

# tRophicPosition two baselines, full Bayesian model example: Comparing bilagay TP across coastal kelp forests in N Chile

*Claudio Quezada-Romegialli & Chris Harrod*

*October 5 2017*

In this vignette, we show how to calculate trophic position with a two baselines, full Bayesian model. We will use a dataset gathered and compiled by Chris Harrod and his group (Docmac et al. 2017) from the Bilagay *Cheilodactylus variegatus*, an abundant marine fish species common to the coastal kelp forests in N Chile. This example is thoroughly covered in the vignette calculating trophic position for multiple species (see it using `browseVignettes("tRophicPosition")` code). So, here we will focus specifically on the calculation of trophic position and preparation of the plot included in our paper (citation of paper).

## Introduction

This dataset demonstrates how to calculate a trophic position for one species separated into different groups, in this case different locations, a common situation when the interest is to compare a species across space and/or time. Here we use data from a marine benthic rockfish, which, along with filter feeding and grazing molluscs (representing pelagic and benthic baselines respectively) were sampled from a series of different locations in the Northern Chilean coast.

We know from the literature that bilagay diet is extremely variable and reflects what is locally available (Pérez-Matus et al. 2012), whereas local variation in upwelling intensity (Figueroa & Moffat 2000) means that baseline  $\delta^{15}\text{N}$  values show considerable spatiotemporal variation (see Docmac et al. 2017) precluding direct comparisons of bilagay  $\delta^{15}\text{N}$  as a means of estimating TP. As such, we need to account for variation in baseline  $\delta^{15}\text{N}$  variation before calculating TP at each site. Also, we know that pelagic and benthic baselines are isotopically quite different in terms of their  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values, so we will use a two baselines Full model, that includes trophic discrimination for C.

In this case our question is, given the potential influences of local differences in diet and upwelling intensity, are there measurable differences in bilagay TP between the different locations along the S-N latitudinal gradient?

## Installing tRophicPosition

First of all, you need to install JAGS for your platform, and then install the stable version of `tRophicPosition` from CRAN:

```
install.packages("tRophicPosition")
```

After that, load the package with:

```
library(tRophicPosition)
```

```
## This is tRophicPosition 0.7.2
```

## Bilagay across N Chile - two baselines full model example

We start by loading the dataset into R. We have included the dataset for convenience:

```
data("Bilagay")
```

As usual, the first thing to do is to check its structure:

```
dplyr::tbl_df(Bilagay)
```

```
## # A tibble: 841 x 7
##   Study Location      Spp      FG d13C d15N  NS
##   <fctr> <fctr>      <fctr>    <fctr> <dbl> <dbl> <int>
## 1 MEL    CHI Echinolittorina peruviana Benthic_BL -10.1 16.3 4
## 2 MEL    CHI Echinolittorina peruviana Benthic_BL -11.7 16.3 4
## 3 MEL    CHI Echinolittorina peruviana Benthic_BL -9.8 16.4 4
## 4 MEL    CHI Echinolittorina peruviana Benthic_BL -11.0 16.6 4
## 5 MEL    CHI Echinolittorina peruviana Benthic_BL -7.4 12.5 4
## 6 MEL    CHI Echinolittorina peruviana Benthic_BL -10.1 14.2 4
## 7 MEL    CHI Echinolittorina peruviana Benthic_BL -10.9 16.2 4
## 8 MEL    CHI Echinolittorina peruviana Benthic_BL -11.8 16.9 4
## 9 MEL    CHI Echinolittorina peruviana Benthic_BL -8.7 15.2 4
## 10 MEL   CHI Echinolittorina peruviana Benthic_BL -10.0 16.3 4
## # ... with 831 more rows
```

Here we see that the dataset is organised as a number of columns:

- Study is a factor with 3 levels, each representing one Study;
- Location is a factor with 10 levels, each representing a different sampling location;
- Spp is a factor with 11 levels, each representing a species;
- FG is a factor with 3 levels, representing three different functional groups: Bilagay (fish species), Benthic\_BL (benthic baseline) and Pelagic\_BL (pelagic baseline);
- d13C and d15N are numerical variables, representing  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  isotope values respectively; and
- NS a numerical variable, representing latitudinal arrangement of sampling locations.

