dietr Tutorial

Samuel R. Borstein

04 September, 2019

1: Introduction

This is a tutorial for using the R package `dietr`. `dietr` uses diet or food item data to calculate trophic levels following the procedures described in `TrophLab` (Pauly et al., 2000), which currently is only avaialble as a Microsoft Access database program. Our implementation is very easy to use and extremely fast as users can specify all their data as dataframes. It also differs from TrophLab in that users can specify a taxonomic hierarchy and measure trophic levels from their data at various levels (e.x. individual,population, species,genus, etc.). `dietr` also works well with FishBase (Froese & Pauly, 2018) data and can use diet and food item data obtained in R using the rfishbase package (Boettiger et al., 2012). In addition to estimating trophic levels, `dietr` can also estimate various electivity indices used in diet studies.

For this tutorial we will refer to diet data as quantitative stomach content data where the proportion of prey items are known (e.x. percent volume or weight of prey items, be cautious using percent frequency as various literature suggest it may be misleading). In this case, trophic level is simply estimated by adding one to the sum of trophic levels of the prey items consumed weighted by their contribution to the diet. The trophic level of consumer i($Troph_i$) is defined by the equation below:

$$Troph_i = 1 + \sum_{j=1}^{G} DC_{ij} \times Troph_j$$

Here $Troph_j$ is the trophic level of the jth prey item consumed in the diet of i, DC_{ij} is the fraction of prey item j in the diet of i, and G is the number of prey species in the diet.

As estimates of prey trophic levels may not be exact, the standard error around the estimate of the trophic level can be calculated with equation 2. TrophLab gets around this by assigning a standard error for each trophic level and using the below equation to measure the standard error (which they also refer to as the omnivory index). Here, the standard error of the focal species i, $s.e._i$, is calculated as the square root of the sum of the product of the standard errors of the prey items $s.e._j$ to the second weighted by their respective contribution in the diet, DC_{ij} over 100.

$$s.e_i = \sqrt{\frac{\sum_{j=1}^{G} DC_{ij} * s.e._j^2}{100}}$$

For estimating trophic levels from food items found in the diet that don't have proportions, a random sampling and ranking of the food items is used to get an estimate of the trophic level. The simulated proportion of prey items for calculating trophic level, `P` is calculated using the following equation:

$$log_{10}P = 2 - 1.9log_{10}R - 0.16log_{10}G$$

Here, `R` is the rank of the food item and `G` is the number of food items, up to 10. If more than 10 food items are listed, then a subsample of ten are randomly selected. The trophic level is then calculated using the following equation:

$$Troph = \sum(P_i * Troph_i) / \sum P_i$$

Here, P_i is the simulated proportion of the prey item i in the diet and $Troph_i$ is the trophic level of prey item i. This procedure is repeated n times and the mean of these n simulations is taken as the trophic level. In cases where only a single prey item is in the diet, the procedure is much simpler and the estimated trophic level is simply calculated by adding 1 to the trophic level of the single prey item.

The estimate of the standard error around the trophic level from food item data is defined below. Here, the standard error is the square root of the sum of the product of the standard error of a prey item squared and the contribution of the prey item minus 1 divided by the sum of the contribution of the prey items, P, minus the total number of prey items, G. For food item data, the random sampling routine and calculation to estimate trophic level and standard error is repeated 100 times, with the final trophic level and standard error being the mean of these 100 calculations.

$$s.e_i = \sqrt{\frac{(s.e_1)^2 * (P_1 - 1) + (s.e._2)^2 * (P_2 - 1)...(s.e_G)^2 * (P_G - 1)}{100}}$$

In this tutorial we will cover the basics of how to use `dietr` to measure trophic levels. We will first discuss how to read in data from FishBase using rfishbase and how to pass that data into `dietr`. We will then show how one can input their own raw data and use the package to estimate trophic levels.

