_sp2, Casearia_sp3, ...
Node labels:
  magnoliales_to_asterales, poales_to_asterales, , , , celastrales_to_malpighiales, ...

Rooted; includes branch lengths.
```

Species pairwise distance matrix format for FD

To perform FD analysis, the species-pairwise distance matrix (Gower distance computed from species traits) for species observed in the pooled data is required in a matrix/data.frame format. For the dataset `Fish_incidence_data`, the distance matrix for all observed species (including species in both time periods) is stored in the file `fish_dist_matrix`; for the dataset `Brazil_rainforest_abun_data`, the distance matrix for all species (including species in both Edge and Interior habitats) is stored in the file `Brazil_rainforest_dist_matrix`. The distance matrix for the first 3 Brazil rainforest tree species is shown below.

```
data("Brazil_rainforest_distance_matrix")
Brazil_rainforest_distance_matrix
```

```
                       Carpotroche_brasiliensis Astronium_concinnum Astronium_graveolens
Carpotroche_brasiliensis                    0.000               0.522                0.522
Astronium_concinnum                         0.522               0.000                0.000
Astronium_graveolens                        0.522               0.000                0.000
```

MAIN FUNCTION iNEXT3D(): RAREFACTION/EXTRAPOLATION

We first describe the main function `iNEXT3D()` with default arguments:

```
iNEXT3D(data, diversity = 'TD', q = c(0,1,2), datatype = "abundance",
        size = NULL, endpoint = NULL, knots = 40, nboot = 50, conf = 0.95, nT = NULL,
        PDtree = NULL, PDreftime = NULL, PDtype = 'meanPD',
        FDdistM, FDtype = 'AUC', FDtau = NULL, FDcut_number = 50)
```

The arguments of this function are briefly described below, and will be explained in more details by illustrative examples in later text. This main function computes standardized 3D diversity estimates of order q = 0, 1 and 2, the sample coverage estimates, and related statistics for K (if `knots = K` in the specified argument) evenly-spaced knots (sample sizes) between size 1 and the `endpoint`, where the endpoint is described below. Each knot represents a particular sample size for which 3D diversity estimates will be calculated. By default, `endpoint` = double the reference sample size for abundance data or double the total sampling units for incidence data. For example, if `endpoint = 10`, `knot = 4` is specified, diversity estimates will be computed for a sequence of samples with sizes (1, 4, 7, 10).

Argument	Description
`data`	a. For `datatype = "abundance"`, data can be input as a vector of species abundances (for a single assemblage), matrix/data.frame (species by assemblages), or a list of species abundance vectors. b. For `datatype = "incidence_raw"`, data can be input as a list of matrices/data.frames (species by sampling units); data can also be input as a single matrix/data.frame by merging all sampling units across assemblages based on species identity; in this case, the number of sampling units (nT, see below) must be specified.
`diversity`	selection of diversity type: `'TD'` = Taxonomic diversity, `'PD'` = Phylogenetic diversity, and `'FD'` = Functional diversity.
`q`	a numerical vector specifying the diversity orders. Default is `c(0, 1, 2)`.
`datatype`	data type of input data: individual-based abundance data (`datatype = "abundance"`), or species by sampling-units incidence/occurrence matrix (`datatype = "incidence_raw"`) with all entries being 0 (non-detection) or 1 (detection).
`size`	an integer vector of sample sizes (number of individuals or sampling units) for which diversity estimates will be computed. If `NULL`, then diversity estimates will be computed for those sample sizes determined by the specified/default `endpoint` and `knots`.
`endpoint`	an integer specifying the sample size that is the `endpoint` for rarefaction/extrapolation. If NULL, then `endpoint` = double the reference sample size.
`knots`	an integer specifying the number of equally-spaced `knots` (say K, default is 40) between size 1 and the `endpoint`; each knot represents a particular sample size for which diversity estimate will be calculated. If the `endpoint` is smaller than the reference sample size, then `iNEXT3D()` computes only the rarefaction estimates for approximately K evenly spaced `knots`. If the `endpoint` is larger than the reference sample size, then `iNEXT3D()` computes rarefaction estimates for approximately K/2 evenly spaced `knots` between sample size 1 and the reference sample size, and computes extrapolation estimates for approximately K/2 evenly spaced `knots` between the reference sample size and the `endpoint`.
`nboot`	a positive integer specifying the number of bootstrap replications when assessing sampling uncertainty and constructing confidence intervals. Enter 0 to skip the bootstrap procedures. Default is 50.
`conf`	a positive number < 1 specifying the level of confidence interval. Default is 0.95.
`nT`	(required only when `datatype = "incidence_raw"` and input data in a single matrix/data.frame) a vector of nonnegative integers specifying the number of sampling units in each assemblage. If assemblage names are not specified(i.e., `names(nT) = NULL`), then assemblages are automatically named as “assemblage1”, “assemblage2”,…, etc.
`PDtree`	(required argument for `diversity = "PD"`), a phylogenetic tree in Newick format for all observed species in the pooled assemblage.
`PDreftime`	(argument only for `diversity = "PD"`), a vector of numerical values specifying reference times for PD. Default is `NULL` (i.e., the age of the root of PDtree).
`PDtype`	(argument only for `diversity = "PD"`), select PD type: `PDtype = "PD"` (effective total branch length) or `PDtype = "meanPD"` (effective number of equally divergent lineages). Default is `"meanPD"`, where `meanPD = PD/tree depth`.
`FDdistM`	(required argument for `diversity = "FD"`), a species pairwise distance matrix for all species in the pooled assemblage.
	(argument only for `diversity = "FD"`), select FD type: `FDtype = "tau_values"` for FD

FDtype	under specified threshold values, or FDtype = "AUC" (area under the curve of tau-profile) for an overall FD which integrates all threshold values between zero and one. Default is "AUC".
FDtau	(argument only for diversity = "FD" and FDtype = "tau_values"), a numerical vector between 0 and 1 specifying tau values (threshold levels). If NULL (default), then threshold is set to be the mean distance between any two individuals randomly selected from the pooled assemblage (i.e., quadratic entropy).
FDcut_number	(argument only for diversity = "FD" and FDtype = "AUC"), a numeric number to cut [0, 1] interval into equal-spaced sub-intervals to obtain the AUC value by integrating the tau-profile. Equivalently, the number of tau values that will be considered to compute the integrated AUC value. Default is FDcut_number = 50. A larger value can be set to obtain more accurate AUC value.

For each dimension of diversity (TD, PD, FD), the main function iNEXT3D() returns the iNEXT3D object, which can be further used to make plots using the function ggiNEXT3D() to be described below. The "iNEXT3D" object includes three lists:

1. $TDInfo ($PDInfo,or $FDInfo) for summarizing data information.
2. $TDiNextEst ($PDiNextEst, or $FDiNextEst) for showing diversity estimates along with related statistics for a series of rarefied and extrapolated samples; there are two data frames ($size_based and $coverage_based) conditioning on standardized sample size or sample coverage, respectively.
3. $TDAsyEst ($PDAsyEst, or $FDAsyEst) for showing asymptotic diversity estimates along with related statistics.

FUNCTION ggiNEXT3D(): GRAPHIC DISPLAYS

The function ggiNEXT3D(), which extends ggplot2 with default arguments, is described as follows:

```
ggiNEXT3D(output, type = 1:3, facet.var = "Assemblage", color.var = "Order.q")
```

Here output is the iNEXT3D() object. Three types of curves are allowed for 3D diversity:

1. Sample-size-based R/E curve (type = 1): This curve plots diversity estimates with confidence intervals as a function of sample size.
2. Sample completeness curve (type = 2): This curve plots the sample coverage with respect to sample size.
3. Coverage-based R/E curve (type = 3): This curve plots the diversity estimates with confidence intervals as a function of sample coverage.

The argument facet.var = "Order.q", facet.var = "Assemblage", facet.var = "Both", or facet.var = "None" is used to create a separate plot for each value of the specified variable.

The ggiNEXT3D() function is a wrapper with the package ggplot2 to create a rarefaction/extrapolation sampling curve in a single line of code. The figure object is of class "ggplot", so it can be manipulated by using the ggplot2 tools.

TAXONOMIC DIVERSITY (TD): RAREFACTION/EXTRAPOLATION VIA EXAMPLES

EXAMPLE 1: TD rarefaction/extrapolation for abundance data

Based on the dataset (Brazil_rainforest_abun_data) included in the package, the following commands return all numerical results for TD. The first list of the output ($TDInfo) returns basic data information including the name of the Assemblage, sample size (n), observed species richness (S.obs), sample coverage estimate of the reference sample with size n (SC(n)), sample coverage estimate of the extrapolated sample with size 2n (SC(2n)) as well as the first five species abundance frequency counts in the reference sample (f1-f5). The output is identical to that based on the function DataInfo3D() by specifying diversity = 'TD' and datatype = "abundance"; see later text). Thus, if only data information is required, the simpler function DataInfo3D() (see later text) can be used to obtain the same output. More information about the observed diversity (for any order q between 0 and 2) can be obtained by function ObsAsy3D(), which will be introduced later.

```
data(Brazil_rainforest_abun_data)
output_TD_abun <- iNEXT3D(Brazil_rainforest_abun_data, diversity = 'TD', q = c(0,1,2),
                          datatype = "abundance")
output_TD_abun$TDInfo
```

```
$TDInfo
  Assemblage    n S.obs SC(n) SC(2n)  f1 f2 f3 f4 f5
1       Edge 1794   319 0.939  0.974 110 48 38 28 13
```

```
2   Interior 2074   356 0.941  0.973 123 48 41 32 19
```

The second list of the output (`$TDiNextEst`) includes size- and coverage-based standardized diversity estimates and related statistics computed for 40 knots by default (for example in the "Edge" assemblage, corresponding to the target sample size `m` = 1, 95, 189, …, 1699, 1794, 1795, 1899, …, 3588), which locates the reference sample size at the mid-point of the selected knots. There are two data frames (`$size_based` and `$coverage_based`).

The first data frame (`$size_based`) includes the name of the Assemblage, diversity order (`Order.q`), the target sample size (`m`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the size `m` is less than, equal to, or greater than the reference sample size), the diversity estimate of order q (`qTD`), the lower and upper confidence limits of diversity (`qTD.LCL` and `qTD.UCL`) conditioning on the sample size, and the corresponding sample coverage estimate (`SC`) along with the lower and upper confidence limits of sample coverage (`SC.LCL` and `SC.UCL`). These sample coverage estimates with confidence intervals are used for plotting the sample completeness curve. If the argument `nboot` is greater than zero, then a bootstrap method is applied to obtain the confidence intervals for the diversity and sample coverage estimates. Otherwise, all confidence intervals will not be computed. Here only the first six rows of the `$size_based` output are displayed:

```
output_TD_abun$TDiNextEst$size_based
```

```
  Assemblage Order.q   m      Method     qTD qTD.LCL qTD.UCL    SC SC.LCL SC.UCL
1       Edge       0   1 Rarefaction   1.000   1.000   1.000 0.012  0.010  0.013
2       Edge       0  95 Rarefaction  66.306  65.043  67.569 0.484  0.468  0.500
3       Edge       0 189 Rarefaction 106.743 104.052 109.434 0.638  0.622  0.653
4       Edge       0 284 Rarefaction 137.029 133.025 141.033 0.718  0.704  0.733
5       Edge       0 378 Rarefaction 161.010 155.820 166.200 0.768  0.755  0.782
6       Edge       0 472 Rarefaction 181.073 174.781 187.366 0.803  0.790  0.816
```

The second data frame (`$coverage_based`) includes the name of assemblage, the diversity order (`Order.q`), the target sample coverage value (`SC`), the corresponding sample size (`m`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the coverage `SC` is less than, equal to, or greater than the reference sample coverage), the diversity estimate of order q (`qTD`), the lower and upper confidence limits of diversity (`qTD.LCL` and `qTD.UCL`) conditioning on the target sample coverage value. Here only the first six rows of the `$coverage_based` output are displayed below: (Note for a fixed coverage value, the confidence interval in the `$coverage_based` table is wider than the corresponding interval in the `$size_based` table. This is because, for a given coverage value, the sample size needed to attain a fixed coverage value varies with bootstrap replication, leading to higher uncertainty on the resulting diversity estimate.)

```
output_TD_abun$TDiNextEst$coverage_based
```

```
  Assemblage Order.q    SC   m      Method     qTD qTD.LCL qTD.UCL
1       Edge       0 0.012   1 Rarefaction   1.000   0.970   1.030
2       Edge       0 0.484  95 Rarefaction  66.306  61.976  70.636
3       Edge       0 0.638 189 Rarefaction 106.743  99.830 113.657
4       Edge       0 0.718 284 Rarefaction 137.029 127.987 146.072
5       Edge       0 0.768 378 Rarefaction 161.010 150.075 171.946
6       Edge       0 0.803 472 Rarefaction 181.073 168.376 193.771
```

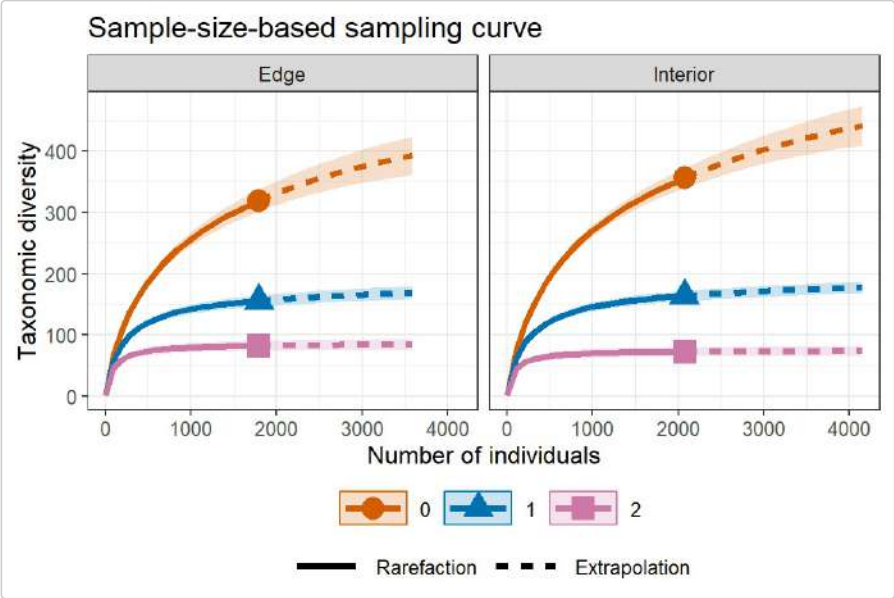
The third list of the output (`$TDAsyEst`) includes the name of the Assemblage, diversity label (`qTD`, species richness for q = 0, Shannon diversity for q = 1, and Simpson diversity for q = 2), the observed diversity (`TD_obs`), asymptotic diversity estimate (`TD_asy`) and its estimated bootstrap standard error (`s.e.`) as well as the confidence intervals for asymptotic diversity (`qTD.LCL` and `qTD.UCL`). These statistics are computed only for q = 0, 1 and 2. More detailed information about asymptotic and observed diversity estimates for any order q between 0 and 2 can be obtained from function `ObsAsy3D()`. The output for `$TDAsyEst` is shown below:

```
output_TD_abun$TDAsyEst
```

```
  Assemblage               qTD  TD_obs  TD_asy   s.e. qTD.LCL qTD.UCL
1       Edge  Species richness 319.000 444.971 28.910 388.309 501.634
2       Edge Shannon diversity 155.386 178.000  4.920 168.357 187.642
3       Edge Simpson diversity  82.023  85.905  3.753  78.550  93.261
4   Interior  Species richness 356.000 513.518 28.411 457.834 569.202
5   Interior Shannon diversity 163.514 186.983  6.553 174.139 199.827
6   Interior Simpson diversity  72.153  74.718  4.713  65.481  83.955
```

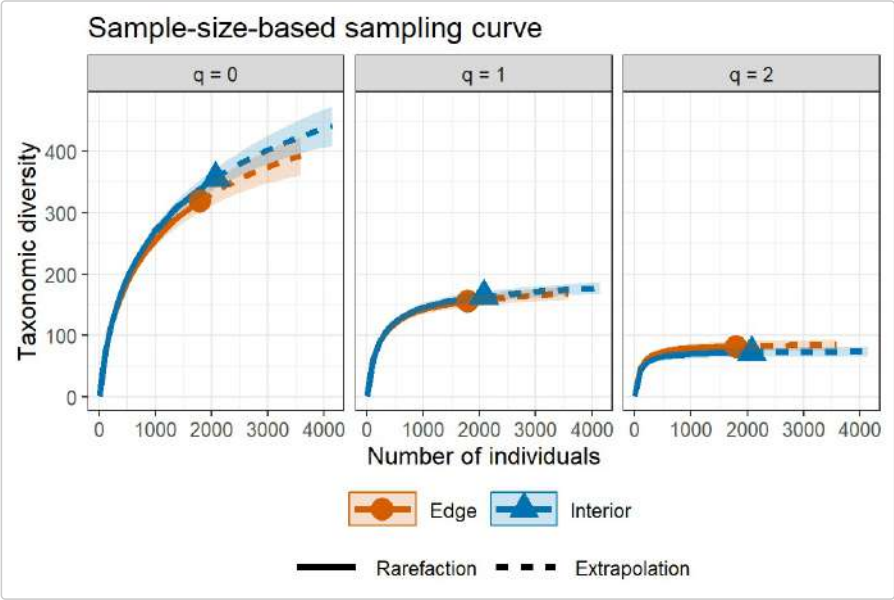
The `ggiNEXT3D` function can be used to make graphical displays for rarefaction and extrapolation sampling curves. When `facet.var = "Assemblage"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each assemblage; within each assemblage, different color curves represent diversity of different orders. An example for showing sample-size-based rarefaction/extrapolation curves (`type = 1`) is given below:

```
# TD sample-size-based R/E curves, separating by "Assemblage"
ggiNEXT3D(output_TD_abun, type = 1, facet.var = "Assemblage")
```



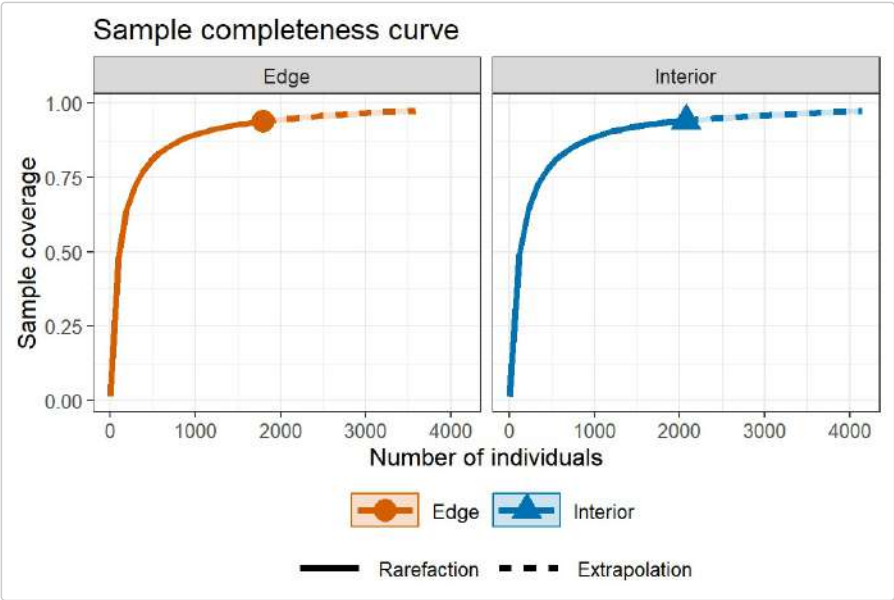
When `facet.var = "Order.q"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each diversity order; within each plot, different color curves represent different assemblages. An example is shown below:

```
# TD sample-size-based R/E curves, separating by "Order.q"
ggiNEXT3D(output_TD_abun, type = 1, facet.var = "Order.q")
```



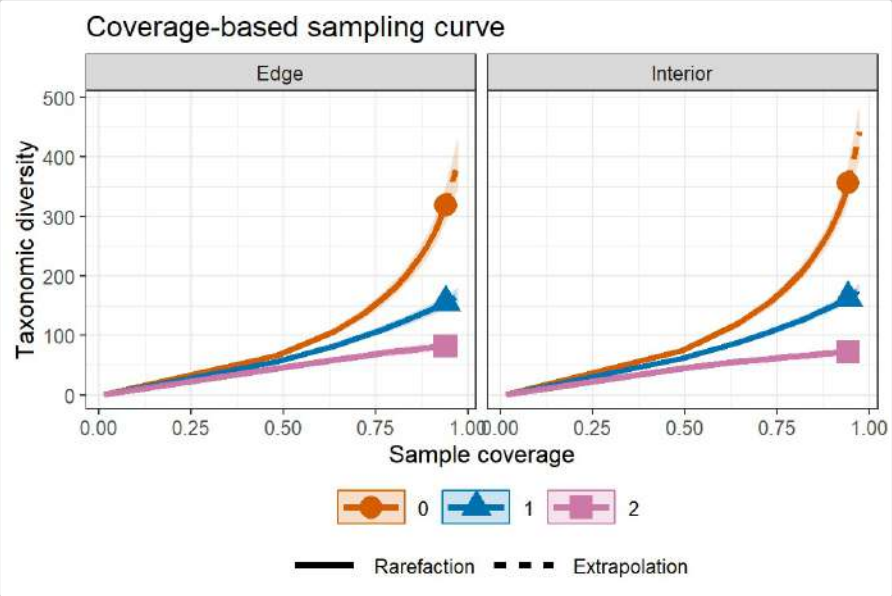
The following commands return the sample completeness (sample coverage) curve (`type = 2`) in which different colors represent different assemblages.