Now we will extract  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  isotope values for Bilagay, Benthic\_BL and Pelagic\_BL from each location, and combine them into a list:

```
# First we arrange the dataset according to location along a north-south  
# latitudinal gradient
```

```
Bilagay <- dplyr::arrange(Bilagay, NS)
```

```
# And then we extract isotope data
```

```
BilagayList <- extractIsotopeData(Bilagay,  
  speciesColumn = "FG",  
  b1 = "Pelagic_BL",  
  b2 = "Benthic_BL",  
  baselineColumn = "FG",  
  communityColumn = "Location",  
  deltaC = TDF(author = "McCutchan",  
    element = "C",  
    type = "muscle"),  
  deltaN = TDF(author = "McCutchan",  
    element = "N",  
    type = "muscle"))
```

```
## You selected McCutchan's et al (2003)
```

```
##           Muscle tissue d15N: 15 values with 2.9 mean +- 0.32 se
## You selected McCutchan's et al (2003)
##           Muscle tissue d13C: 18 values with 1.3 mean +- 0.3 se
```

Above you see that we used Pelagic\_BL functional group observations as baseline 1, and Benthic\_BL functional group observations as baseline 2, while using Location as the community grouping variable. Also, we used TDF(author = "McCutchan", element = "C", type = "muscle") and TDF(author = "McCutchan", element = "N", type = "muscle") for both deltaC and deltaN trophic discrimination factors (TDF). Using the latter approach to get TDF, we use the function on the fly to indicate what TDF values we are using for both elements.

We will check each dataset with plot() and summary(), to see if everything is fine:

```
for (location in BilagayList){
  plot(location)
  summary(location)
}
```

And now we calculate the trophic position for Bilagay in each location. We will use 20,000 iterations for the adaptive phase, 20,000 iterations as burnin (iterations discarded at the beginning of posterior sampling) and 20,000 actual iterations. Also we will use 5 chains to calculate trophic position, while using a two baselines, full Bayesian model.

We will save all results into Bilagay\_TPmodels. To make things faster, we will calculate trophic position with a parallel processing approach that maximises the efficient use of computing power available to the user:

```
# First we create a cluster with the cores available
cluster <- parallel::makePSOCKcluster(parallel::detectCores())

# Then we run the model in parallel, nested in system.time()
# in order to know how much time it takes to finish calculations
system.time(Bilagay_TPmodels <- parallel::parLapply(cluster,
  BilagayList,
  multiModelTP,
  adapt = 20000,
  n.iter = 20000,
  burnin = 20000,
  n.chains = 5,
  model = "twoBaselinesFull"))
```

```
##      user  system elapsed
##    0.01    0.01    78.62
```