2: Installation

2.1: Installation From CRAN

In order to install the stable CRAN version of the `dietr` package:

```
install.packages("dietr")
```

2.2: Installation of Development Version From GitHub

While we recommend use of the stable CRAN version of this package, we recommend using the package `devtools` to temporarily install the development version of the package from GitHub if for any reason you wish to use it:

```
#1. Install 'devtools' if you do not already have it installed:
install.packages("devtools")

#2. Load the 'devtools' package and temporarily install the development version of
#'dietr' from GitHub:
library(devtools)
dev_mode(on=T)
install_github("sborstein/dietr")  # install the package from GitHub
library(dietr)# load the package

#3. Leave developers mode after using the development version of 'dietr' so it will not
#remain on your system permanently.
dev_mode(on=F)
```

3: Using dietr

To load `dietr` and all of its functions/data:

```
library(dietr)
```

3.0: Basic data necessary to run dietr

To run `dietr` you will need the following data as inputs. First is diet or food item data, which our functions call in DietTroph and FoodTroph as `DietItems` and `FoodItems` respectively. These should be organized from most inclusive to least inclusive from left to right and have column names.

The second is a dataframe of trophic levels of the prey, or as we name them in the functions, `PreyValues`. We include a few different datasets with prey values users can use, though you can also supply your own. `FishBasePreyVals` are the values FishBase/TrophLab use to calculate trophic levels. `CortesPreyVals` are standardized diet prey values for sharks from Cortes, 1999. Both of these can be loaded into R from `dietr` using the `data()` function:

```
data(FishBasePreyVals)#Load the Fishbase trophic levels of prey items
data(CortesPreyVals)#Load the Cortes (1992) standardized trophic levels of prey items for elasmobranchs
```

The last data object we need is a data frame we call `Taxonomy`. Columns of this dataframe should move from least inclusive to most inclusive from left to right. This data is used to assign individuals to groups for measuring hierarchical trophic levels (ex. trophic levels for an individuals, populations, species, etc.). This can be as simple as just a single column data frame if you only want to measure trophic levels for each individual and not at a higher level.

3.1: Using dietr to calculate trophic levels from FishBase diet data

Our first example will use FishBase diet data. For this, we will get data for two species: The Goliath Grouper, *Epinephelus itajara*, and the Schoolmaster snapper, *Lutjanus apodus*.

First, let us load rfishbase and use the diet function to obtain diet data.

This object contains a lot of info, most of which is metadata for the diet records. It is always good to look at this metadata to assess possible issues, but we will want to strip out the metadata leaving us with just the columns containing data we need to calculate trophic level. We can do this by using `dietr`'s `ConvertFishbaseDiet` function. This function has two arguments. `FishBaseDiet` is the data frame we obtained through `rfishbase`. ExcludeStage is an argument that excludes records of a certain life-history stage. For example, we may want to exclude larvae and immature fishes from our data. We could do this with the following code:

We can see that `cleaned.diets` object we created is a list containing two data frames. The first one, called DietItems, contains the information about diet composition, with `FoodI`, `FoodII`,`FoodIII`, `Stage` and `Volume` being the fields with descriptors for the various diets. In this case, each diet item has its own row in the data frame. The first column, `Individual`, contains the fish base diet reference number unique to that study. By including this, one can have multiple studies of the same species and then pool the data using the Taxonomy, which is the second data frame in our list. The `Species` column in the `DietItems` data frame contains the species name. For example, from the `cleaned.diets$DietItems` object we created above, we can see that *Lutjanus apodus* with the diet record number of 1337 ate three different items, bony fish, crabs, and other benthic crustaceans. If we look at `cleaned.diets$Taxonomy` we will see it is a data frame of two columns. For FishBase data, our function returns each individual diet study for a species in the column `Individual`. The next column has the respective species name. So, for this example, using this Taxonomy, we can expect the trophic level for *Lutjanus apodus* to be calculated for two studies, and then also return a single values for the species.