```
# Sample completeness curves for abundance data, separating by "Assemblage"
ggiNEXT3D(output_TD_abun, type = 2, color.var = "Assemblage")
```

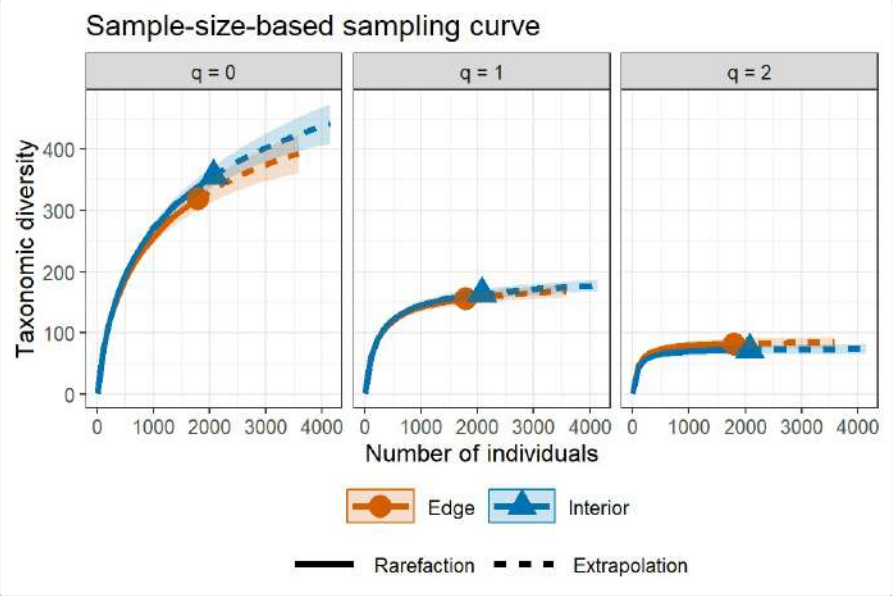


The following commands return the coverage-based rarefaction/extrapolation sampling curves in which different color curves represent three diversity orders within each assemblage (`facet.var = "Assemblage"`), or represent two assemblages within each diversity order (`facet.var = "Order.q"`), respectively.

```
# TD coverage-based R/E curves, separating by "Assemblage"
ggiNEXT3D(output_TD_abun, type = 3, facet.var = "Assemblage")
```




```
# TD coverage-based R/E curves, separating by "Order.q"
ggiNEXT3D(output_TD_abun, type = 3, facet.var = "Order.q")
```



EXAMPLE 2: TD rarefaction/extrapolation for incidence data

Based on the dataset (`Fish_incidence_data`) included in the package, the following commands return all numerical results for `TD`. The first list of the output (`$TDInfo`) returns basic data information including the name of the Assemblage, number of sampling units (`T`), total number of incidences (`U`), observed species richness (`S.obs`), sample coverage estimate of the reference sample with size T (`SC(T)`), sample coverage estimate of the extrapolated sample with size 2T (`SC(2T)`) as well as the first five species incidence frequency counts in the reference sample (`Q1-Q5`). The output is identical to that based on the function `DataInfo3D()` by specifying `diversity = 'TD'` and `datatype = "incidence_raw"`; see later text). Thus, if only data information is required, the simpler function `DataInfo3D()` (see later text) can be used to obtain the same output. More information about the observed diversity (for any order q between 0 and 2) can be obtained by function `ObsAsy3D()`, which will be introduced later.

```
data(Fish_incidence_data)
output_TD_inci <- iNEXT3D(Fish_incidence_data, diversity = 'TD', q = c(0, 1, 2),
                          datatype = "incidence_raw")
output_TD_inci$TDInfo
```

```
$TDInfo
```

```
  Assemblage  T   U S.obs SC(T) SC(2T) Q1 Q2 Q3 Q4 Q5
1  2013-2015 36 532    50 0.980  0.993 11  6  4  1  3
2  2016-2018 36 522    53 0.976  0.989 13  5  5  2  3
```

The second list of the output (`$TDiNextEst`) includes size- and coverage-based standardized diversity estimates and related statistics computed for 40 knots by default (for example in the `"2013-2015"` time period, corresponding to the target number of sample units `mT` = 1, 2, 4, …, 34, 36, 37, 38, …, 72), which locates the reference sampling units at the mid-point of the selected knots. There are two data frames (`$size_based` and `$coverage_based`).

The first data frame (`$size_based`) includes the name of the Assemblage, diversity order (`Order.q`), the target number of sampling units (`mT`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the target number of sample units `mT` is less than, equal to, or greater than the number of sampling units in the reference sample), the diversity estimate of order q (`qTD`), the lower and upper confidence limits of diversity (`qTD.LCL` and `qTD.UCL`) conditioning on the sample size, and the corresponding sample coverage estimate (`SC`) along with the lower and upper confidence limits of sample coverage (`SC.LCL` and `SC.UCL`). These sample coverage estimates with confidence intervals are used for plotting the sample completeness curve. If the argument `nboot` is greater than zero, then a bootstrap method is applied to obtain the confidence intervals for the diversity and sample coverage estimates. Otherwise, all confidence intervals will not be computed. Here only the first six rows of the `$size_based` output are displayed:

```
output_TD_inci$TDiNextEst$size_based
```

```
  Assemblage Order.q mT      Method    qTD qTD.LCL qTD.UCL    SC SC.LCL SC.UCL
1  2013-2015       0  1 Rarefaction 14.778  13.921  15.635 0.606  0.575  0.636
2  2013-2015       0  2 Rarefaction 20.603  19.460  21.746 0.749  0.724  0.773
3  2013-2015       0  4 Rarefaction 27.079  25.501  28.658 0.851  0.833  0.868
4  2013-2015       0  6 Rarefaction 31.121  29.209  33.034 0.894  0.880  0.909
5  2013-2015       0  8 Rarefaction 34.042  31.847  36.237 0.919  0.906  0.931
6  2013-2015       0 10 Rarefaction 36.319  33.873  38.765 0.934  0.923  0.945
```

The second data frame (`$coverage_based`) includes the name of assemblage, the diversity order (`Order.q`), the target sample coverage value (`SC`), the corresponding number of sampling units (`mT`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the coverage `SC` is less than, equal to, or greater than the reference sample coverage), the diversity estimate of order q (`qTD`), the lower and upper confidence limits of diversity (`qTD.LCL` and `qTD.UCL`) conditioning on the target sample coverage value. Here only the first six rows of the `$coverage_based` output are displayed below: (Note for a fixed coverage value, the confidence interval in the `$coverage_based` table is wider than the corresponding interval in the `$size_based` table. This is because, for a given coverage value, the sample size needed to attain a fixed coverage value varies with bootstrap replication, leading to higher uncertainty on the resulting diversity estimate.)

```
output_TD_inci$TDiNextEst$coverage_based
```

```
  Assemblage Order.q    SC mT      Method    qTD qTD.LCL qTD.UCL
1  2013-2015       0 0.606  1 Rarefaction 14.778  13.769  15.787
2  2013-2015       0 0.749  2 Rarefaction 20.603  18.962  22.244
3  2013-2015       0 0.851  4 Rarefaction 27.079  24.751  29.408
4  2013-2015       0 0.894  6 Rarefaction 31.121  28.310  33.933
5  2013-2015       0 0.919  8 Rarefaction 34.042  30.836  37.247
6  2013-2015       0 0.934 10 Rarefaction 36.319  32.771  39.867
```

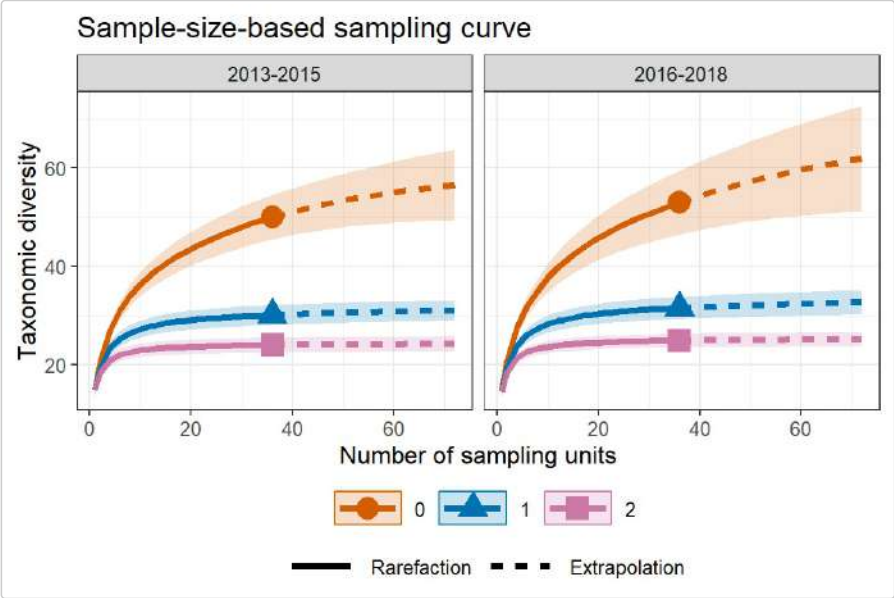
The third list of the output (`$TDAsyEst`) includes the name of the Assemblage, diversity label (`qTD`, species richness for q = 0, Shannon diversity for q = 1, and Simpson diversity for q = 2), the observed diversity (`TD_obs`), asymptotic diversity estimate (`TD_asy`) and its estimated bootstrap standard error (`s.e.`) as well as the confidence intervals for asymptotic diversity (`qTD.LCL` and `qTD.UCL`). These statistics are computed only for q = 0, 1 and 2. More detailed information about asymptotic and observed diversity estimates for any order q between 0 and 2 can be obtained from function `ObsAsy3D()`. The output is shown below:

```
output_TD_inci$TDAsyEst
```

```
  Assemblage              qTD TD_obs TD_asy   s.e. qTD.LCL qTD.UCL
1  2013-2015  Species richness 50.000 59.803 18.179  24.173  95.433
2  2013-2015 Shannon diversity 30.089 31.542  1.173  29.243  33.840
3  2013-2015 Simpson diversity 23.961 24.394  0.885  22.659  26.128
4  2016-2018  Species richness 53.000 69.431  9.946  49.937  88.924
5  2016-2018 Shannon diversity 31.534 33.393  1.388  30.674  36.113
6  2016-2018 Simpson diversity 24.889 25.409  0.848  23.746  27.072
```

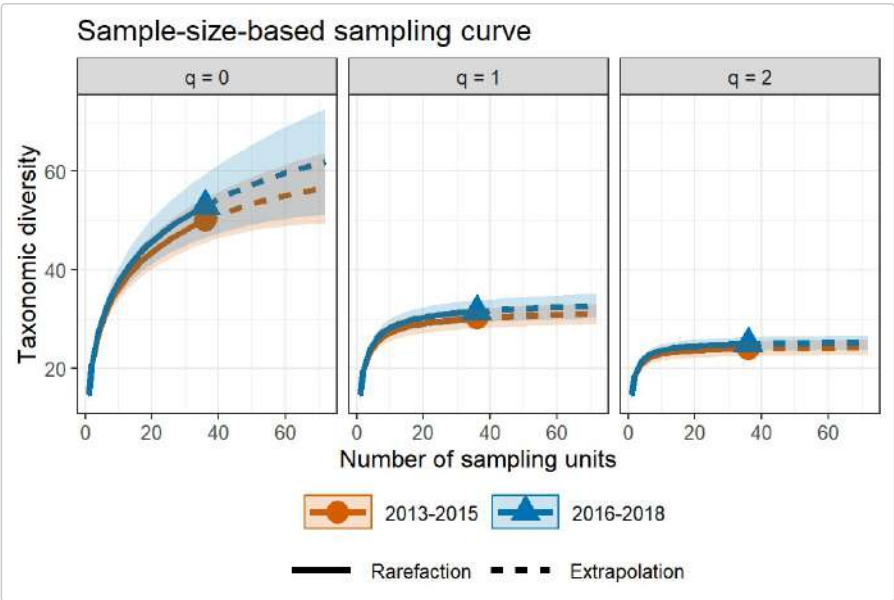
The `ggiNEXT3D` function can be used to make graphical displays for rarefaction and extrapolation sampling curves. When `facet.var = "Assemblage"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each assemblage; within each assemblage, different color curves represent diversity of different orders. An example for showing sample-size-based rarefaction/extrapolation curves (`type = 1`) for incidence data is given below:

```
# TD sample-size-based R/E curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_TD_inci, type = 1, facet.var = "Assemblage")
```



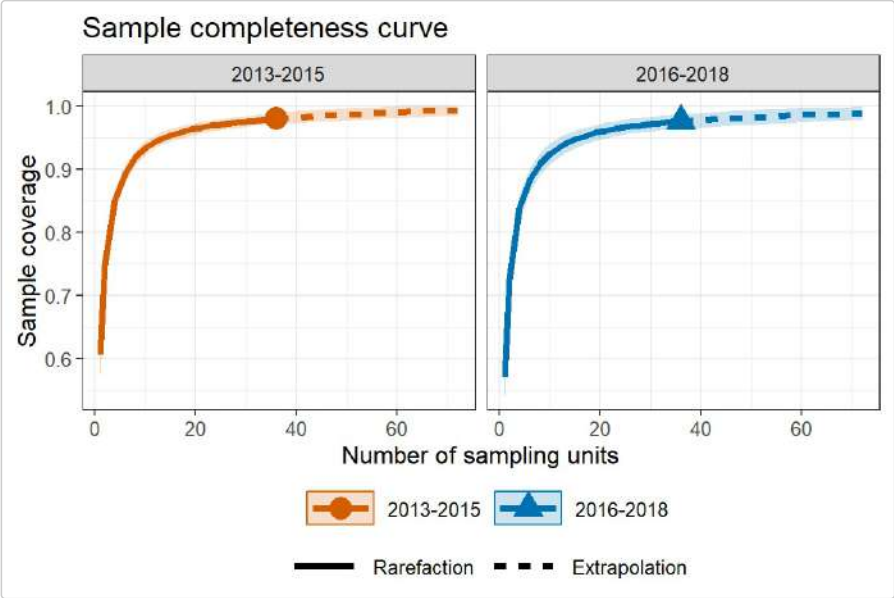
When `facet.var = "Order.q"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each diversity order; within each plot, different color curves represent different assemblages. An example is shown below:

```
# TD sample-size-based R/E curves for incidence data, separating by "Order.q"
ggiNEXT3D(output_TD_inci, type = 1, facet.var = "Order.q")
```



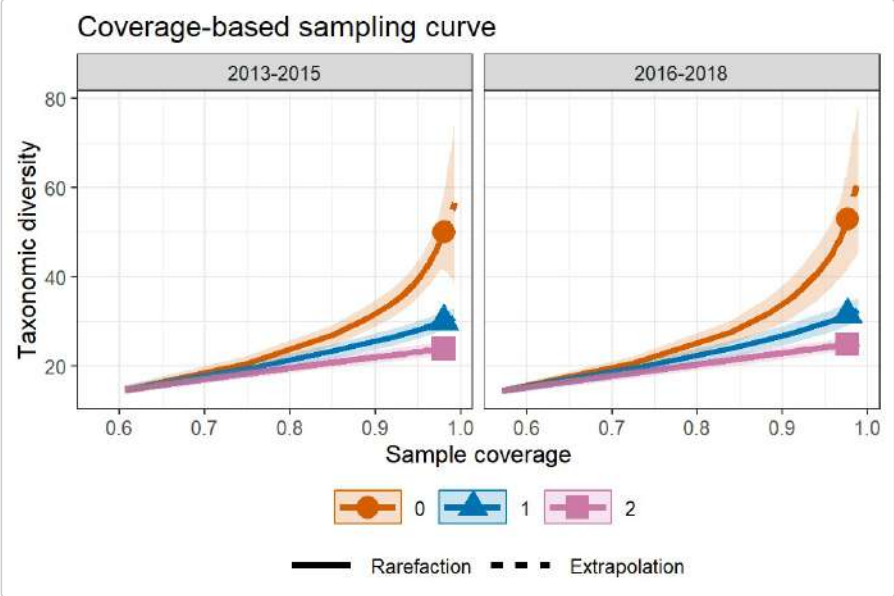
The following commands return the sample completeness (sample coverage) curve (`type = 2`) in which different colors are used for different assemblages.

```
# Sample completeness curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_TD_inci, type = 2, color.var = "Assemblage")
```

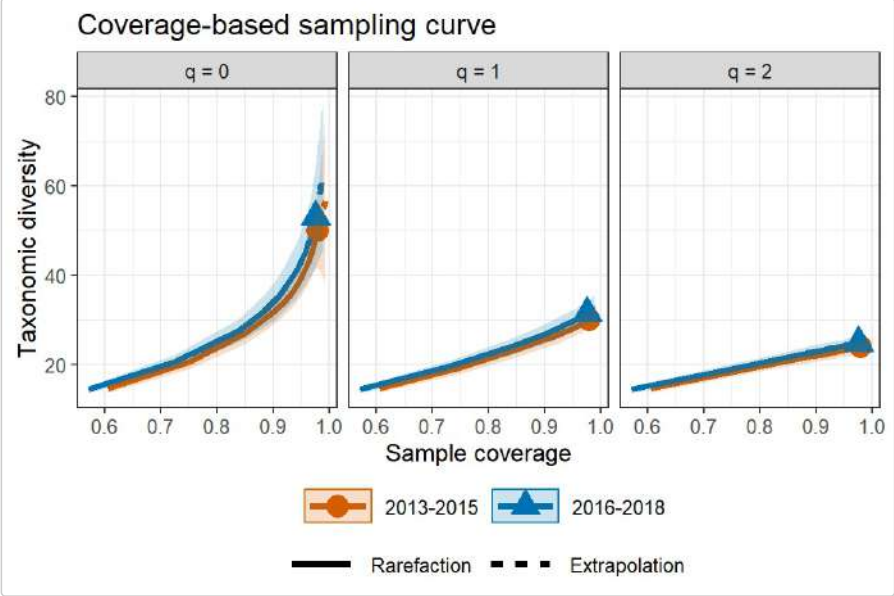


The following commands return the coverage-based rarefaction/extrapolation sampling curves in which different color curves represent three diversity orders within each assemblage (`facet.var = "Assemblage"`), or represent two assemblages within each diversity order (`facet.var = "Order.q"`), respectively.