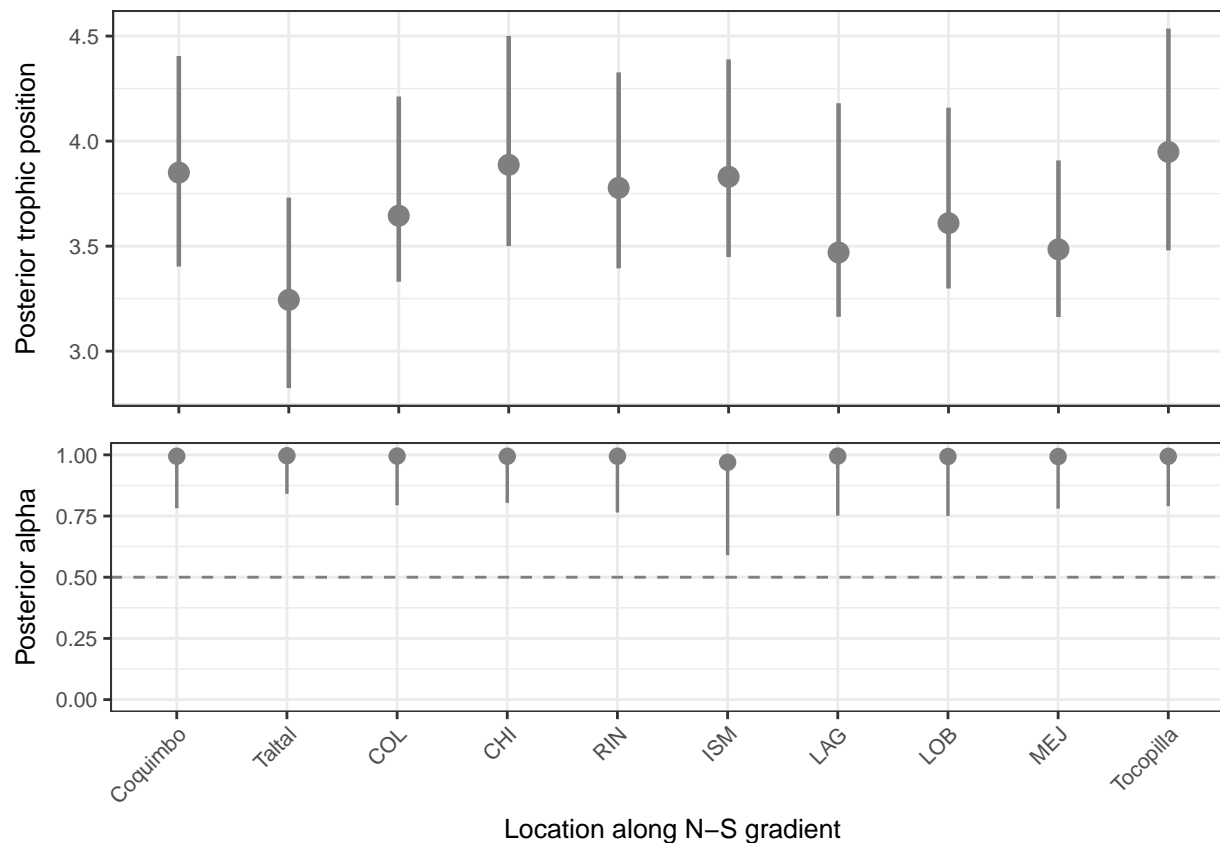
```
parallel::stopCluster(cluster)
```

Now we will prepare the plot to describe trophic position, alpha and the mean +- SD of baseline in each location that we used in our paper (citation).

First, we get a summary from parallel calculations of TP, and then we prepare the plot and save it into the variable TP\_alpha:

```
ggplot_df <- fromParallelTP(Bilagay_TPmodels, get = "summary")

TP_alpha <- credibilityIntervals(ggplot_df, x = "community", xlab = "Location along N-S gradient")
```



And now we prepare a plot summarising isotopic values of the pelagic baseline in these sampling locations:

```
# First we select only Bilagay isotope values
```

```
Bilagay_obs <- dplyr::filter(Bilagay, FG == "Bilagay")
```

```
## Warning: package 'bindrcpp' was built under R version 3.4.1
```

```
# Then we summarise the isotope data grouping by Location and ordering by NS
```

```
Bilagay_obs_summary <- summariseIsotopeData(df = Bilagay_obs,
                                             grouping = "Location",
                                             ordering = "NS" )
```

```
# And then we prepare a modified credibilityIntervals plot,
```

```
# using mean_d15N +/- SD by Location, and save it into the
```