We can now measure the trophic level from diet items using the function `DietTroph`. Here, we will specify our `DietItems` and `Taxonomy` using the `cleaned.diets` object we generated above. We will also specify the `PreyValues` as being the included `FishBasePreyVals` that are part of the package. As the columns for the `PreyVals` and `DietItems` use a calssification of "FoodI","FoodII","FoodIII","Stage", we will specify these as a vector for the `PreyClass` argument.

```
data(FishBasePreyVals)#load the FishBase prey values that are part of the dietr package
#Calculate trophic level with DietTroph function
my.TL<-DietTroph(DietItems = converted.diet$DietItems,PreyValues = FishBasePreyVals, Taxonomy = convert
```

We can see that the `my.TL` object we just created contains a list of length two, each with a data frame, with the names of these data frames matching the `Taxonomy` we provided. We can see that the first data frame in this list, named `Individual` contains the trophic levels for each individual study that we input, while the data frame named `Species` provides the mean trophic level and SE and the number of studies across all individuals for each species.

3.2: Using dietr to calculate trophic levels from FishBase food item data

Our second example will estimate trophic level from food item data. The process is extremely similar to the above. For this example, let's get some data from rfishbase using the `fooditems` function for the same two species we did above.

```
#Get some food item data from rfishbase
my.food<-rfishbase::fooditems(c("Lutjanus apodus","Epinephelus itajara"))
```

In order to use this in `dietr`, we will need to convert it using the function `ConvertFishbaseFood`. This function is basically identical to the `ConvertFishbaseDiet` function we used above, except our data frame is from the `fooditems` function. We will then use `FoodTroph` function to calculate the trophic level. It is important to note that as this method randomly samples food items and ranks them for calculating the trophic level, that each time you run the function you will get a slightly different result, though they should largey be close in value.

```
#Convert FishBase food item data to a format usable for FoodTroph
converted.foods<-ConvertFishbaseFood(my.food)
#Calculate trophic level from food items
my.TL<-FoodTroph(FoodItems = converted.foods$FoodItems,PreyValues = FishBasePreyVals, Taxonomy = convert
```

3.3: Using dietr to calculate trophic levels from your own data

`dietr` was written with flexability in mind so it is easy for users to use their own data they collected to estimate trophic levels. For instance, users may want to analyze their own data or may be interested in using a different set of prey values.

3.4: Electivity Indices

While `dietr` can estimate trophic levels from food item and diet composition data as highlighted above, it can also measure a number of popular elevtivity indices used in studies of trophic ecology. The `dietr` function `electivity` implements Ivlev's (1961), Strauss' (1979), Jacob's Q and D (1974), Chesson's (1983)(Which is similar to Manl'y Alpha (1974)), and Vanderploeg & Scavia (1979) electivity indices.

4: Final Comments

Further information on the functions and their usage can be found in the helpfiles `help(package=dietr)`. For any further issues and questions send an email with subject 'dietr support' to sam@borstein.com or post to the issues section on GitHub(https://github.com/sborstein/dietr/issues).

5: References

Boettiger C, Lang DT, and Wainwright PC. 2012. rfishbase: exploring, manipulating and visualizing FishBase data from R. Journal of Fish Biology 81:2030-2039.

Chesson, J. 1983. The estimation and analysis of preference and its relatioship to foraging models. Ecology 64:1297-1304.

Cortes E. 1999. Standardized diet compositions and trophic levels of sharks. ICES Journal of marine science 56:707-717.

Froese R, and Pauly D. 2019. FishBase. http://www.fishbase.org/2019).

Ivlev, U. 1961. Experimental ecology of the feeding of fish. Yale University Press, New Haven.

Jacobs, J. 1974. Quantitative measurement of food selection. Oecologia 14:413-417.

Manly, B. 1974. A model for certain types of selection experiments. Biometrics 30:281-294.

Pauly D, Froese R, Sa-a P, Palomares M, Christensen V, and Rius J. 2000. TrophLab manual. ICLARM, Manila, Philippines.

Strauss, R. E. 1979. Reliability Estimates for Ivlev's Electivity Index, the Forage Ratio, and a Proposed Linear Index of Food Selection. Transactions of the American Fisheries Society 108:344-352.

Vanderploeg, H., and D. Scavia. 1979. Two electivity indices for feeding with special reference to zooplankton grazing. Journal of the Fisheries Board of Canada 36:362-365.