```
# TD coverage-based R/E curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_TD_inci, type = 3, facet.var = "Assemblage")
```




```
# TD coverage-based R/E curves for incidence data, separating by "Order.q"
ggiNEXT3D(output_TD_inci, type = 3, facet.var = "Order.q")
```



PHYLOGENETIC DIVERSITY (PD): RAREFACTION/EXTRAPOLATION VIA EXAMPLES

EXAMPLE 3: PD rarefaction/extrapolation for abundance data

Based on the dataset (`Brazil_rainforest_abun_data`) and the phylogentic tree (`Brazil_rainforest_phylo_tree`) included in the package, the following commands return all numerical results for `PD`. The first list of the output (`$PDInfo`) returns basic data information including the name of the Assemblage, sample size (`n`), observed species richness (`S.obs`), sample coverage estimate of the reference sample with size n (`SC(n)`), sample coverage estimate of the extrapolated sample with size 2n (`SC(2n)`), the observed total branch length in the phylogenetic tree spanned by all observed specise (`PD.obs`), the number of singletons and doubletons in the node/branch abundance set (`f1*,f2*`), the total branch length of those singletons and doubletons in the node/branch abundance set (`g1,g2`), and the reference time (`Reftime`). The output is identical to that based on the function `DataInfo3D()` by specifying `diversity = 'PD'` and `datatype = "abundance"`; see later text). Thus, if only data information is required, the simpler function `DataInfo3D()` (see later text) can be used to obtain the same output. More information about the observed diversity (for any order q between 0 and 2) can be obtained by function `ObsAsy3D()`, which will be introduced later.

The required argument for performing PD analysis is `PDtree`. For example, the phylogenetic tree for all observed species (including species in both Edge and Interior habitats) is stored in `Brazil_rainforest_phylo_tree`. Then we enter the argument `PDtree = Brazil_rainforest_phylo_tree`. Two optional arguments are: `PDtype` and `PDreftime`. There are two options for `PDtype`: `"PD"` (effective total branch length) or `"meanPD"` (effective number of equally divergent lineages, `meanPD` = PD/tree depth). Default is `PDtype = "meanPD"`. `PDreftime` is a numerical value specifying a reference time for computing phylogenetic diversity. By default (`PDreftime = NULL`), the reference time is set to the tree depth, i.e., age of the root of the phylogenetic tree. Run the following code to perform PD analysis.

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_phylo_tree)
data <- Brazil_rainforest_abun_data
tree <- Brazil_rainforest_phylo_tree
output_PD_abun <- iNEXT3D(data, diversity = 'PD', q = c(0, 1, 2), datatype = "abundance",
                          nboot = 20, PDtree = tree)
output_PD_abun$PDInfo
```

```
$PDInfo
# A tibble: 2 x 11
  Assemblage     n S.obs `SC(n)` `SC(2n)` PD.obs `f1*` `f2*`    g1    g2 Reftime
  <chr>      <int> <int>   <dbl>    <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
1 Edge        1794   319   0.939    0.974  24516   110    52  6578  2885     400
2 Interior    2074   356   0.941    0.973  27727   123    56  7065  3656     400
```

The second list of the output (`$PDiNextEst`) includes size- and coverage-based standardized diversity estimates and related statistics computed for 40 knots by default (for example in the “Edge” assemblage, corresponding to the target sample size `m` = 1, 95, 189, …, 1699, 1794, 1795, 1899, …, 3588), which locates the reference sample size at the mid-point of the selected knots. There are two data frames (`$size_based` and `$coverage_based`).

The first data frame (`$size_based`) includes the name of the Assemblage, diversity order (`Order.q`), the target sample size (`m`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the size `m` is less than, equal to, or greater than the reference sample size), the diversity estimate of order q (`qPD`), the lower and upper confidence limits of diversity (`qPD.LCL` and `qPD.UCL`) conditioning on the sample size, the corresponding sample coverage estimate (`SC`) along with the lower and upper confidence limits of sample coverage (`SC.LCL` and `SC.UCL`), the reference time (`Reftime`) and the type of PD (`Type`). These sample coverage estimates with confidence intervals are used for plotting the sample completeness curve. If the argument `nboot` is greater than zero, then a bootstrap method is applied to obtain the confidence intervals for the diversity and sample coverage estimates. Otherwise, all confidence intervals will not be computed. Here only the first six rows of the `$size_based` output are displayed:

```
output_PD_abun$PDiNextEst$size_based
```

```
  Assemblage Order.q   m      Method    qPD qPD.LCL qPD.UCL    SC SC.LCL SC.UCL Reftime   Type
1       Edge       0   1 Rarefaction  1.000   0.984   1.016 0.012  0.011  0.013     400 meanPD
2       Edge       0  95 Rarefaction 18.547  17.956  19.137 0.484  0.469  0.499     400 meanPD
3       Edge       0 189 Rarefaction 26.723  25.867  27.579 0.638  0.624  0.652     400 meanPD
4       Edge       0 284 Rarefaction 32.305  31.275  33.336 0.718  0.706  0.731     400 meanPD
5       Edge       0 378 Rarefaction 36.498  35.336  37.661 0.768  0.757  0.780     400 meanPD
6       Edge       0 472 Rarefaction 39.882  38.610  41.153 0.803  0.792  0.814     400 meanPD
```

The second data frame (`$coverage_based`) includes the name of assemblage, the diversity order (`Order.q`), the target sample coverage value (`SC`), the corresponding sample size (`m`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the coverage `SC` is less than, equal to, or greater than the reference sample coverage), the diversity estimate of order q (`qPD`), the lower and upper confidence limits of diversity (`qPD.LCL` and `qPD.UCL`) conditioning on the target sample coverage value, the reference times (`Reftime`) and the

type of PD (`Type`). Here only the first six rows of the `$coverage_based` output are displayed below: (Note for a fixed coverage value, the confidence interval in the `$coverage_based` table is wider than the corresponding interval in the `$size_based` table. This is because, for a given coverage value, the sample size needed to attain a fixed coverage value varies with bootstrap replication, leading to higher uncertainty on the resulting diversity estimate.)

```
output_PD_abun$PDiNextEst$coverage_based
```

```
  Assemblage Order.q     SC   m      Method    qPD qPD.LCL qPD.UCL Reftime   Type
1       Edge       0 0.012   1 Rarefaction  1.000   0.983   1.017     400 meanPD
2       Edge       0 0.484  95 Rarefaction 18.547  17.553  19.541     400 meanPD
3       Edge       0 0.638 189 Rarefaction 26.723  25.350  28.097     400 meanPD
4       Edge       0 0.718 284 Rarefaction 32.305  30.674  33.936     400 meanPD
5       Edge       0 0.768 378 Rarefaction 36.498  34.671  38.325     400 meanPD
6       Edge       0 0.803 472 Rarefaction 39.882  37.898  41.866     400 meanPD
```

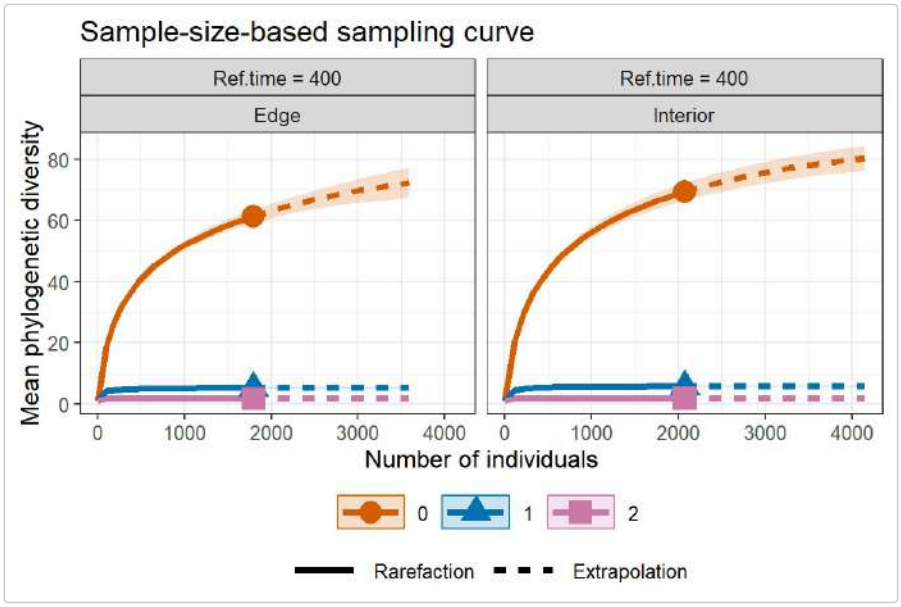
The third list of the output (`$PDAsyEst`) includes the name of the Assemblage, PD (or meanPD) for q = 0, 1, and 2 (`qPD`), the observed diversity (`PD_obs`), asymptotic diversity estimates (`PD_asy`), estimated asymptotic bootstrap standard error (`s.e.`) as well as the confidence intervals for asymptotic diversity with q = 0, 1, and 2 (`qPD.LCL` and `qPD.UCL`), the reference times (`Reftime`) and the type of PD (`Type`). These statistics are computed only for q = 0, 1 and 2. More detailed information about asymptotic and observed diversity estimates for any order q between 0 and 2 can be obtained from function `ObsAsy3D()`. The output is shown below:

```
output_PD_abun$PDAsyEst
```

```
  Assemblage      qPD PD_obs PD_asy  s.e. qPD.LCL qPD.UCL Reftime   Type
1       Edge q = 0 PD 61.290 80.027 5.580  69.091  90.964     400 meanPD
2       Edge q = 1 PD  5.246  5.372 0.095   5.184   5.559     400 meanPD
3       Edge q = 2 PD  1.797  1.798 0.022   1.754   1.841     400 meanPD
4   Interior q = 0 PD 69.318 86.375 4.457  77.640  95.110     400 meanPD
5   Interior q = 1 PD  5.721  5.854 0.093   5.672   6.036     400 meanPD
6   Interior q = 2 PD  1.914  1.915 0.023   1.869   1.961     400 meanPD
```

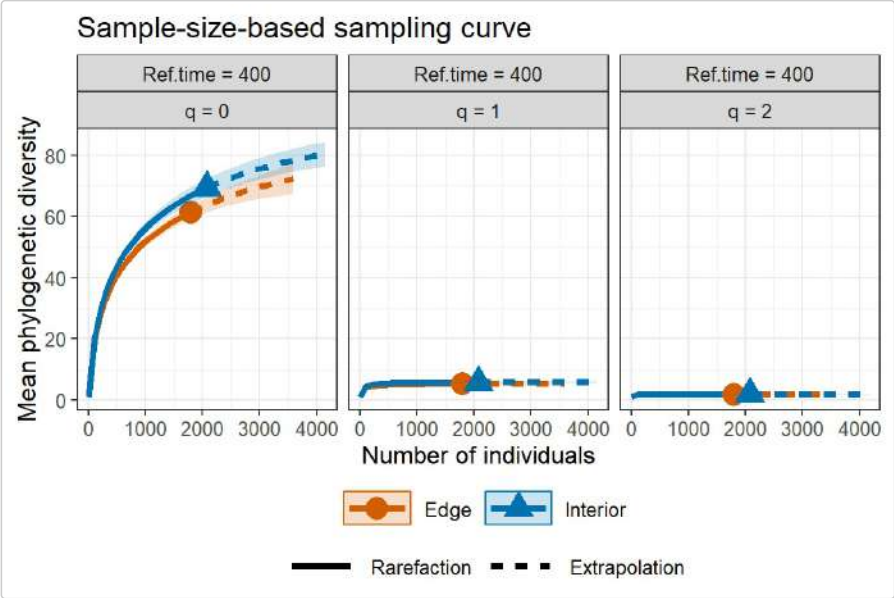
The `ggiNEXT3D` function can be used to make graphical displays for rarefaction and extrapolation sampling curves. When `facet.var = "Assemblage"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each assemblage; within each assemblage, different color curves represent diversity of different orders. An example for showing sample-size-based rarefaction/extrapolation curves (`type = 1`) is given below:

```
# PD sample-size-based R/E curves, separating by "Assemblage"
ggiNEXT3D(output_PD_abun, type = 1, facet.var = "Assemblage")
```



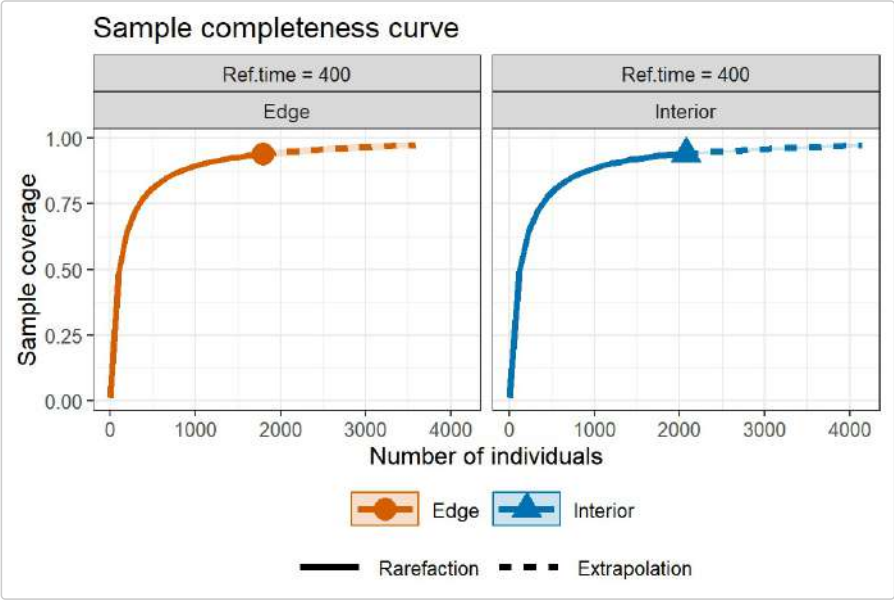
When `facet.var = "Order.q"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each diversity order; within each plot, different color curves represent different assemblages. An example is shown below:

```
# PD sample-size-based R/E curves, separating by "Order.q"
ggiNEXT3D(output_PD_abun, type = 1, facet.var = "Order.q")
```



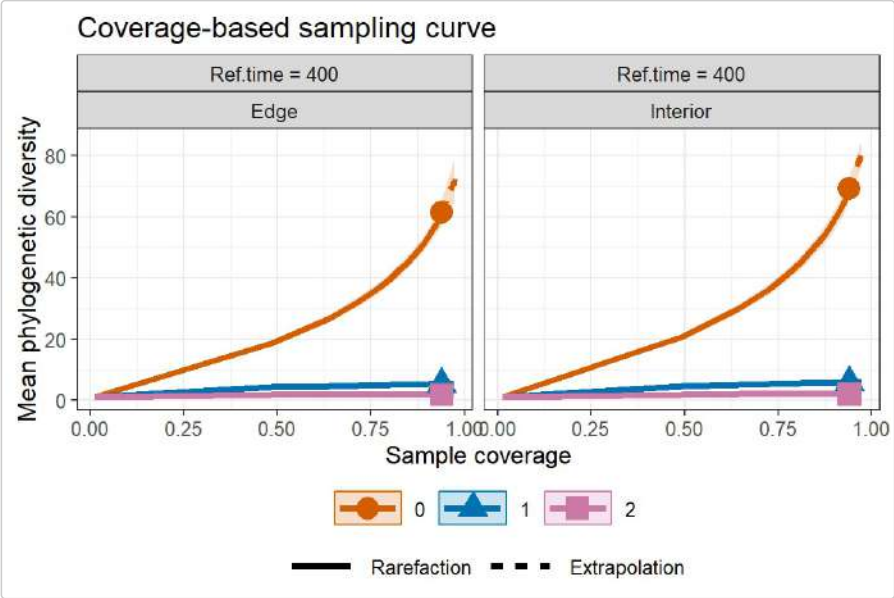
The following commands return the sample completeness (sample coverage) curve (`type = 2`) in which different colors are used for different assemblages.

```
# Sample completeness curves for abundance data, separating by "Assemblage"
ggiNEXT3D(output_PD_abun, type = 2, color.var = "Assemblage")
```

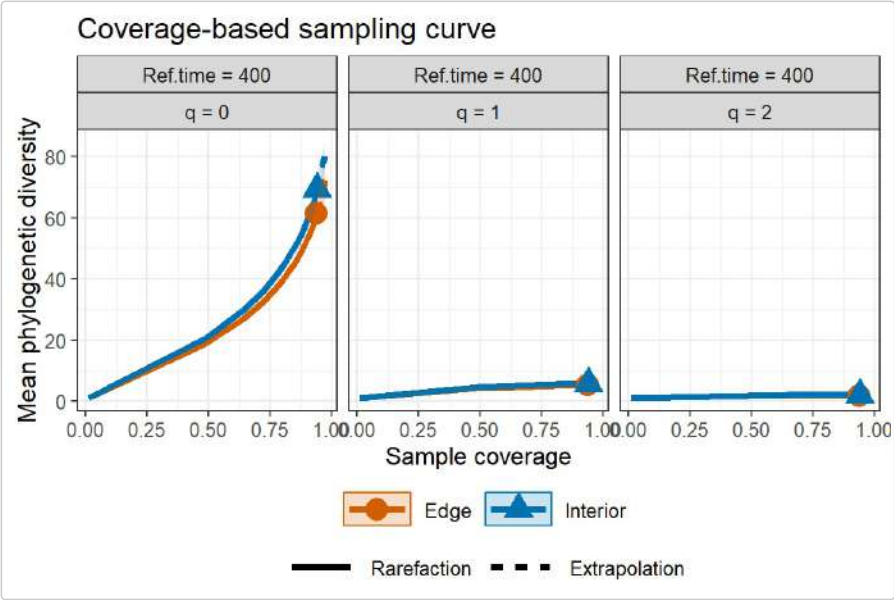


The following commands return the coverage-based rarefaction/extrapolation sampling curves in which different color curves represent three diversity orders within each assemblage (`facet.var = "Assemblage"`), or represent two assemblages within each diversity order (`facet.var = "Order.q"`), respectively.

```
# PD coverage-based R/E curves, separating by "Assemblage"
ggiNEXT3D(output_PD_abun, type = 3, facet.var = "Assemblage")
```




```
# PD coverage-based R/E curves, separating by "Order.q"
ggiNEXT3D(output_PD_abun, type = 3, facet.var = "Order.q")
```



EXAMPLE 4: PD rarefaction/extrapolation for incidence data

Based on the dataset (`Fish_incidence_data`) included in the package and the phylogentic tree (`Fish_phylo_tree`), the following commands return all numerical results for `PD`. The first list of the output (`$PDInfo`) returns basic data information including the name of the Assemblage, number of sampling units (`T`), total number of incidences (`U`), observed species richness (`S.obs`), sample coverage estimate of the reference sample with size T (`SC(T)`), sample coverage estimate of the extrapolated sample with size 2T (`SC(2T)`), the observed total branch length in the phylogenetic tree spanned by all observed species (`PD.obs`), the singletons/doubletons in the sample branch incidence (`Q1*,Q2*`), the total branch length of those singletons/doubletons in the sample branch incidence (`R1,R2`), and the reference time (`Reftime`). The output is identical to that based on the function `DataInfo3D()` by specifying `diversity = 'PD'` and `datatype = "incidence_raw"`; see later text). Thus, if only data information is required, the simpler function `DataInfo3D()` (see later text) can be used to obtain the same output. More information about the observed diversity (for any order q between 0 and 2) can be obtained by function `ObsAsy3D()`, which will be introduced later.

