Original version (July 2023)

A Brief Introduction to iNEXT.3D Online: Software for Interpolation and Extrapolation in three dimensions.

Anne Chao, K. H. Hu

Institute of Statistics, National Tsing Hua University, Hsin-Chu, Taiwan 30043

Overview

iNEXT.3D (iNterpolation and EXTrapolation in three Dimensions) Online is the R-based interactive online version of iNEXT.3D available via the link https://chao.shinyapps.io/iNEXT_3D/or http://chao.stat.nthu.edu.tw/wordpress/software_download/. Clicking these links, you will be directed to the online interface window. Users do not need to learn/understand R to run iNEXT.3D Online. The interactive web application was built using Shiny (a web application framework).

iNEXT.3D is the extension of shiny iNEXT Online (Chao and Jost 2012, 2015; Chao et al., 2014, 2019). It features two statistical analyses (non-asymptotic and asymptotic) for "Taxonomic diversity", "Phylogenetic diversity", and "Functional diversity" based on Hill numbers:

(1) A non-asymptotic approach based on interpolation and extrapolation
iNEXT.3D computes the estimated three kinds of diversities 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 numbers for q = 0, 1 and 2. iNEXT.3D offers three types of R/E sampling

curves for three kinds of diversities:

- 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 diversity

iNEXT.3D computes the estimated asymptotic diversity profiles including q-profile, time-profile, and tau-profile. Note that time profile for only when Phylogenetic diversity. Tau profile for only when Functional diversity and the FD type is "tau values".

How to cite

If you publish your work based on results from iNEXT.3D Online, you should make references to the following reference:

- 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.

Data

iNEXT.3D Online supports three types of data:

- Individual-based abundance data: Data for each assemblage/site include sample species abundances in an empirical sample of individuals (called a reference sample). When there are N assemblages and S species, input a txt file consist of a S by N matrix (in addition to the site labels and the species names). Input a txt file consist of species by assemblages matrix.

- Sampling-unit-based incidence data: There are two kinds of input data:

  (1) Incidence-raw data : For one assemblage, input a txt file for a reference sample consist of a species-by-sampling-unit matrix; when there are N assemblages, input N txt file consist of a species-by-sampling-unit matrix.

(2) Incidence-frequency data (for only when Taxonomic and Functional diversity): Input data for each assemblage consist of species sample incidence frequencies (row sums of each incidence raw matrix). When there are N assemblages and S species, input a txt file consist of a (S + 1) by N matrix (in addition to the site labels and the species names). The first row must be the total number of sampling units, followed by the species incidence frequencies.

See iNEXT.3D Online User's Guide for data input formats.

Running procedures for "iNEXT.3D" analysis

Species identification names are irrelevant in species/taxonomic diversity assessment so they are not essential for upload file (but necessary for phylogenetic and functional diversity). The following Steps 1, 2 and 3 are required procedures (selections are placed in blocks); Step 4 and 5 are optional.

Step 1. Select an analysis method (Taxonomic diversity or Phylogenetic diversity or Functional diversity) from the top menu of iNEXT.3D Online window.

Step 2. "Data Setting" (on the left hand side of the window screen)

(2a) Select Demo data to load demo data or Upload data to load your own data,

(2b) Select data type: For Taxonomic diversity and Functional diversity, we can select abundance data, incidence raw data or incidence frequency data; For Phylogenetic diversity, we can select abundance data or incidence raw data; see above for data input formats.

(2c) For abundance data, there are three demo data sets. Select one: Brazil rainforest, Beetles' data or Hinkley's fish abu. data.

(2d) Then the names/labels for the uploaded assemblages/sites will be automatically shown in the window below; you can select one set or multiple sets for comparison.

(2e) For Phylogenetic diversity and Functional diversity, there are one demo phylogenetic tree and distance matrix, respectively.

Step 3. Press the Run! button to get the output.

Step 4. Optional: "General setting" for statistical procedures. Only for panel "Rarefaction and Extrapolation" and "Asymptotic and Empirical Analysis".

For panel "Rarefaction and Extrapolation":

(4a) Optional: For Phylogenetic diversity, select the diversity type, including PD (effective total branch length) or meanPD (effective number of equally divergent lineages). (Default is meanPD); For Functional diversity, select the threshold type, including single threshold (under specified threshold) and AUC (for an overall FD which integrates all threshold values between zero and one). (Default of single threshold is d_{mean} (quadratic entropy: the mean distance between any two individuals randomly selected from the pooled assemblage)).

(4b) Optional: select a diversity order of q. (Default is q = 0).