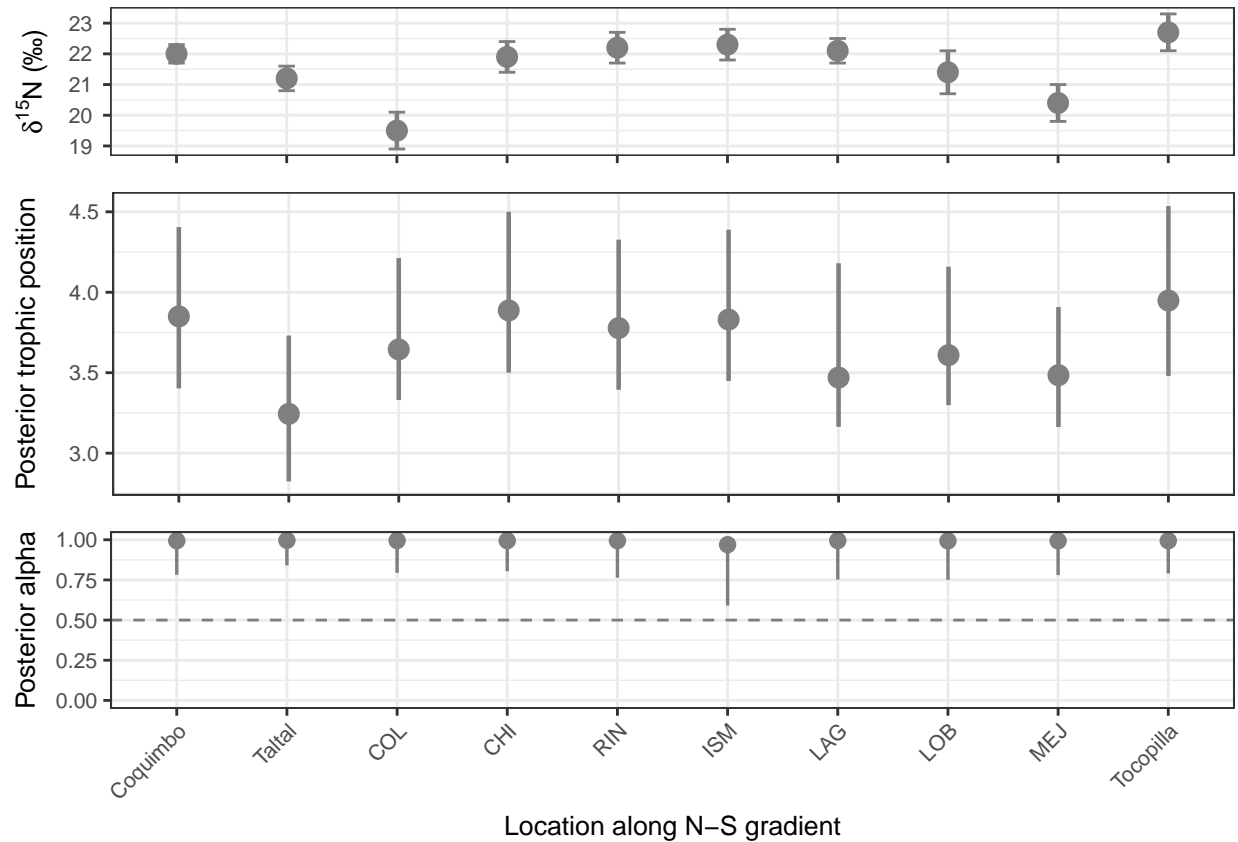
```
# object ggplot_blg
```

```
ggplot_blg <- credibilityIntervals(Bilagay_obs_summary,
                                   x = "Location",
                                   y1 = "mean_d15N",
                                   y1min = "mean_d15N - SD_d15N",
                                   y1max = "mean_d15N + SD_d15N",
                                   ylab1 = expression(paste(delta^{15}, "N (\u2030)")),
                                   plotAlpha = FALSE,
                                   plot = FALSE)
```

```
# We modify it by deleting both text and title from axis x, tweaking the text of axis y,  
# and adding a ggplot2 errorbar aesthetic
```

```
ggplot_blg <- ggplot_blg + ggplot2::theme(axis.title.x = ggplot2::element_blank(),
axis.text.x = ggplot2::element_blank(),
axis.text.y = ggplot2::element_text(
size = 8,
margin = ggplot2::margin(0,4,0,0,"pt"))) +
ggplot2::geom_errorbar(width = 0.15, color = "grey50", size = .5)

# And finally we plot all of them with grid.arrange()
gridExtra::grid.arrange(ggplot_blg, TP_alpha, nrow = 2, heights = c(3,11))
```



## References

1. Docmac, F., Araya, M., Hinojosa, I.A., Dorador, C. & Harrod, C. (2017). Habitat coupling writ large: pelagic-derived materials fuel benthivorous macroalgal reef fishes in an upwelling zone. *Ecology*, 98, 2267-2272.
2. Figueroa, D. & Moffat, C. (2000). On the influence of topography in the induction of coastal upwelling along the Chilean coast. *Geophysical Research Letters*, 27, 3905-3908.
3. Pérez-Matus, A., Pledger, S., Díaz, F.J., Ferry, L.A. & Vásquez, J.A. (2012). Plasticity in feeding selectivity and trophic structure of kelp forest associated fishes from northern Chile. *Revista chilena de historia natural*, 85, 29-48.