The required argument for performing PD analysis is `PDtree`. For example, the phylogenetic tree for all observed species (including species in both `"2013-2015"` and `"2016-2018"` time periods) is stored in `Fish_phylo_tree`. Then we enter the argument `PDtree = Fish_phylo_tree`. Two optional arguments are: `PDtype` and `PDreftime`. There are two options for `PDtype`: `"PD"` (effective total branch length) or `"meanPD"` (effective number of equally divergent lineages, `meanPD` = PD/tree depth). Default is `PDtype = "meanPD"`. `PDreftime` is a numerical value specifying a reference time for computing phylogenetic diversity. By default (`PDreftime = NULL`), the reference time is set to the tree depth, i.e., age of the root of the phylogenetic tree. Run the following code to perform PD analysis.

```
data(Fish_incidence_data)
data(Fish_phylo_tree)
data <- Fish_incidence_data
tree <- Fish_phylo_tree
```

```
output_PD_inci <- iNEXT3D(data, diversity = 'PD', q = c(0, 1, 2),
                          datatype = "incidence_raw", nboot = 20, PDtree = tree)
output_PD_inci$PDInfo
```

```
$PDInfo
# A tibble: 2 x 12
  Assemblage     T     U S.obs `SC(T)` `SC(2T)` PD.obs `Q1*` `Q2*`    R1    R2 Reftime
  <chr>      <int> <int> <int>   <dbl>    <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
1 2013-2015     36   532    50   0.98     0.993   9.62    11     7 0.69  1.23    0.977
2 2016-2018     36   522    53   0.976    0.989   9.44    13     6 0.368 0.345   0.977
```

The second list of the output (`$PDiNextEst`) includes size- and coverage-based standardized diversity estimates and related statistics computed for 40 knots by default (for example in the `"2013-2015"` time period, corresponding to the target number of sample units `mT` = 1, 2, 4, …, 34, 36, 37, 38, …, 72), which locates the reference sampling units at the mid-point of the selected knots. There are two data frames (`$size_based` and `$coverage_based`).

The first data frame (`$size_based`) includes the name of the Assemblage, diversity order (`Order.q`), the target number of sample units (`mT`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the target number of sample units `mT` is less than, equal to, or greater than the number of sampling units in the reference sample), the diversity estimate of order q (`qPD`), the lower and upper confidence limits of diversity (`qPD.LCL` and `qPD.UCL`) conditioning on the sample size, the corresponding sample coverage estimate (`SC`) along with the lower and upper confidence limits of sample coverage (`SC.LCL` and `SC.UCL`), the reference time (`Reftime`) and the type of PD (`Type`). These sample coverage estimates with confidence intervals are used for plotting the sample completeness curve. If the argument `nboot` is greater than zero, then a bootstrap method is applied to obtain the confidence intervals for the diversity and sample coverage estimates. Otherwise, all confidence intervals will not be computed. Here only the first six rows of the `$size_based` output are displayed:

```
output_PD_inci$PDiNextEst$size_based
```

```
  Assemblage Order.q mT      Method   qPD qPD.LCL qPD.UCL    SC SC.LCL SC.UCL Reftime   Type
1  2013-2015       0  1 Rarefaction 5.744   5.541   5.946 0.606  0.577  0.635   0.977 meanPD
2  2013-2015       0  2 Rarefaction 6.813   6.581   7.045 0.749  0.727  0.770   0.977 meanPD
3  2013-2015       0  4 Rarefaction 7.716   7.488   7.945 0.851  0.837  0.865   0.977 meanPD
4  2013-2015       0  6 Rarefaction 8.130   7.865   8.394 0.894  0.881  0.908   0.977 meanPD
5  2013-2015       0  8 Rarefaction 8.389   8.079   8.700 0.919  0.905  0.932   0.977 meanPD
6  2013-2015       0 10 Rarefaction 8.589   8.237   8.942 0.934  0.921  0.947   0.977 meanPD
```

The second data frame (`$coverage_based`) includes the name of assemblage, the diversity order (`Order.q`), the target sample coverage value (`SC`), the corresponding number of sample units (`mT`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the coverage `SC` is less than, equal to, or greater than the reference sample coverage), the diversity estimate of order q (`qPD`), the lower and upper confidence limits of diversity (`qPD.LCL` and `qPD.UCL`) conditioning on the target sample coverage value, the reference time (`Reftime`) and the type of PD (`Type`). Here only the first six rows of the `$coverage_based` output are displayed below: (Note for a fixed coverage value, the confidence interval in the `$coverage_based` table is wider than the corresponding interval in the `$size_based` table. This is because, for a given coverage value, the sample size needed to attain a fixed coverage value varies with bootstrap replication, leading to higher uncertainty on the resulting diversity estimate.)

```
output_PD_inci$PDiNextEst$coverage_based
```

```
  Assemblage Order.q    SC mT      Method   qPD qPD.LCL qPD.UCL Reftime   Type
1  2013-2015       0 0.606  1 Rarefaction 5.744   5.542   5.946   0.977 meanPD
2  2013-2015       0 0.749  2 Rarefaction 6.813   6.598   7.028   0.977 meanPD
3  2013-2015       0 0.851  4 Rarefaction 7.716   7.492   7.941   0.977 meanPD
4  2013-2015       0 0.894  6 Rarefaction 8.130   7.852   8.407   0.977 meanPD
5  2013-2015       0 0.919  8 Rarefaction 8.389   8.055   8.724   0.977 meanPD
6  2013-2015       0 0.934 10 Rarefaction 8.589   8.204   8.975   0.977 meanPD
```

The third list of the output (`$PDAsyEst`) includes the name of the Assemblage, PD (or meanPD) for q = 0, 1, and 2 (`qPD`), the observed diversity (`PD_obs`), asymptotic diversity estimate (`PD_asy`) and its estimated bootstrap standard error (`s.e.`), the confidence intervals for asymptotic diversity (`qPD.LCL` and `qPD.UCL`), the reference time (`Reftime`) and the type of PD (`Type`) . These statistics are computed only for q = 0, 1 and 2. More detailed information about asymptotic and observed diversity estimates for any order q between 0 and 2 can be obtained from function `ObsAsy3D()`. The output is shown below:

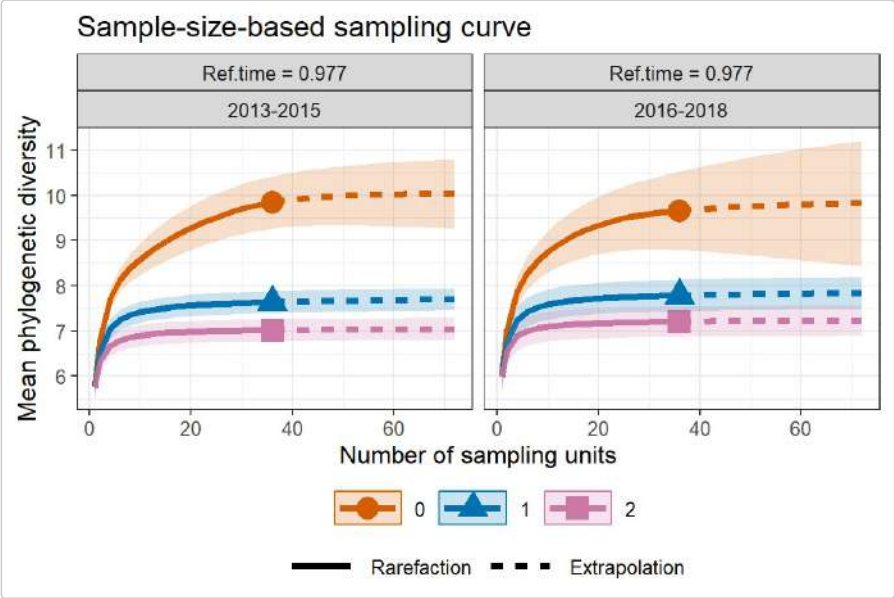
```
output_PD_inci$PDAsyEst
```

```
  Assemblage      qPD PD_obs PD_asy  s.e. qPD.LCL qPD.UCL Reftime   Type
1  2013-2015 q = 0 PD  9.847 10.039 0.702   8.663  11.416   0.977 meanPD
2  2013-2015 q = 1 PD  7.635  7.729 0.157   7.421   8.037   0.977 meanPD
3  2013-2015 q = 2 PD  7.013  7.057 0.152   6.760   7.355   0.977 meanPD
4  2016-2018 q = 0 PD  9.659  9.854 0.796   8.295  11.413   0.977 meanPD
```

```
5  2016-2018 q = 1 PD  7.781  7.859 0.141   7.583   8.136   0.977 meanPD
6  2016-2018 q = 2 PD  7.202  7.244 0.116   7.016   7.471   0.977 meanPD
```

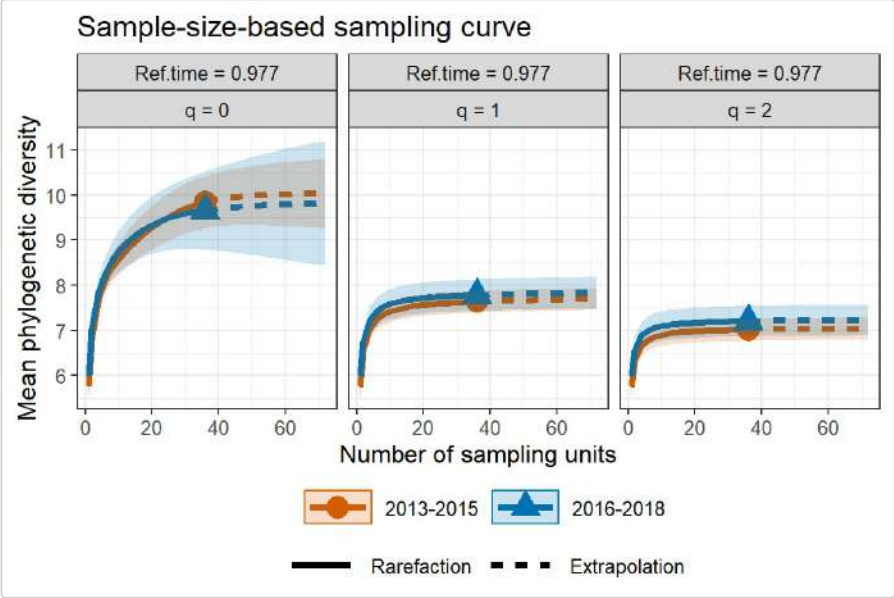
The `ggiNEXT3D` function can be used to make graphical displays for rarefaction and extrapolation sampling curves. When `facet.var = "Assemblage"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each assemblage; within each assemblage, different color curves represent diversity of different orders. An example for showing sample-size-based rarefaction/extrapolation curves (`type = 1`) is given below:

```
# PD sample-size-based R/E curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_PD_inci, type = 1, facet.var = "Assemblage")
```



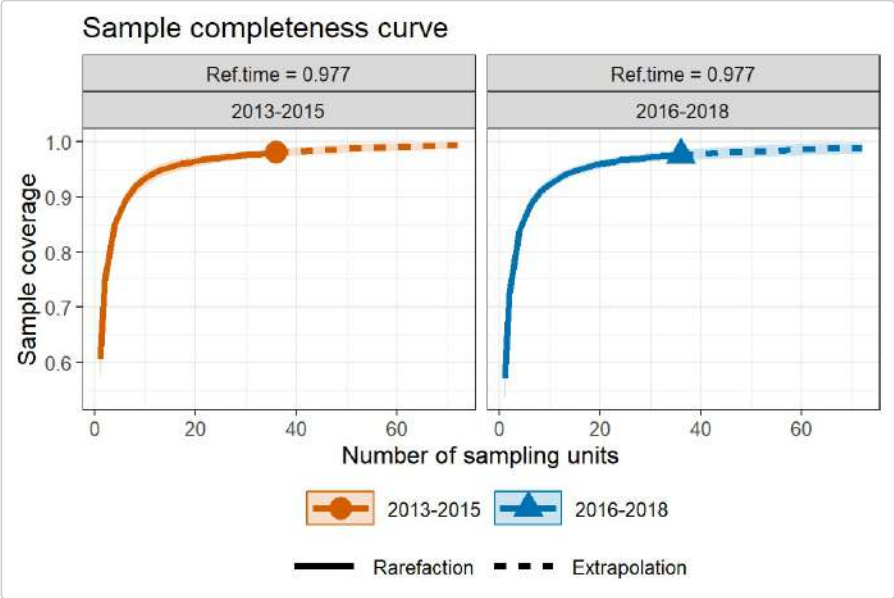
When `facet.var = "Order.q"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each diversity order; within each plot, different color curves represent different assemblages. An example is shown below:

```
# PD sample-size-based R/E curves for incidence data, separating by "Order.q"
ggiNEXT3D(output_PD_inci, type = 1, facet.var = "Order.q")
```



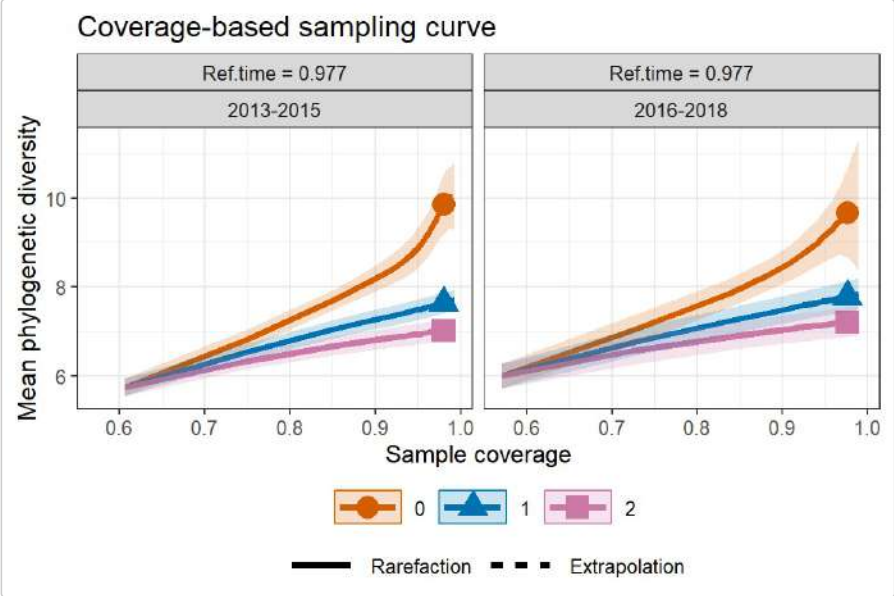
The following commands return the sample completeness (sample coverage) curve (`type = 2`) in which different colors are used for different assemblages.

```
# Sample completeness curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_PD_inci, type = 2, color.var = "Assemblage")
```

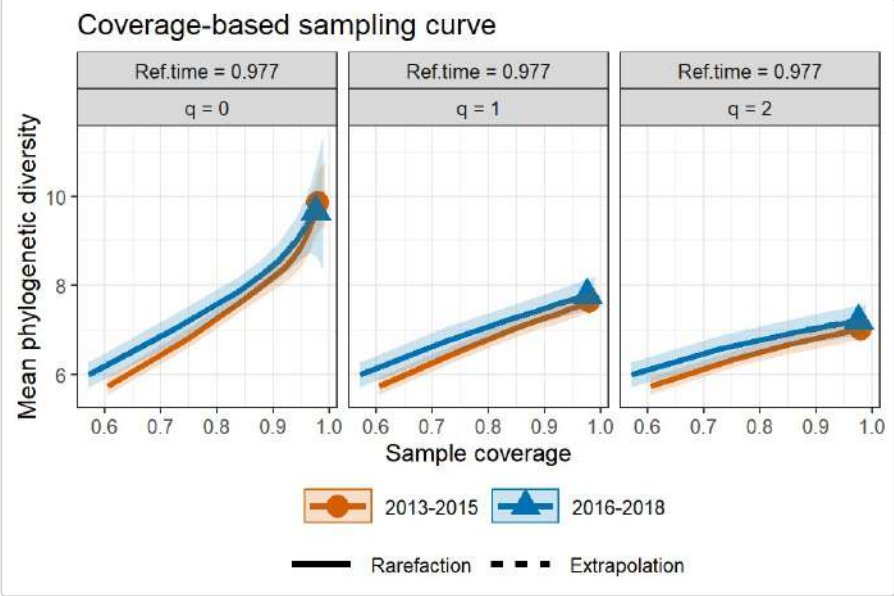


The following commands return the coverage-based rarefaction/extrapolation sampling curves in which different color curves represent three diversity orders within each assemblage (`facet.var = "Assemblage"`), or represent two assemblages within each diversity order (`facet.var = "Order.q"`), respectively.

```
# PD coverage-based R/E curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_PD_inci, type = 3, facet.var = "Assemblage")
```




```
# PD coverage-based R/E curves for incidence data, separating by "Order.q"
ggiNEXT3D(output_PD_inci, type = 3, facet.var = "Order.q")
```



FUNCTIONAL DIVERSITY (FD): RAREFACTION/EXTRAPOLATION VIA EXAMPLES

EXAMPLE 5: FD rarefaction/extrapolation for abundance data

Based on the dataset (`Brazil_rainforest_abun_data`) and the the distance matrix (`Brazil_rainforest_distance_matrix`) included in the package, the following commands return all numerical results for `FD`. The first list of the output (`$FDInfo`) returns basic data information including the name of the Assemblage, sample size (`n`), observed species richness (`S.obs`), sample coverage estimate of the reference sample with size n (`SC(n)`), sample coverage estimate of the extrapolated sample with size 2n (`SC(2n)`), and the minimum, mean, and maximum distance among all non-diagonal elements in the distance matrix(`dmin, dmean, dmax`). The output is identical to that based on the function `DataInfo3D()` by specifying `diversity = 'FD'` and `datatype = "abundance"`; see later text). Thus, if only data information is required, the simpler function `DataInfo3D()` (see later text) can be used to obtain the same output. More information about the observed diversity (for any order q between 0 and 2) can be obtained by function `ObsAsy3D()`, which will be introduced later.

The required argument for performing FD analysis is `FDdistM`. For example, the distance matrix for all species (including species in both Edge and Interior habitats) is stored in `Brazil_rainforest_distance_matrix`. Then we enter the argument `FDdistM = Brazil_rainforest_distance_matrix` Three optional arguments are (1) `FDtype`: `FDtype = "AUC"` means FD is computed from the area under the curve of a tau-profile by integrating all plausible threshold values between zero and one; `FDtype = "tau_values"` means FD is computed under specific threshold values to be specified in the argument `FD_tau`. (2) `FD_tau`: a numerical value specifying the tau value (threshold level) that will be used to compute FD. If `FDtype = "tau_values"` and `FD_tau = NULL`, then the threshold level is set to be the mean distance between any two individuals randomly selected from the pooled data over all data (i.e., quadratic entropy).