(4c) Optional:choose specify endpoint or specify sample sizes. You can specify the endpoint of the extrapolation range or specify the sample sizes for which diversity estimates will be calculated. (Default endpoint = double of the minimum reference sample size, and the default number of knots = 40). If you choose to specify the sample sizes, then type in those sample sizes in the window space.

(4d) Optional: specify the number of bootstrap replications to compute s.e. and confidence intervals for each estimator. (Default bootstrap replications = 0). We type in "0" to skip all bootstrapping to save running time.

(4e) Optional: specify the level for the confidence interval (Default level is 0.95).

(4f) Optional: Only for Phylogenetic diversity, specify the reference time. (Default is tree height).

For panel "Asymptotic and Empirical Analysis":

(4a) Optional: For Phylogenetic diversity, select the profile type, including q profile or Time profile. (Default is q profile); For Functional diversity only for single threshold, select the profile type, including q profile and Tau profile. (Default is q profile)

(4b) Optional: For Phylogenetic diversity, select the diversity type, including PD (effective total branch length) or meanPD (effective number of equally divergent lineages). (Default is meanPD); For Functional diversity, select the threshold type, including single threshold (under specified threshold) and AUC (for an overall FD which integrates all threshold values between zero and one).

(Default of single threshold is d_{mean} (quadratic entropy: the mean distance between any two individuals randomly selected from the pooled assemblage)).

(4c) Optional: specify the number of bootstrap replications to compute s.e. and confidence intervals for each estimator. (Default bootstrap replications = 0). We type in "0" to skip all bootstrapping to save running time.

(4d) Optional: specify the level for the confidence interval (Default level is 0.95).

(4e) Optional: Only for Phylogenetic diversity, specify the reference time. (Default is tree height).

(4f) Optional: Type in diversity order q. (Default is 0, 0.25, 0.5, …, 2.75, 3) for q-profile; type in time when Time values profile; and type in tau values when Tau profile.

Step 5. Press the Run! button to get the output.

Note 1: We use a bootstrap resampling method to compute the s.e. and confidence interval of any estimator involved in the analysis. The default number of bootstrap replications is 0 to save running time. In this case, the computation of s.e. and confidence intervals will be skipped so that the output can be promptly shown..
You may specify a larger number (say, 50) to obtain more accurate results for

publication purposes, but it will take a longer time to get the output.

Note 2: The bootstrap resampling procedures vary with trial, meaning that two different runs for the same data may result in different s.e. estimates and different confidence intervals.

Output for "iNEXT.3D"

Along the second row (output) menu, there are five output selection tabs:

- In the "**Data Summary**" tab panel, basic data information, such as sample size, observed species richness and estimated sample coverage are shown for the reference sample under Taxonomic diversity panel. More details in iNEXT.3D Online UserGuide.

- In the "**Rarefaction and Extrapolation**" tab panel, various estimates for interpolated or extrapolated samples and their confidence intervals are displayed. These results can be downloaded by clicking "Download" at the bottom of the displayed figure. Furthermore, three types of sampling curves are shown:

  (1) Sample-size-based rarefaction and extrapolation sampling curve.
  (2) Sample completeness curve.
  (3) Coverage-based rarefaction and extrapolation sampling curve.

  These figures can be downloaded by clicking "Download" at the bottom of the displayed figures.

- In the "**Asymptotic and Empirical Analysis**" tab panel, the estimated asymptotic and empirical (observed) diversity profiles for diversity order are shown. Below the table, all numerical values of diversity are tabulated along with confidence intervals. Furthermore, the sampling curve is shown:

  - ✧ Asymptotic and empirical diversity profiles. (q profile/ time profile (Only for PD)/ tau profile (Only for FD with threshold type is under specified thresholds)).

  These figures can be downloaded by clicking "Download" at the bottom of the displayed figures.

- In the "**Introduction**" panel, users can view a brief introduction to iNEXT.3D Online and a summary of the running procedures.

- In the "**User Guide**" panel, a link will direct users to this user guide.

References

The following papers for pertinent background on rarefaction/extrapolation and related statistical analyses. These papers can be directly downloaded from Anne Chao's website.

Chao, A., Chiu, C.-H., Villéger, S., Sun, I.-F., Thorn, S., Lin, Y.-C., Chiang, J. M. and Sherwin, W. B. (2019). An attribute-diversity approach to functional diversity, functional beta diversity, and related (dis)similarity measures. *Ecological Monographs*, 89, e01343. 10.1002/ecm.1343.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K. and Ellison, A.M. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs*, **84**, 45–67.

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. & Jost, L. (2012) Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, **93**, 2533–2547.

Chao, A. and Jost, L. (2015). Estimating diversity and entropy profiles via discovery. rates of new species. *Methods in Ecology and Evolution*, **6**, 873–882.