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_distance_matrix)
data <- Brazil_rainforest_abun_data
distM <- Brazil_rainforest_distance_matrix
output_FD_abun <- iNEXT3D(data, diversity = 'FD', datatype = "abundance", nboot = 10,
                          FDdistM = distM, FDtype = 'AUC')
output_FD_abun$FDInfo
```

```
$FDInfo
  Assemblage    n S.obs SC(n) SC(2n) dmin dmean  dmax
1       Edge 1794   319 0.939  0.974    0 0.372 0.776
2   Interior 2074   356 0.941  0.973    0 0.329 0.776
```

The second list of the output (`$FDiNextEst`) includes size- and coverage-based standardized diversity estimates and related statistics computed for 40 knots by default (for example in the “Edge” assemblage, corresponding to the target sample size `m` = 1, 95, 189, …, 1699, 1794, 1795, 1899, …, 3588), which locates the reference sample size at the mid-point of the selected knots. There are two data frames (`$size_based` and `$coverage_based`).

The first data frame (`$size_based`) includes the name of the Assemblage, diversity order (`Order.q`), the target sample size (`m`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the size `m` is less than, equal to, or greater than the reference sample size), the diversity estimate of order q (`qFD`), the lower and upper confidence limits of diversity (`qFD.LCL` and `qFD.UCL`) conditioning on the sample size, and the corresponding sample coverage estimate (`SC`) along with the lower and upper confidence limits of sample coverage (`SC.LCL` and `SC.UCL`). These sample coverage estimates with confidence intervals are used for plotting the sample completeness curve. If the argument `nboot` is greater than zero, then a bootstrap method is applied to obtain the confidence intervals for the diversity and sample coverage estimates. Otherwise, all confidence intervals will not be computed. Here only the first six rows of the `$size_based` output are displayed:

```
output_FD_abun$FDiNextEst$size_based
```

```
  Assemblage Order.q   m      Method    qFD qFD.LCL qFD.UCL    SC SC.LCL SC.UCL
1       Edge       0   1 Rarefaction  1.000   1.000   1.000 0.012  0.010  0.013
2       Edge       0  95 Rarefaction 10.900  10.442  11.358 0.484  0.466  0.502
3       Edge       0 189 Rarefaction 12.993  12.117  13.868 0.638  0.619  0.657
4       Edge       0 284 Rarefaction 14.129  12.888  15.371 0.718  0.702  0.735
5       Edge       0 378 Rarefaction 14.860  13.304  16.416 0.768  0.755  0.782
6       Edge       0 472 Rarefaction 15.383  13.549  17.216 0.803  0.792  0.814
```

The second data frame (`$coverage_based`) includes the name of assemblage, the diversity order (`Order.q`), the target sample coverage value (`SC`), the corresponding sample size (`m`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the coverage `SC` is less than, equal to, or greater than the reference sample coverage), the diversity estimate of order q (`qFD`), and the lower and upper confidence limits of diversity (`qFD.LCL` and `qFD.UCL`) conditioning on the target sample coverage value. Here only the first six rows of the `$coverage_based` output are displayed below: (Note for a fixed coverage value, the confidence interval in the `$coverage_based` table is wider than the corresponding interval in the `$size_based` table. This is because, for a given coverage value, the sample size needed to attain a fixed coverage value varies with bootstrap replication,

leading to higher uncertainty on the resulting diversity estimate.)

```
output_FD_abun$FDiNextEst$coverage_based
```

```
  Assemblage Order.q     SC   m      Method    qFD qFD.LCL qFD.UCL
1       Edge       0 0.012   1 Rarefaction  1.000   1.000   1.000
2       Edge       0 0.484  95 Rarefaction 10.900  10.472  11.328
3       Edge       0 0.638 189 Rarefaction 12.993  12.328  13.657
4       Edge       0 0.718 284 Rarefaction 14.129  13.209  15.049
5       Edge       0 0.768 378 Rarefaction 14.860  13.696  16.025
6       Edge       0 0.803 472 Rarefaction 15.383  13.991  16.775
```

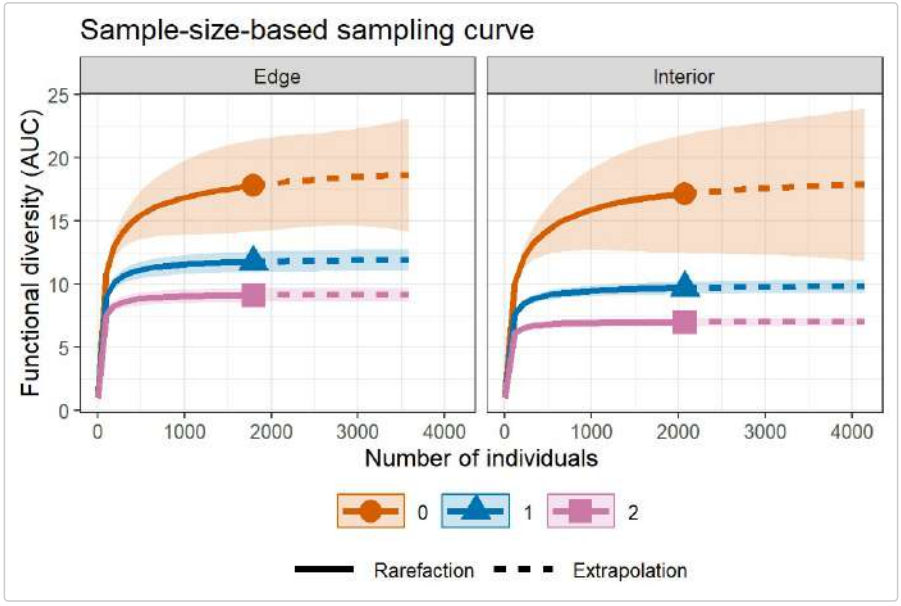
The third list of the output (`$FDAsyEst`) includes the name of the Assemblage, FD for q = 0, 1, and 2 (`qFD`), the observed diversity (`FD_obs`), asymptotic diversity estimate (`FD_asy`) and its estimated bootstrap standard error (`s.e.`) as well as the confidence intervals for asymptotic diversity (`qFD.LCL` and `qFD.UCL`). These statistics are computed only for q = 0, 1 and 2. More detailed information about asymptotic and observed diversity estimates for any order q between 0 and 2 can be obtained from function `ObsAsy3D()`. The output is shown below:

```
output_FD_abun$FDAsyEst
```

```
  Assemblage           qFD FD_obs FD_asy  s.e. qFD.LCL qFD.UCL
1       Edge q = 0 FD(AUC) 17.851 19.008 4.997   9.214  28.801
2       Edge q = 1 FD(AUC) 11.781 12.037 0.521  11.016  13.057
3       Edge q = 2 FD(AUC)  9.139  9.228 0.397   8.451  10.006
4   Interior q = 0 FD(AUC) 17.168 18.208 8.415   1.716  34.700
5   Interior q = 1 FD(AUC)  9.716  9.922 0.276   9.381  10.463
6   Interior q = 2 FD(AUC)  7.007  7.055 0.148   6.766   7.345
```

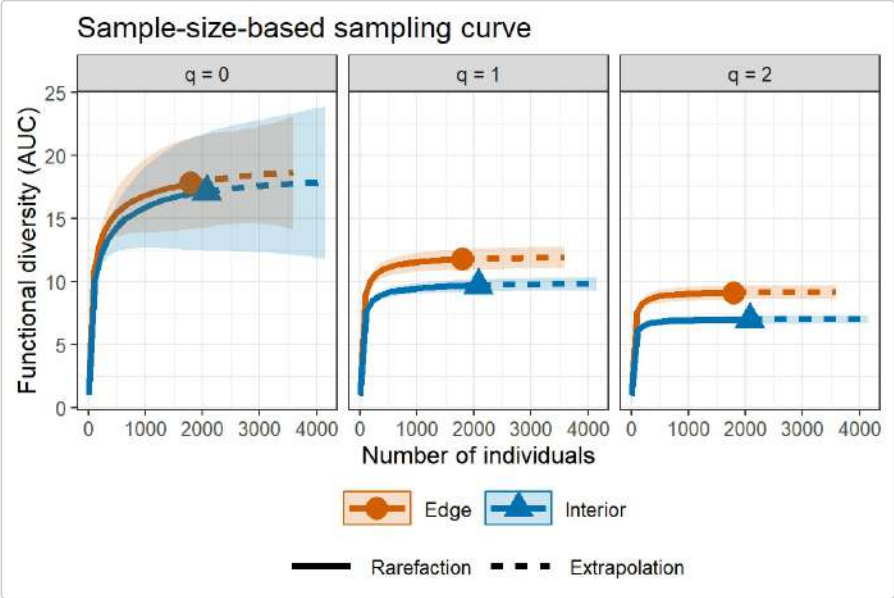
The `ggiNEXT3D` function can be used to make graphical displays for rarefaction and extrapolation sampling curves. When `facet.var = "Assemblage"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each assemblage; within each assemblage, different color curves represent diversity of different orders. An example for showing sample-size-based rarefaction/extrapolation curves (`type = 1`) is given below:

```
# FD sample-size-based R/E curves, separating by "Assemblage"
ggiNEXT3D(output_FD_abun, type = 1, facet.var = "Assemblage")
```



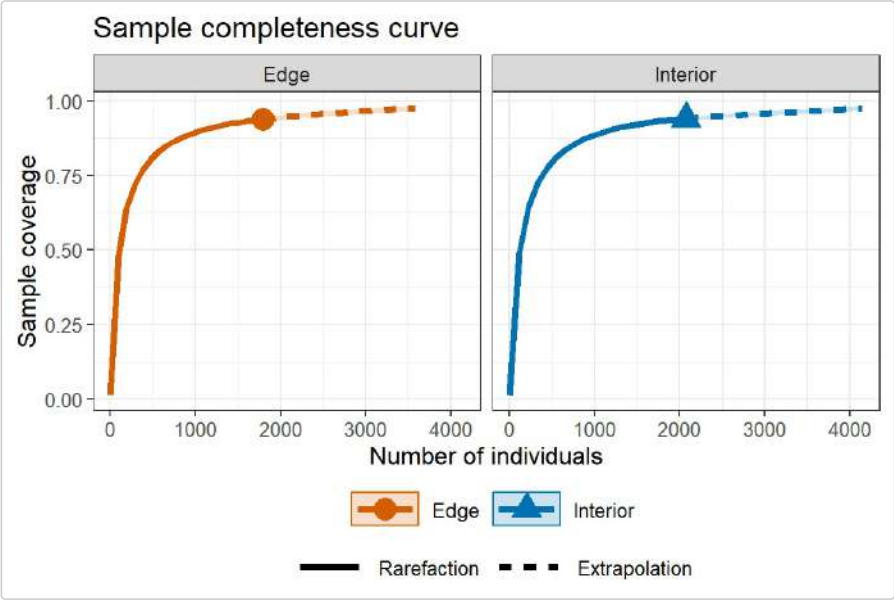
When `facet.var = "Order.q"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each diversity order; within each plot, different color curves represent different assemblages. An example is shown below:

```
# FD sample-size-based R/E curves, separating by "Order.q"
ggiNEXT3D(output_FD_abun, type = 1, facet.var = "Order.q")
```



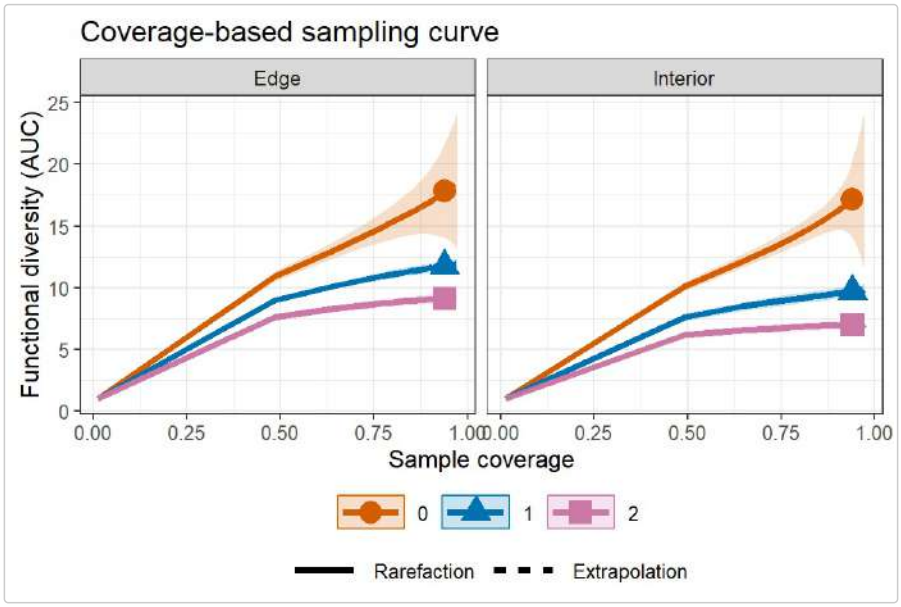
The following commands return the sample completeness (sample coverage) curve (`type = 2`) in which different colors are used for different assemblages.

```
# Sample completeness curves for abundance data, separating by "Assemblage"
ggiNEXT3D(output_FD_abun, type = 2, color.var = "Assemblage")
```

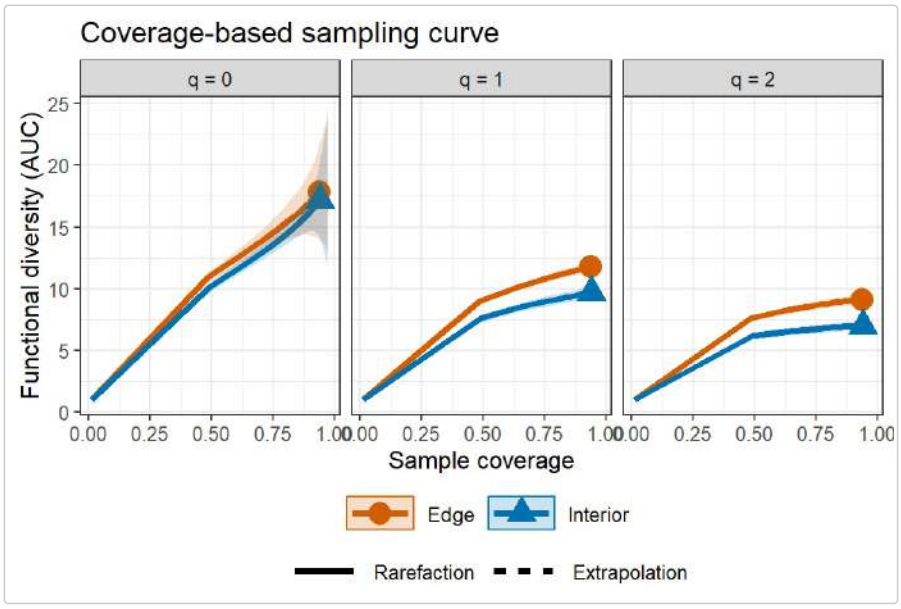


The following commands return the coverage-based rarefaction/extrapolation sampling curves in which different color curves represent three diversity orders within each assemblage (`facet.var = "Assemblage"`), or represent two assemblages within each diversity order (`facet.var = "Order.q"`), respectively.

```
# FD coverage-based R/E curves, separating by "Assemblage"
ggiNEXT3D(output_FD_abun, type = 3, facet.var = "Assemblage")
```




```
# FD coverage-based R/E curves, separating by "Order.q"
ggiNEXT3D(output_FD_abun, type = 3, facet.var = "Order.q")
```



EXAMPLE 6: FD rarefaction/extrapolation for incidence data

Based on the dataset (`Fish_incidence_data`) and the the distance matrix (`Fish_distance_matrix`) included in the package, the following commands return all numerical results for `FD`. The first list of the output (`$FDInfo`) returns basic data information including the name of the Assemblage, number of sampling units (`T`), total number of incidences (`U`), observed species richness (`S.obs`), sample coverage estimate of the reference sample with size T (`SC(T)`), sample coverage estimate of the reference sample with size 2T (`SC(2T)`), and the minimum, mean, and maximum distance among all non-diagonal elements in the distance matrix(`dmin, dmean, dmax`). The output is identical to that based on the function `DataInfo3D()` by specifying `diversity = 'FD'` and `datatype = "incidence_raw"`; see later text). Thus, if only data information is required, the simpler function `DataInfo3D()` (see later text) can be used to obtain the same output. More information about the observed diversity (for any order q between 0 and 2) can be obtained by function `ObsAsy3D()`, which will be introduced later.

The required argument for performing FD analysis is `FDdistM`. For example, the distance matrix for all species (including species in both `"2013-2015"` and `"2016-2018"` time periods) is stored in `Fish_distance_matrix`. Then we enter the argument `FDdistM = Fish_distance_matrix` Three optional arguments are (1) `FDtype`: `FDtype = "AUC"` means FD is computed from the area under the curve of a tau-profile by integrating all plausible threshold values between zero and one; `FDtype = "tau_values"` means FD is computed under specific threshold values to be specified in the argument FD_tau. (2) `FD_tau`: a numerical value specifying the tau value (threshold level) that will be used to compute FD. If `FDtype = "tau_values"` and `FD_tau = NULL`, then the threshold level is set to be the mean distance between any two individuals randomly selected from the pooled data over all data (i.e., quadratic entropy).

```
data(Fish_incidence_data)
data(Fish_distance_matrix)
data <- Fish_incidence_data
distM <- Fish_distance_matrix
output_FD_inci <- iNEXT3D(data, diversity = 'FD', datatype = "incidence_raw", nboot = 20,
```

```
                          FDdistM = distM, FDtype = 'AUC')
output_FD_inci$FDInfo
```

```
$FDInfo
  Assemblage  T   U S.obs SC(T) SC(2T)  dmin dmean  dmax
1  2013-2015 36 532    50 0.980  0.993 0.006 0.240 0.733
2  2016-2018 36 522    53 0.976  0.989 0.006 0.237 0.733
```

The second list of the output (`$FDiNextEst`) includes size- and coverage-based standardized diversity estimates and related statistics computed for 40 knots by default (for example in the `"2013-2015"` time period, corresponding to the target number of sample units `mT` = 1, 2, 4, …, 34, 36, 37, 38, …, 72), which locates the reference sampling units at the mid-point of the selected knots. There are two data frames (`$size_based` and `$coverage_based`).

The first data frame (`$size_based`) includes the name of the Assemblage, diversity order (`Order.q`), the target number of sample units (`mT`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the target number of sample units `mT` is less than, equal to, or greater than the number of sampling units in the reference sample), the diversity estimate of order q (`qFD`), the lower and upper confidence limits of diversity (`qFD.LCL` and `qFD.UCL`) conditioning on the sample size, and the corresponding sample coverage estimate (`SC`) along with the lower and upper confidence limits of sample coverage (`SC.LCL` and `SC.UCL`). These sample coverage estimates with confidence intervals are used for plotting the sample completeness curve. If the argument `nboot` is greater than zero, then a bootstrap method is applied to obtain the confidence intervals for the diversity and sample coverage estimates. Otherwise, all confidence intervals will not be computed. Here only the first six rows of the `$size_based` output are displayed:

```
output_FD_inci$FDiNextEst$size_based
```

```
  Assemblage Order.q mT      Method    qFD qFD.LCL qFD.UCL    SC SC.LCL SC.UCL
1  2013-2015       0  1 Rarefaction 14.778  13.862  15.694 0.606  0.575  0.637
2  2013-2015       0  2 Rarefaction 15.318  14.403  16.234 0.749  0.723  0.774
3  2013-2015       0  4 Rarefaction 15.888  14.972  16.803 0.851  0.832  0.869
4  2013-2015       0  6 Rarefaction 16.224  15.301  17.146 0.894  0.880  0.909
5  2013-2015       0  8 Rarefaction 16.463  15.530  17.396 0.919  0.906  0.931
6  2013-2015       0 10 Rarefaction 16.652  15.706  17.598 0.934  0.923  0.945
```

The second data frame (`$coverage_based`) includes the name of assemblage, the diversity order (`Order.q`), the target sample coverage value (`SC`), the corresponding number of sample units (`mT`), the `Method` (`Rarefaction`, `Observed`, or `Extrapolation`, depending on whether the coverage `SC` is less than, equal to, or greater than the reference sample coverage), the diversity estimate of order q (`qFD`), and the lower and upper confidence limits of diversity (`qFD.LCL` and `qFD.UCL`) conditioning on the target sample coverage value. Here only the first six rows of the `$coverage_based` output are displayed below: (Note for a fixed coverage value, the confidence interval in the `$coverage_based` table is wider than the corresponding interval in the `$size_based` table. This is because, for a given coverage value, the sample size needed to attain a fixed coverage value varies with bootstrap replication, leading to higher uncertainty on the resulting diversity estimate.)

```
output_FD_inci$FDiNextEst$coverage_based
```

```
  Assemblage Order.q    SC mT      Method    qFD qFD.LCL qFD.UCL
1  2013-2015       0 0.606  1 Rarefaction 14.778  14.179  15.376
2  2013-2015       0 0.749  2 Rarefaction 15.318  14.741  15.896
3  2013-2015       0 0.851  4 Rarefaction 15.888  15.243  16.533
4  2013-2015       0 0.894  6 Rarefaction 16.224  15.515  16.932
5  2013-2015       0 0.919  8 Rarefaction 16.463  15.699  17.226
6  2013-2015       0 0.934 10 Rarefaction 16.652  15.843  17.461
```

The third list of the output (`$FDAsyEst`) includes the name of the Assemblage, FD for q = 0, 1, and 2 (`qFD`), the observed diversity (`FD_obs`), asymptotic diversity estimate (`FD_asy`) and its estimated bootstrap standard error (`s.e.`), and the confidence intervals for asymptotic diversity (`qFD.LCL` and `qFD.UCL`). These statistics are computed only for q = 0, 1 and 2. More detailed information about asymptotic and observed diversity estimates for any order q between 0 and 2 can be obtained from function `ObsAsy3D()`. The output is shown below:

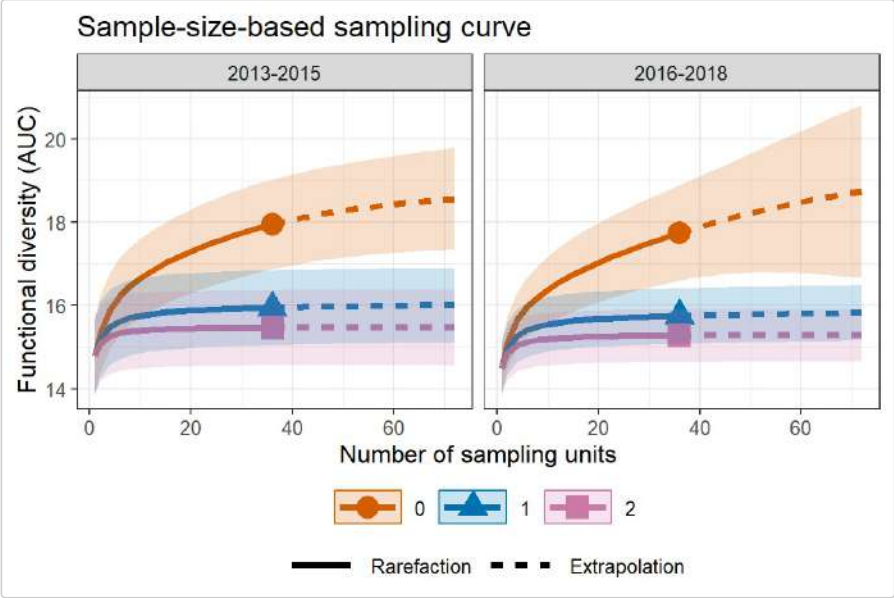
```
output_FD_inci$FDAsyEst
```

```
  Assemblage           qFD FD_obs FD_asy  s.e. qFD.LCL qFD.UCL
1  2013-2015 q = 0 FD(AUC) 17.904 18.906 1.386  16.188  21.623
2  2013-2015 q = 1 FD(AUC) 15.944 16.043 0.469  15.124  16.961
3  2013-2015 q = 2 FD(AUC) 15.463 15.490 0.455  14.598  16.383
4  2016-2018 q = 0 FD(AUC) 17.739 19.770 4.931  10.106  29.434
5  2016-2018 q = 1 FD(AUC) 15.749 15.867 0.607  14.678  17.056
6  2016-2018 q = 2 FD(AUC) 15.275 15.305 0.532  14.262  16.348
```

The `ggiNEXT3D` function can be used to make graphical displays for rarefaction and extrapolation sampling curves. When `facet.var = "Assemblage"` is specified in the `ggiNEXT3D` function, it creates a separate plot for

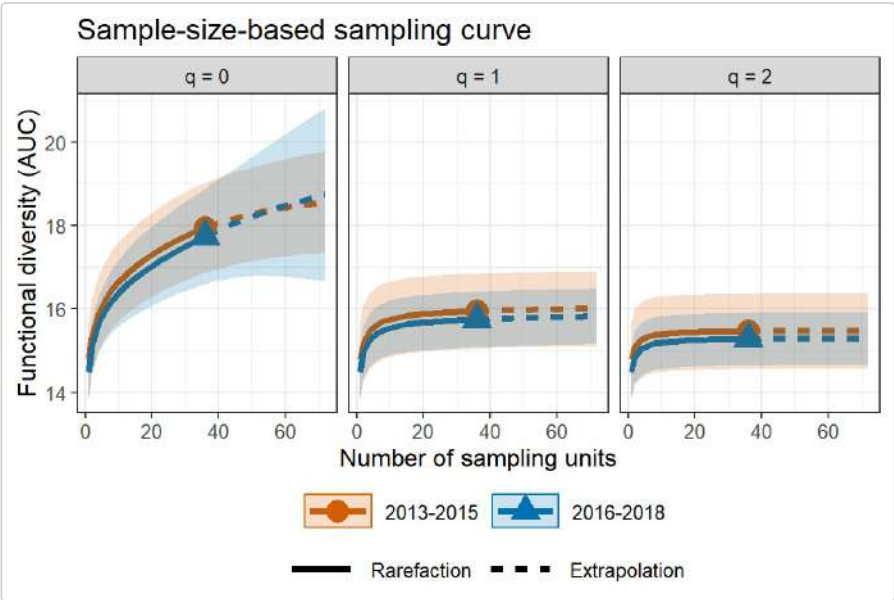
each assemblage; within each assemblage, different color curves represent diversity of different orders. An example for showing sample-size-based rarefaction/extrapolation curves (`type = 1`) is given below:

```
# FD sample-size-based R/E curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_FD_inci, type = 1, facet.var = "Assemblage")
```



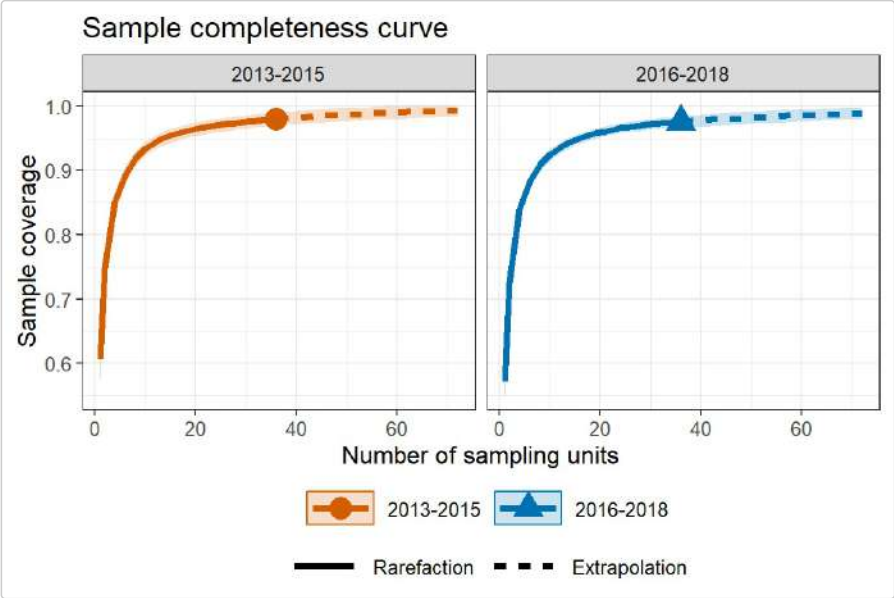
When `facet.var = "Order.q"` is specified in the `ggiNEXT3D` function, it creates a separate plot for each diversity order; within each plot, different color curves represent different assemblages. An example is shown below:

```
# FD sample-size-based R/E curves for incidence data, separating by "Order.q"
ggiNEXT3D(output_FD_inci, type = 1, facet.var = "Order.q")
```



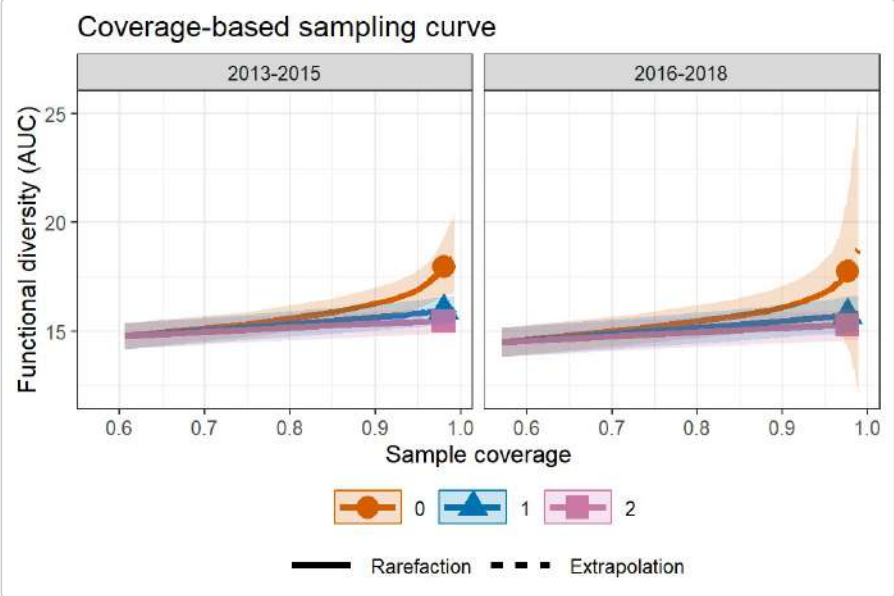
The following commands return the sample completeness (sample coverage) curve (`type = 2`) in which different colors are used for different assemblages.

```
# Sample completeness curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_FD_inci, type = 2, color.var = "Assemblage")
```

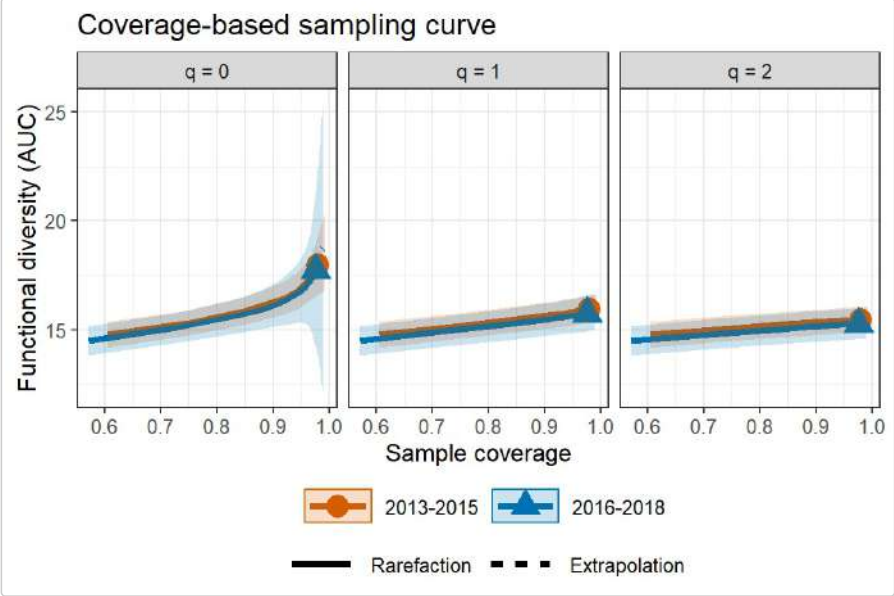


The following commands return the coverage-based rarefaction/extrapolation sampling curves in which different color curves represent three diversity orders within each assemblage (`facet.var = "Assemblage"`), or represent two assemblages within each diversity order (`facet.var = "Order.q"`), respectively.

```
# FD coverage-based R/E curves for incidence data, separating by "Assemblage"
ggiNEXT3D(output_FD_inci, type = 3, facet.var = "Assemblage")
```




```
# FD coverage-based R/E curves for incidence data, separating by "Order.q"
ggiNEXT3D(output_FD_inci, type = 3, facet.var = "Order.q")
```



FUNCTION DataInfo3D(): DATA INFORMATION

The function `DataInfo3D()` provides basic data information for the reference sample in each individual assemblage. The function `DataInfo3D()` with default arguments is shown below:

```
DataInfo3D(data, diversity = "TD", datatype = "abundance",
           nT = NULL, PDtree, PDreftime = NULL,
           FDdistM, FDtype = "AUC", FDtau = NULL)
```

All arguments in the above function are the same as those for the main function `iNEXT3D`. Running the `DataInfo3D()` function returns basic data information including sample size, observed species richness, two sample coverage estimates (`SC(n)` and `SC(2n)`) as well as other relevant information in each of the three dimensions of diversity. We use Brazil_rainforest_abun_data and Fish_incidence_data to demo the function for each dimension of diversity.

TAXONOMIC DIVERSITY (TD): Basic data information for abundance data

```
data(Brazil_rainforest_abun_data)
DataInfo3D(Brazil_rainforest_abun_data, diversity = 'TD', datatype = "abundance")
```

```
  Assemblage    n S.obs SC(n) SC(2n)  f1 f2 f3 f4 f5
1       Edge 1794   319 0.939  0.974 110 48 38 28 13
2   Interior 2074   356 0.941  0.973 123 48 41 32 19
```

Output description:

- `Assemblage` = assemblage name.
- `n` = number of observed individuals in the reference sample (sample size).
- `S.obs` = number of observed species in the reference sample.
- `SC(n)` = sample coverage estimate of the reference sample with size n.
- `SC(2n)` = sample coverage estimate of the reference sample with size 2n.
- `f1`-`f5` = the first five species abundance frequency counts in the reference sample.

TAXONOMIC DIVERSITY (TD): Basic data information for incidence data

```
data(Fish_incidence_data)
DataInfo3D(Fish_incidence_data, diversity = 'TD', datatype = "incidence_raw")
```

```
  Assemblage  T   U S.obs SC(T) SC(2T) Q1 Q2 Q3 Q4 Q5
1  2013-2015 36 532    50 0.980  0.993 11  6  4  1  3
2  2016-2018 36 522    53 0.976  0.989 13  5  5  2  3
```

Output description:

- `Assemblage` = assemblage name.
- `T` = number of sampling units in the reference sample (sample size for incidence data).
- `U` = total number of incidences in the reference sample.
- `S.obs` = number of observed species in the reference sample.
- `SC(T)` = sample coverage estimate of the reference sample with size T.
- `SC(2T)` = sample coverage estimate of the reference sample with size 2T.
- `Q1`-`Q5` = the first five species incidence frequency counts in the reference sample.

PHYLOGENETIC DIVERSITY (PD): Basic data information for abundance data

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_phylo_tree)
data <- Brazil_rainforest_abun_data
tree <- Brazil_rainforest_phylo_tree
DataInfo3D(data, diversity = 'PD', datatype = "abundance", PDtree = tree)
```

```
# A tibble: 2 x 11
  Assemblage     n S.obs `SC(n)` `SC(2n)` PD.obs `f1*` `f2*`    g1    g2 Reftime
  <chr>      <int> <int>   <dbl>    <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
1 Edge        1794   319   0.939    0.974  24516   110    52  6578  2885     400
2 Interior    2074   356   0.941    0.973  27727   123    56  7065  3656     400
```

Output description:

- `Assemblage`, `n`, `S.obs`, `SC(n)` and `SC(2n)`: definitions are the same as in the TD abundance output and thus are omitted.
- `PD.obs` = the observed total branch length in the phylogenetic tree spanned by all observed species.
- `f1*`,`f2*` = the number of singletons and doubletons in the node/branch abundance set.
- `g1`,`g2` = the total branch length of those singletons/doubletons in the node/branch abundance set.
- `Reftime` = reference time for phylogenetic diversity (the age of the root of phylogenetic tree).

PHYLOGENETIC DIVERSITY (PD): Basic data information for incidence data

```
data(Fish_incidence_data)
data(Fish_phylo_tree)
data <- Fish_incidence_data
tree <- Fish_phylo_tree
DataInfo3D(data, diversity = 'PD', datatype = "incidence_raw", PDtree = tree)
```

```
# A tibble: 2 x 12
  Assemblage     T     U S.obs `SC(T)` `SC(2T)` PD.obs `Q1*` `Q2*`    R1    R2 Reftime
  <chr>      <int> <int> <int>   <dbl>    <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
1 2013-2015     36   532    50   0.98     0.993   9.62    11     7 0.69  1.23    0.977
2 2016-2018     36   522    53   0.976    0.989   9.44    13     6 0.368 0.345   0.977
```

Output description:

- `Assemblage`, `T`, `U`, `S.obs`, `SC(T)` and `SC(2T)`: definitions are the same as in the TD incidence output and thus are omitted.
- `PD.obs` = the observed total branch length in the phylogenetic tree spanned by all observed species.
- `Q1*`,`Q2*` = the singletons/doubletons in the sample branch incidence.
- `R1`,`R2` = the total branch length of those singletons/doubletons in the sample branch incidence.
- `Reftime` = reference time.

FUNCTIONAL DIVERSITY (FD): Basic data information for abundance data

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_distance_matrix)
data <- Brazil_rainforest_abun_data
distM <- Brazil_rainforest_distance_matrix
DataInfo3D(data, diversity = 'FD', datatype = "abundance",
           FDdistM = distM, FDtype = 'AUC')
```

```
  Assemblage    n S.obs SC(n) SC(2n) dmin dmean  dmax
1       Edge 1794   319 0.939  0.974    0 0.372 0.776
2   Interior 2074   356 0.941  0.973    0 0.329 0.776
```

Output description:

- `Assemblage`, `n`, `S.obs`, `SC(n)` and `SC(2n)`: definitions are the same as in TD abundance output and thus are omitted.
- `dmin` = the minimum distance among all non-diagonal elements in the distance matrix.
- `dmean` = the mean distance between any two individuals randomly selected from each assemblage.
- `dmax` = the maximum distance among all elements in the distance matrix.

FUNCTIONAL DIVERSITY (FD): Basic data information for incidence data

```
data(Fish_incidence_data)
data(Fish_distance_matrix)
data <- Fish_incidence_data
distM <- Fish_distance_matrix
DataInfo3D(data, diversity = 'FD', datatype = "incidence_raw",
           FDdistM = distM, FDtype = 'AUC')
```

```
  Assemblage  T   U S.obs SC(T) SC(2T)  dmin dmean  dmax
1  2013-2015 36 532    50 0.980  0.993 0.006 0.240 0.733
2  2016-2018 36 522    53 0.976  0.989 0.006 0.237 0.733
```

Output description:

- `Assemblage`, `T`, `U`, `S.obs`, `SC(T)` and `SC(2T)`: definitions are the same as in the TD incidence output and

thus are omitted.

- `dmin` = the minimum distance among all non-diagonal elements in the distance matrix.
- `dmean` = the mean distance between any two individuals randomly selected from each assemblage.
- `dmax` = the maximum distance among all elements in the distance matrix.

FUNCTION estimate3D(): POINT ESTIMATION

`estimate3D` is used to compute 3D diversity (TD, PD, FD) estimates with q = 0, 1, 2 under any specified levels of sample size (when `base = "size"`) and sample coverage values (when `base = "coverage"`) for abundance data (`datatype = "abundance"`) or incidence data (`datatype = "incidence_raw"`). When `base = "size"`, `level` can be specified with a particular vector of sample sizes (greater than 0); if `level = NULL`, this function computes the diversity estimates for the minimum sample size among all samples extrapolated to the double reference sizes. When `base = "coverage"`, `level` can be specified with a particular vector of sample coverage values (between 0 and 1); if `level = NULL`, this function computes the diversity estimates for the minimum sample coverage among all samples extrapolated to the double reference sizes. All arguments in the function are the same as those for the main function `iNEXT3D`.

```
estimate3D(data, diversity = "TD", q = c(0, 1, 2), datatype = "abundance",
           base = "coverage", level = NULL, nboot = 50, conf = 0.95,
           nT = NULL, PDtree, PDreftime = NULL, PDtype = "meanPD",
           FDdistM, FDtype = "AUC", FDtau = NULL, FDcut_number = 50)
```

TAXONOMIC DIVERSITY (TD): point estimation

Example 7a: TD for abundance data with two target coverage values (93% and 97%)

The following commands return the TD estimates with two specified levels of sample coverage (93% and 97%) based on the `Brazil_rainforest_abun_data`.

```
data(Brazil_rainforest_abun_data)
output_est_TD_abun <- estimate3D(Brazil_rainforest_abun_data, diversity = 'TD', q = c(0,1,2),
                                 datatype = "abundance", base = "coverage", level = c(0.93,
        0.97))
output_est_TD_abun
```

```
   Assemblage Order.q   SC        m        Method     qTD   s.e. qTD.LCL qTD.UCL
1        Edge       0 0.93 1547.562   Rarefaction 302.879 12.456 278.465 327.293
2        Edge       0 0.97 3261.971 Extrapolation 383.307 18.571 346.909 419.705
3        Edge       1 0.93 1547.562   Rarefaction 152.374  4.504 143.547 161.202
4        Edge       1 0.97 3261.971 Extrapolation 166.837  4.992 157.052 176.622
5        Edge       2 0.93 1547.562   Rarefaction  81.437  3.760  74.069  88.806
6        Edge       2 0.97 3261.971 Extrapolation  83.726  3.953  75.978  91.474
7    Interior       0 0.93 1699.021   Rarefaction 331.917 12.276 307.858 355.977
8    Interior       0 0.97 3883.447 Extrapolation 433.807 18.549 397.452 470.162
9    Interior       1 0.93 1699.021   Rarefaction 159.330  4.855 149.814 168.847
10   Interior       1 0.97 3883.447 Extrapolation 175.739  5.128 165.689 185.790
11   Interior       2 0.93 1699.021   Rarefaction  71.611  3.922  63.924  79.297
12   Interior       2 0.97 3883.447 Extrapolation  73.326  4.068  65.353  81.299
```

Example 7b: TD for incidence data with two target coverage values (97.5% and 99%)

The following commands return the TD estimates with two specified levels of sample coverage (97.5% and 99%) for the `Fish_incidence_data`.

```
data(Fish_incidence_data)
output_est_TD_inci <- estimate3D(Fish_incidence_data, diversity = 'TD', q = c(0, 1, 2),
                                 datatype = "incidence_raw", base = "coverage",
                                 level = c(0.975, 0.99))
output_est_TD_inci
```

```
   Assemblage Order.q    SC     mT        Method    qTD   s.e. qTD.LCL qTD.UCL
1   2013-2015       0 0.975 29.169   Rarefaction 47.703  3.264  41.306  54.100
2   2013-2015       0 0.990 58.667 Extrapolation 54.914  4.665  45.771  64.057
3   2013-2015       1 0.975 29.169   Rarefaction 29.773  1.197  27.427  32.118
4   2013-2015       1 0.990 58.667 Extrapolation 30.751  1.214  28.372  33.130
5   2013-2015       2 0.975 29.169   Rarefaction 23.861  0.825  22.245  25.478
6   2013-2015       2 0.990 58.667 Extrapolation 24.126  0.840  22.479  25.773
7   2016-2018       0 0.975 34.825   Rarefaction 52.574  6.997  38.860  66.288
8   2016-2018       0 0.990 76.971 Extrapolation 62.688 14.646  33.983  91.393
```

```
9   2016-2018       1 0.975 34.825   Rarefaction 31.479  1.223  29.082  33.875
10  2016-2018       1 0.990 76.971 Extrapolation 32.721  1.186  30.397  35.046
11  2016-2018       2 0.975 34.825   Rarefaction 24.872  0.755  23.392  26.352
12  2016-2018       2 0.990 76.971 Extrapolation 25.163  0.743  23.708  26.618
```

PHYLOGENETIC DIVERSITY (PD): point estimation

Example 8a: PD for abundance data with two target sample sizes (1500 and 3500)

The following commands return the PD estimates with two specified levels of sample sizes (1500 and 3500) for the `Brazil_rainforest_abun_data`.

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_phylo_tree)
data <- Brazil_rainforest_abun_data
tree <- Brazil_rainforest_phylo_tree
output_est_PD_abun <- estimate3D(data, diversity = 'PD', datatype = "abundance",
                                 base = "size", level = c(1500, 3500), PDtree = tree)
output_est_PD_abun
```

```
   Assemblage Order.q    m        Method    SC    qPD  s.e. qPD.LCL qPD.UCL Reftime   Type
1        Edge       0 1500   Rarefaction 0.928 58.370 1.007  56.396  60.344      400 meanPD
2        Edge       0 3500 Extrapolation 0.973 71.893 2.233  67.516  76.270      400 meanPD
3        Edge       1 1500   Rarefaction 0.928  5.224 0.103   5.021   5.426      400 meanPD
4        Edge       1 3500 Extrapolation 0.973  5.320 0.105   5.115   5.526      400 meanPD
5        Edge       2 1500   Rarefaction 0.928  1.797 0.024   1.749   1.844      400 meanPD
6        Edge       2 3500 Extrapolation 0.973  1.797 0.024   1.749   1.845      400 meanPD
7    Interior       0 1500   Rarefaction 0.922 63.555 0.917  61.758  65.353      400 meanPD
8    Interior       0 3500 Extrapolation 0.965 78.004 1.749  74.576  81.431      400 meanPD
9    Interior       1 1500   Rarefaction 0.922  5.675 0.113   5.454   5.896      400 meanPD
10   Interior       1 3500 Extrapolation 0.965  5.784 0.114   5.560   6.008      400 meanPD
11   Interior       2 1500   Rarefaction 0.922  1.913 0.032   1.851   1.976      400 meanPD
12   Interior       2 3500 Extrapolation 0.965  1.914 0.032   1.852   1.977      400 meanPD
```

Example 8b: PD for incidence data with two target coverage values (97.5% and 99%)

The following commands return the PD estimates with two specified levels of sample coverage (97.5% and 99%) for the `Fish_incidence_data`.

```
data(Fish_incidence_data)
data(Fish_phylo_tree)
data <- Fish_incidence_data
tree <- Fish_phylo_tree
output_est_PD_inci <- estimate3D(data, diversity = 'PD', datatype = "incidence_raw",
                                 base = "coverage", level = c(0.975, 0.99), PDtree = tree)
output_est_PD_inci
```

```
   Assemblage Order.q    SC     mT        Method    qPD  s.e. qPD.LCL qPD.UCL   Reftime   Type
1   2013-2015       0 0.975 29.169   Rarefaction  9.672 0.381   8.926  10.419 0.9770115 meanPD
2   2013-2015       0 0.990 58.667 Extrapolation 10.018 0.616   8.810  11.226 0.9770115 meanPD
3   2013-2015       1 0.975 29.169   Rarefaction  7.612 0.149   7.320   7.905 0.9770115 meanPD
4   2013-2015       1 0.990 58.667 Extrapolation  7.680 0.147   7.393   7.967 0.9770115 meanPD
5   2013-2015       2 0.975 29.169   Rarefaction  7.003 0.147   6.715   7.290 0.9770115 meanPD
6   2013-2015       2 0.990 58.667 Extrapolation  7.030 0.146   6.745   7.315 0.9770115 meanPD
7   2016-2018       0 0.975 34.825   Rarefaction  9.646 0.464   8.737  10.556 0.9770115 meanPD
8   2016-2018       0 0.990 76.971 Extrapolation  9.831 0.896   8.075  11.587 0.9770115 meanPD
9   2016-2018       1 0.975 34.825   Rarefaction  7.779 0.130   7.524   8.033 0.9770115 meanPD
10  2016-2018       1 0.990 76.971 Extrapolation  7.835 0.140   7.561   8.109 0.9770115 meanPD
11  2016-2018       2 0.975 34.825   Rarefaction  7.201 0.121   6.963   7.439 0.9770115 meanPD
12  2016-2018       2 0.990 76.971 Extrapolation  7.224 0.124   6.982   7.466 0.9770115 meanPD
```

FUNCTIONAL DIVERSITY (FD): point estimation

Example 9a: FD for abundance data with two target coverage values (93% and 97%)

The following commands return the FD estimates with two specified levels of sample coverage (93% and 97%) for the `Brazil_rainforest_abun_data`.

```
data(Brazil_rainforest_abun_data)
```

```
data(Brazil_rainforest_distance_matrix)
data <- Brazil_rainforest_abun_data
distM <- Brazil_rainforest_distance_matrix
output_est_FD_abun <- estimate3D(data, diversity = 'FD', datatype = "abundance",
                                 base = "coverage", level = c(0.93, 0.97), nboot = 10,
                                 FDdistM = distM, FDtype = 'AUC')
output_est_FD_abun
```

```
   Assemblage Order.q   SC        m        Method    qFD  s.e. qFD.LCL qFD.UCL
1        Edge       0 0.93 1547.562   Rarefaction 17.590 2.069  13.534  21.645
2        Edge       0 0.97 3261.971 Extrapolation 18.578 2.740  13.207  23.949
3        Edge       1 0.93 1547.562   Rarefaction 11.732 0.311  11.123  12.341
4        Edge       1 0.97 3261.971 Extrapolation 11.920 0.313  11.307  12.534
5        Edge       2 0.93 1547.562   Rarefaction  9.120 0.261   8.609   9.632
6        Edge       2 0.97 3261.971 Extrapolation  9.183 0.264   8.665   9.701
7    Interior       0 0.93 1699.021   Rarefaction 16.890 1.820  13.324  20.457
8    Interior       0 0.97 3883.447 Extrapolation 17.839 4.980   8.079  27.599
9    Interior       1 0.93 1699.021   Rarefaction  9.668 0.258   9.161  10.175
10   Interior       1 0.97 3883.447 Extrapolation  9.834 0.275   9.294  10.374
11   Interior       2 0.93 1699.021   Rarefaction  6.994 0.160   6.680   7.308
12   Interior       2 0.97 3883.447 Extrapolation  7.033 0.162   6.716   7.350
```

Example 9b: FD for incidence data with two target number of sampling units (30 and 70)

The following commands return the FD estimates with two specified levels of sample sizes (30 and 70) for the `Fish_incidence_data`.

```
data(Fish_incidence_data)
data(Fish_distance_matrix)
data <- Fish_incidence_data
distM <- Fish_distance_matrix
output_est_FD_inci <- estimate3D(data, diversity = 'FD', datatype = "incidence_raw",
                                 base = "size", level = c(30, 70), nboot = 10,
                                 FDdistM = distM, FDtype = 'AUC')
output_est_FD_inci
```

```
   Assemblage Order.q mT        Method    SC    qFD  s.e. qFD.LCL qFD.UCL
1   2013-2015       0 30   Rarefaction 0.976 17.748 0.519  16.730  18.766
2   2013-2015       0 70 Extrapolation 0.993 18.550 0.696  17.186  19.914
3   2013-2015       1 30   Rarefaction 0.976 15.929 0.314  15.314  16.545
4   2013-2015       1 70 Extrapolation 0.993 16.006 0.315  15.388  16.624
5   2013-2015       2 30   Rarefaction 0.976 15.459 0.277  14.915  16.003
6   2013-2015       2 70 Extrapolation 0.993 15.477 0.278  14.932  16.022
7   2016-2018       0 30   Rarefaction 0.972 17.503 0.562  16.401  18.606
8   2016-2018       0 70 Extrapolation 0.988 18.705 1.207  16.340  21.070
9   2016-2018       1 30   Rarefaction 0.972 15.729 0.371  15.001  16.457
10  2016-2018       1 70 Extrapolation 0.988 15.816 0.364  15.103  16.530
11  2016-2018       2 30   Rarefaction 0.972 15.268 0.386  14.512  16.025
12  2016-2018       2 70 Extrapolation 0.988 15.290 0.386  14.533  16.046
```

FUNCTION ObsAsy3D: ASYMPTOTIC AND OBSERVED DIVERSITY PROFILES

```
ObsAsy3D(data, diversity = "TD", q = seq(0, 2, 0.2), datatype = "abundance",
         nboot = 50, conf = 0.95, nT = NULL,
         method = c("Asymptotic", "Observed"),
         PDtree, PDreftime = NULL, PDtype = "meanPD",
         FDdistM, FDtype = "AUC", FDtau = NULL, FDcut_number = 50
         )
```

All arguments in the above function are the same as those for the main function `iNEXT3D` (except that the default of `q` here is `seq(0, 2, 0.2)`). The function `ObsAsy3D()` computes observed and asymptotic diversity of order q between 0 and 2 (in increments of 0.2) for 3D diversity; these 3D values with different order q can be used to depict a q-profile in the `ggObsAsy3D` function.

It also computes observed and asymptotic PD for various reference times by specifying the argument `PDreftime`; these PD values with different reference times can be used to depict a time-profile in the `ggObsAsy3D` function.

It also computes observed and asymptotic FD for various threshold tau levels by specifying the argument `FDtau`; these FD values with different threshold levels can be used to depict a tau-profile in the `ggObsAsy3D` function.

For each dimension, by default, both the observed and asymptotic diversity estimates will be computed.

FUNCTION ggObsAsy3D(): GRAPHIC DISPLAYS OF DIVERSITY PROFILES

```
ggObsAsy3D(output, profile = "q")
```

`ggObsAsy3D` is a ggplot2 extension for an `ObsAsy3D` object to plot 3D q-profile (which depicts the observed diversity and asymptotic diversity estimate with respect to order q) for q between 0 and 2 (in increments of 0.2).

It also plots time-profile (which depicts the observed and asymptotic estimate of PD or mean PD with respect to reference times when `diversity = "PD"` specified in the ObsAsy3D function), and tau-profile (which depicts the observed and asymptotic estimate of FD with respect to threshold level tau when `diversity = "FD"` and `FDtype = "tau_values"` specified in the `ObsAsy3D` function) based on the output from the function `ObsAsy3D`.

In the plot of profiles, only confidence intervals of the asymptotic diversity will be shown when both the observed and asymptotic diversity estimates are computed.

TAXONOMIC DIVERSITY (TD): q-profiles

Example 10a: TD q-profiles for abundance data

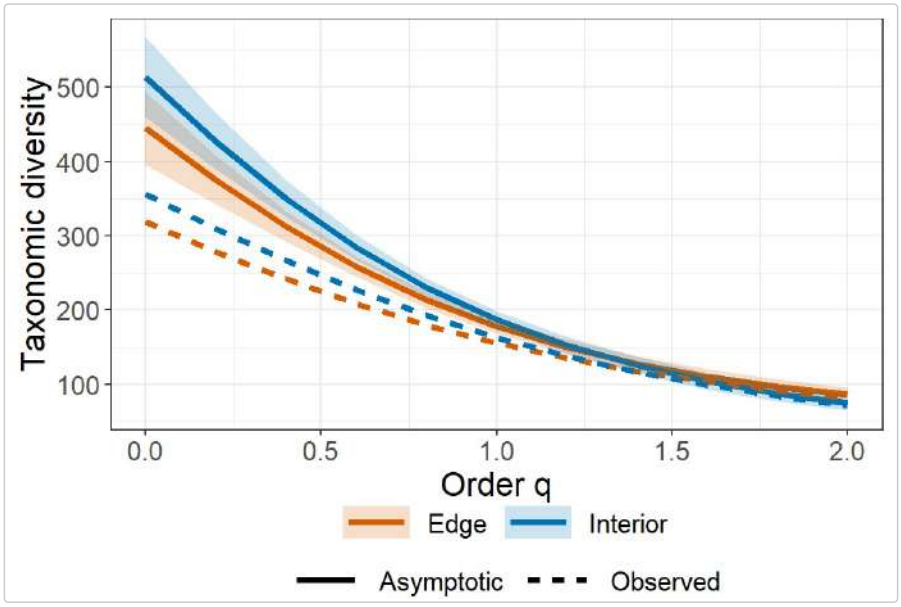
The following commands returns the observed and asymptotic taxonomic diversity ('TD') for the `Brazil_rainforest_abun_data`, along with its confidence interval for diversity order q between 0 to 2. Here only the first ten rows of the output are shown.

```
data(Brazil_rainforest_abun_data)
output_ObsAsy_TD_abun <- ObsAsy3D(Brazil_rainforest_abun_data, diversity = 'TD',
                                  datatype = "abundance")
output_ObsAsy_TD_abun
```

```
   Assemblage Order.q     qTD   s.e. qTD.LCL qTD.UCL     Method
1        Edge     0.0 444.971 25.175 395.629 494.314 Asymptotic
2        Edge     0.2 375.270 16.678 342.582 407.958 Asymptotic
3        Edge     0.4 312.452 10.496 291.880 333.024 Asymptotic
4        Edge     0.6 258.379  6.878 244.900 271.859 Asymptotic
5        Edge     0.8 213.730  5.445 203.057 224.403 Asymptotic
6        Edge     1.0 178.000  5.138 167.930 188.069 Asymptotic
7        Edge     1.2 149.914  5.123 139.874 159.955 Asymptotic
8        Edge     1.4 127.945  5.135 117.879 138.010 Asymptotic
9        Edge     1.6 110.672  5.139 100.599 120.745 Asymptotic
10       Edge     1.8  96.948  5.137  86.880 107.016 Asymptotic
```

The following commands plot the corresponding q-profiles, along with its confidence interval for q between 0 to 2.

```
# q-profile curves
ggObsAsy3D(output_ObsAsy_TD_abun)
```



Example 10b: TD q-profiles for incidence data

The following commands return the observed and asymptotic taxonomic diversity ('TD') estimates for the

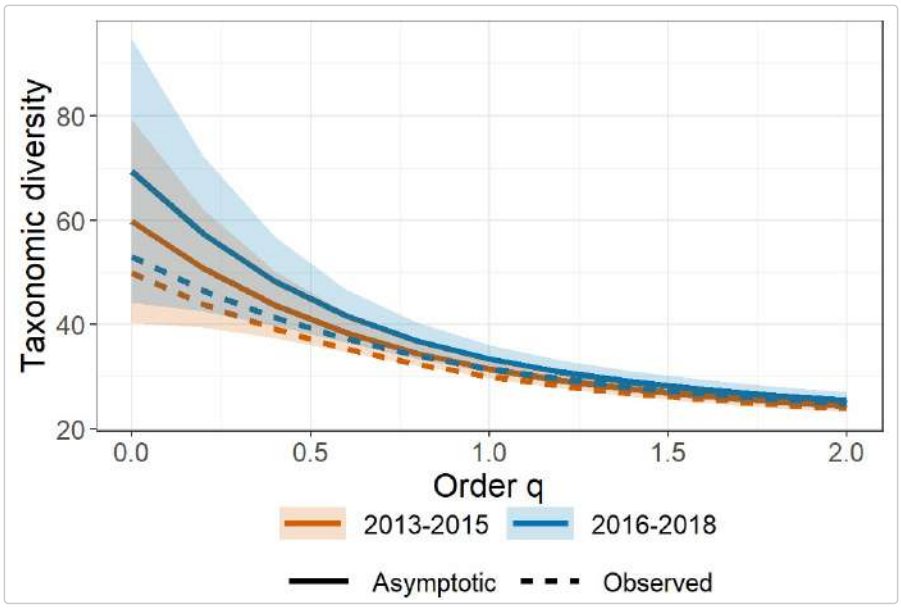
`Fish_incidence_data`, along with its confidence interval for diversity order q between 0 to 2. Here only the first ten rows of the output are shown.

```
data(Fish_incidence_data)
output_ObsAsy_TD_inci <- ObsAsy3D(Fish_incidence_data, diversity = 'TD',
                                  datatype = "incidence_raw")
output_ObsAsy_TD_inci
```

```
   Assemblage Order.q    qTD  s.e. qTD.LCL qTD.UCL     Method
1   2013-2015     0.0 59.803 9.908  40.384  79.223 Asymptotic
2   2013-2015     0.2 50.828 5.806  39.449  62.207 Asymptotic
3   2013-2015     0.4 43.790 3.281  37.359  50.221 Asymptotic
4   2013-2015     0.6 38.458 1.911  34.713  42.204 Asymptotic
5   2013-2015     0.8 34.490 1.248  32.044  36.936 Asymptotic
6   2013-2015     1.0 31.542 0.947  29.685  33.398 Asymptotic
7   2013-2015     1.2 29.328 0.803  27.754  30.902 Asymptotic
8   2013-2015     1.4 27.635 0.724  26.217  29.053 Asymptotic
9   2013-2015     1.6 26.312 0.673  24.992  27.632 Asymptotic
10  2013-2015     1.8 25.255 0.639  24.002  26.509 Asymptotic
```

The following commands plot the corresponding q-profiles, along with its confidence interval for q between 0 to 2.

```
# q-profile curves
ggObsAsy3D(output_ObsAsy_TD_inci)
```



PHYLOGENETIC DIVERSITY (PD): time-profiles and q-profiles

Example 11a: PD time-profiles for abundance data

The following commands return the observed and asymptotic phylogenetic diversity ('PD') estimates for the `Brazil_rainforest_abun_data`, along with its confidence interval for diversity order q = 0, 1, 2 under reference times from 0.01 to 400 (tree height). Here only the first ten rows of the output are shown.

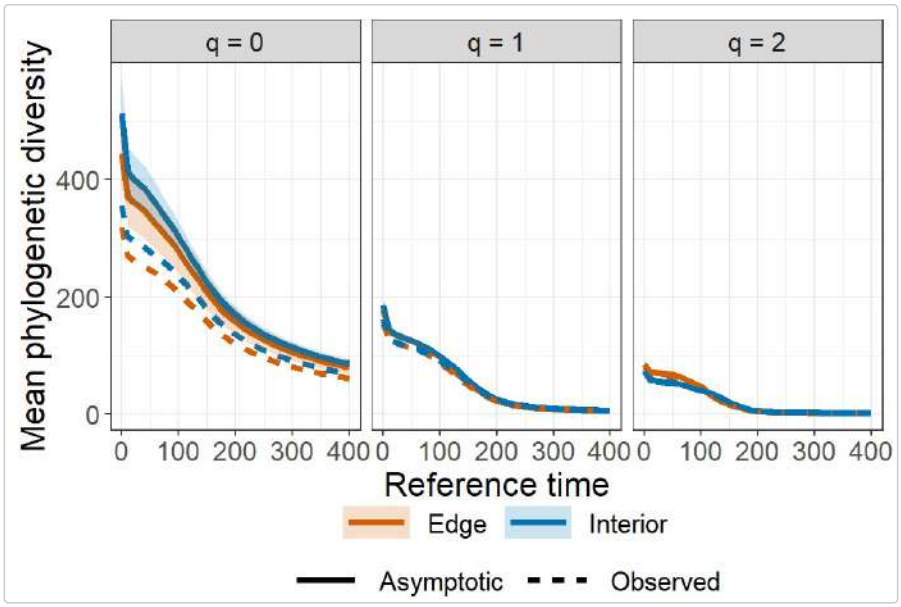
```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_phylo_tree)
data <- Brazil_rainforest_abun_data
tree <- Brazil_rainforest_phylo_tree
output_ObsAsy_PD_abun <- ObsAsy3D(data, diversity = 'PD', q = c(0, 1, 2),
                                  PDreftime = seq(0.01, 400, length.out = 20),
                                  datatype = "abundance", nboot = 20, PDtree = tree)
output_ObsAsy_PD_abun
```

```
   Assemblage Order.q     qPD   s.e. qPD.LCL qPD.UCL     Method Reftime   Type
1        Edge       0 444.971 29.001 388.130 501.812 Asymptotic   0.100 meanPD
2        Edge       1 178.000  5.074 168.055 187.944 Asymptotic   0.100 meanPD
3        Edge       2  85.905  4.149  77.773  94.038 Asymptotic   0.100 meanPD
4    Interior       0 513.518 29.215 456.256 570.779 Asymptotic   0.100 meanPD
5    Interior       1 186.983  5.190 176.812 197.154 Asymptotic   0.100 meanPD
6    Interior       2  74.718  4.210  66.466  82.969 Asymptotic   0.100 meanPD
7        Edge       0 371.100 25.520 321.082 421.117 Asymptotic  10.354 meanPD
8        Edge       1 141.418  3.841 133.891 148.946 Asymptotic  10.354 meanPD
```

```
9        Edge       2  72.848  3.260  66.458  79.238 Asymptotic  10.354 meanPD
10   Interior       0 413.568 22.401 369.663 457.472 Asymptotic  10.354 meanPD
```

The argument `profile = "time"` in the `ggObsAsy3D` function creates a separate plot for each diversity order q = 0, 1, and 2 with x-axis being “Reference time”. Different assemblages will be represented by different color lines.

```
# time-profile curves
ggObsAsy3D(output_ObsAsy_PD_abun, profile = "time")
```



Example 11b: PD q-profiles for incidence data

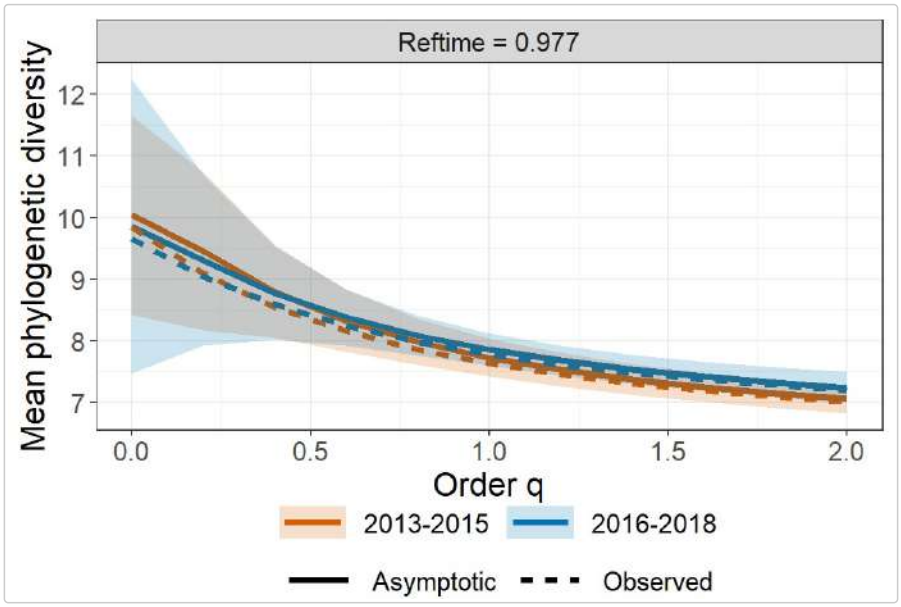
The following commands return the observed and asymptotic taxonomic diversity (‘PD’) estimates for the `Fish_incidence_data`, along with its confidence interval for diversity order q between 0 to 2. Here only the first ten rows of the output are shown.

```
data(Fish_incidence_data)
data(Fish_phylo_tree)
data <- Fish_incidence_data
tree <- Fish_phylo_tree
output_ObsAsy_PD_inci <- ObsAsy3D(data, diversity = 'PD', q = seq(0, 2, 0.2),
                                  datatype = "incidence_raw", nboot = 20, PDtree = tree,
                                  PDreftime = NULL)
output_ObsAsy_PD_inci
```

```
   Assemblage Order.q    qPD  s.e. qPD.LCL qPD.UCL     Method Reftime   Type
1   2013-2015     0.0 10.039 0.823   8.426  11.653 Asymptotic   0.977 meanPD
2   2013-2015     0.2  9.462 0.656   8.177  10.748 Asymptotic   0.977 meanPD
3   2013-2015     0.4  8.802 0.387   8.043   9.561 Asymptotic   0.977 meanPD
4   2013-2015     0.6  8.329 0.257   7.825   8.833 Asymptotic   0.977 meanPD
5   2013-2015     0.8  7.985 0.192   7.608   8.362 Asymptotic   0.977 meanPD
6   2013-2015     1.0  7.729 0.158   7.419   8.039 Asymptotic   0.977 meanPD
7   2013-2015     1.2  7.533 0.139   7.260   7.805 Asymptotic   0.977 meanPD
8   2013-2015     1.4  7.378 0.128   7.126   7.629 Asymptotic   0.977 meanPD
9   2013-2015     1.6  7.252 0.122   7.012   7.492 Asymptotic   0.977 meanPD
10  2013-2015     1.8  7.147 0.119   6.913   7.381 Asymptotic   0.977 meanPD
```

The following commands plot the corresponding q-profiles, along with its confidence interval for q between 0 to 2, for the default reference time = 0.977 (the tree depth).

```
# q-profile curves
ggObsAsy3D(output_ObsAsy_PD_inci, profile = "q")
```



FUNCTIONAL DIVERSITY (FD): tau-profiles and q-profiles

Example 12a: FD tau-profiles for abundance data

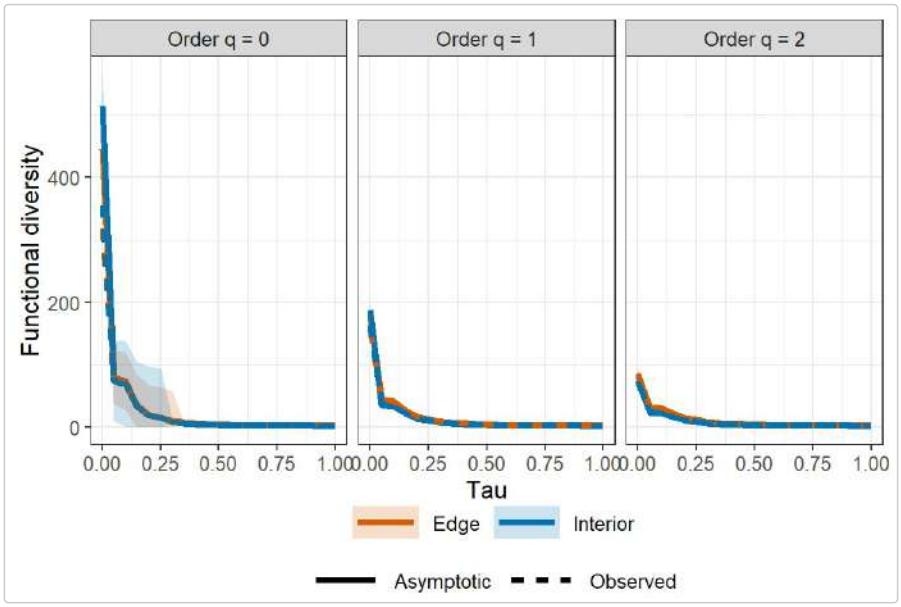
The following commands returns observed and asymptotic functional diversity ('FD') for `Brazil_rainforest_abun_data`, along with its confidence interval at diversity order q = 0, 1, 2 under tau values from 0 to 1. Here only the first ten rows of the output are shown.

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_distance_matrix)
data <- Brazil_rainforest_abun_data
distM <- Brazil_rainforest_distance_matrix
output_ObsAsy_FD_abun_tau <- ObsAsy3D(data, diversity = 'FD', q = c(0, 1, 2),
                                      datatype = "abundance", nboot = 10, FDdistM = distM,
                                      FDtype = 'tau_values', FDtau = seq(0, 1, 0.05))
output_ObsAsy_FD_abun_tau
```

```
   Assemblage Order.q     qFD   s.e. qFD.LCL qFD.UCL     Method  Tau
1        Edge       0 444.971 22.481 400.909 489.034 Asymptotic 0.00
2        Edge       1 178.000  5.377 167.461 188.538 Asymptotic 0.00
3        Edge       2  85.905  4.471  77.143  94.668 Asymptotic 0.00
4        Edge       0  79.904 22.161  36.468 123.340 Asymptotic 0.05
5        Edge       1  45.187  1.216  42.804  47.569 Asymptotic 0.05
6        Edge       2  32.092  0.799  30.526  33.658 Asymptotic 0.05
7        Edge       0  73.276 23.497  27.223 119.328 Asymptotic 0.10
8        Edge       1  42.200  1.137  39.972  44.427 Asymptotic 0.10
9        Edge       2  30.182  0.683  28.843  31.521 Asymptotic 0.10
10       Edge       0  35.372 24.511   0.000  83.413 Asymptotic 0.15
```

The following commands plot the corresponding tau-profiles, along with its confidence interval for diversity order q = 0, 1, 2.

```
# tau-profile curves
ggObsAsy3D(output_ObsAsy_FD_abun_tau, profile = "tau")
```



Example 12b: FD q-profiles for abundance data

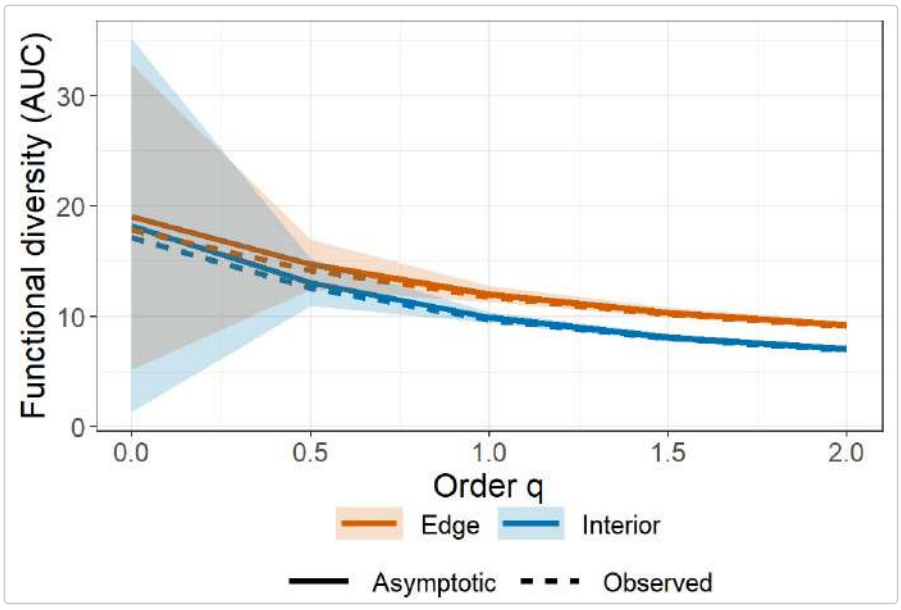
The following commands returns the observed and asymptotic taxonomic diversity ('FD') for the `Brazil_rainforest_abun_data`, along with its confidence interval for diversity order q between 0 to 2 with `FDtype = 'AUC'`. Here only the first ten rows of the output are shown.

```
data(Brazil_rainforest_abun_data)
data(Brazil_rainforest_distance_matrix)
data <- Brazil_rainforest_abun_data
distM <- Brazil_rainforest_distance_matrix
output_ObsAsy_FD_abun <- ObsAsy3D(data, diversity = 'FD', q = seq(0, 2, 0.5),
                                  datatype = "abundance", nboot = 10,
                                  FDdistM = distM, FDtype = 'AUC')
output_ObsAsy_FD_abun
```

```
   Assemblage Order.q    qFD  s.e. qFD.LCL qFD.UCL     Method
1        Edge     0.0 19.008 7.049   5.191  32.824 Asymptotic
2        Edge     0.5 14.698 1.144  12.456  16.941 Asymptotic
3        Edge     1.0 12.037 0.362  11.328  12.746 Asymptotic
4        Edge     1.5 10.345 0.233   9.889  10.802 Asymptotic
5        Edge     2.0  9.228 0.189   8.857   9.600 Asymptotic
6    Interior     0.0 18.208 8.615   1.322  35.094 Asymptotic
7    Interior     0.5 13.071 1.076  10.963  15.179 Asymptotic
8    Interior     1.0  9.922 0.249   9.434  10.410 Asymptotic
9    Interior     1.5  8.103 0.167   7.776   8.430 Asymptotic
10   Interior     2.0  7.055 0.143   6.776   7.335 Asymptotic
```

The following commands plot the corresponding q-profiles, along with its confidence interval for q between 0 to 2.

```
# q-profile curves
ggObsAsy3D(output_ObsAsy_FD_abun, profile = "q")
```



Example 12c: FD q-profiles for incidence data

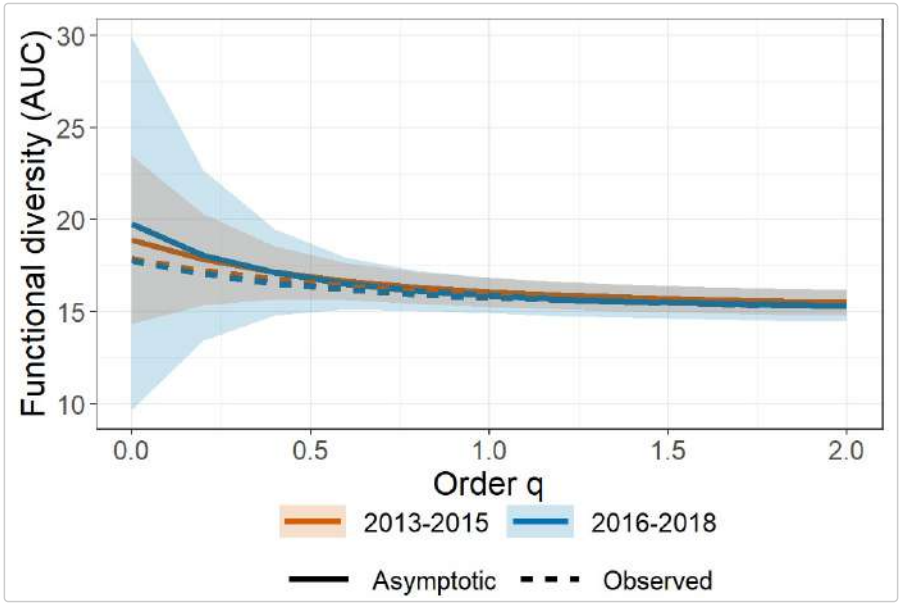
The following commands returns observed and asymptotic functional diversity ('FD') for `Fish_incidence_data`, along with its confidence interval at diversity order q from 0 to 2. Here only the first ten rows of the output are shown.

```
data(Fish_incidence_data)
data(Fish_distance_matrix)
data <- Fish_incidence_data
distM <- Fish_distance_matrix
output_ObsAsy_FD_inci <- ObsAsy3D(data, diversity = 'FD', datatype = "incidence_raw",
                                  nboot = 20, FDdistM = distM, FDtype = 'AUC')
output_ObsAsy_FD_inci
```

```
   Assemblage Order.q    qFD  s.e. qFD.LCL qFD.UCL     Method
1   2013-2015     0.0 18.906 2.329  14.341  23.470 Asymptotic
2   2013-2015     0.2 17.826 1.264  15.348  20.303 Asymptotic
3   2013-2015     0.4 17.115 0.736  15.673  18.557 Asymptotic
4   2013-2015     0.6 16.624 0.518  15.609  17.639 Asymptotic
5   2013-2015     0.8 16.284 0.435  15.430  17.137 Asymptotic
6   2013-2015     1.0 16.043 0.401  15.257  16.828 Asymptotic
7   2013-2015     1.2 15.868 0.383  15.117  16.618 Asymptotic
8   2013-2015     1.4 15.736 0.372  15.007  16.466 Asymptotic
9   2013-2015     1.6 15.635 0.365  14.919  16.351 Asymptotic
10  2013-2015     1.8 15.555 0.360  14.849  16.262 Asymptotic
```

The following commands plot the corresponding q-profiles, along with its confidence interval for q between 0 to 2.

```
# q-profile curves
ggObsAsy3D(output_ObsAsy_FD_inci, profile = "q")
```



License

The iNEXT.3D package is licensed under the GPLv3. To help refine `iNEXT.3D`, your comments or feedback would be welcome (please send them to Anne Chao or report an issue on the iNEXT.3D github iNEXT.3D_github.

References

- Chao, A., Henderson, P. A., Chiu, C.-H., Moyes, F., Hu, K.-H., Dornelas, M. and Magurran, A. E. (2021). Measuring temporal change in alpha diversity: a framework integrating taxonomic, phylogenetic and functional diversity and the iNEXT.3D standardization. Methods in Ecology and Evolution, 12, 1926-1940.
- Hsieh, T. C., Ma, K-H, and Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution, 7, 1451-1456.