Methods in Ecology and Evolution

DR CLAUDIO QUEZADA-ROMEGIALLI (Orcid ID : 0000-0003-3380-389X)

DR BRIAN HAYDEN (Orcid ID : 0000-0002-8524-7373)

DR CHRIS HARROD (Orcid ID : 0000-0002-5353-1556)

Manuscript type: application

Title: *tRophicPosition*, an R package for the Bayesian estimation of trophic position from consumer stable isotope ratios

Running title: *tRophicPosition*

Authors: Claudio Quezada-Romegialli[*,1,2,3], Andrew L. Jackson[4], Brian Hayden[5,6], Kimmo K. Kahilainen[6,7,8], Christelle Lopes[9] & Chris Harrod[1,2,10]

Contact details:

[1] Fish & Stable Isotope Ecology Laboratory, Instituto de Ciencias Naturales Alexander Von Humboldt, Universidad de Antofagasta, Antofagasta, Chile

[2] Núcleo Milenio INVASAL, Concepción, Chile

[3] Current address: Departamento de Biología, Facultad de Ciencias Naturales y Exactas, Universidad de Playa Ancha, Valparaíso, Chile

[4] Department of Zoology, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland

[5] Canadian Rivers Institute and Biology Department, University of New Brunswick, Fredericton, New Brunswick, Canada E3B 5A3

[6] Kilpisjärvi Biological Station, University of Helsinki, Kilpisjärvi, Finland

[7] Department of Environmental Sciences, University of Helsinki, Helsinki, Finland

[8] Current address: Faculty of Biosciences, Fisheries and Economics, Arctic University of Norway, Tromsø, Norway

[9] Univ Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR5558, F-69622 Villeurbanne, France

[10] Universidad de Antofagasta Stable Isotope Facility, Universidad de Antofagasta, Antofagasta, Chile

* Correspondence author. E-mail: claudio.quezada@upla.cl

Abstract

1. Stable isotope analysis provides a powerful tool to identify the energy which sources fuel consumers, to understand trophic interactions, and to infer consumer trophic position (TP), an important concept that describes the ecological role of consumers in food webs.

However, current methods for estimating TP using stable isotopes are limited and do not fulfil the complete potential of the isotopic approach. For instance, researchers typically use point estimates for key parameters including trophic discrimination factors and isotopic baselines, and do not explicitly include variance associated with these parameters when calculating TP.

2. We present '*tRophicPosition*', an R package incorporating a Bayesian model for the calculation of consumer TP at the population level using stable isotopes, with one or two baselines. It combines Markov Chain Monte Carlo simulations through JAGS and statistical and graphical analyses using R. We model consumer and baseline observations using relevant statistical distributions, allowing them to be treated as random variables. The calculation of TP – a random parameter – for one baseline follows standard equations linking ^{15}N enrichment per trophic level and the trophic position of the baseline (e.g. a primary producer or primary consumer). In the case of two baselines, a simple mixing model incorporating δ^{13}C allows for the differentiation between two distinct sources of nitrogen, thus including heterogeneity derived from alternatives sources of δ^{15}N.

3. Methods currently implemented in '*tRophicPosition*' includes loading, plotting and summarising stable isotope data either from multiple sites and/or communities or a local assemblage; loading trophic discrimination factors from an internal database or generating them; defining and initialising a Bayesian model of TP; sampling posterior parameters; analysing, comparing and plotting posterior estimates of TP and other parameters; and calculating a parametric (non-Bayesian) TP estimate. Additionally, full

documentation including examples, multiple vignettes and code are available for download.

Resumen

1. El análisis de isótopos estables es una poderosa herramienta para identificar qué vías energéticas alimentan a los consumidores, para entender las interacciones tróficas, y también para inferir la posición trófica (PT) de los consumidores, un concepto importante que describe el rol ecológico de los consumidores en las tramas tróficas. Sin embargo, los métodos actuales para estimar la PT utilizando isótopos estables están limitados y no permiten alcanzar el potencial de la aproximación isotópica. Por ejemplo, los investigadores típicamente utilizan estimaciones puntuales de parámetros claves incluyendo factores de discriminación trófica y líneas base isotópicas, y no incorporan explícitamente la varianza asociada con estos parámetros cuando calculan la PT.

2. Presentamos *'tRophicPosition'*, un paquete de R para isótopos estables que incorpora un modelo Bayesiano para el cálculo de la PT en consumidores al nivel ecológico de población, con una o dos líneas base. Este paquete combina simulaciones de cadenas de Markov Monte Carlo a través de JAGS y utilizando los análisis estadísticos y gráficos de R. Modelamos las observaciones de consumidores y líneas base utilizando distribuciones estadísticas relevantes, por lo que son tratadas como variables aleatorias. El cálculo de la PT – un parámetro aleatorio – para una línea base considera ecuaciones estándar que

vinculan el enriquecimiento en ^{15}N por nivel trófico y el nivel trófico de la línea base (p.e. un productor primario o un consumidor primario). En el caso de dos líneas base, un modelo de mezcla simple incorporando δ^{13}C permite la diferenciación entre dos fuentes distintas de nitrógeno, incluyendo de esta forma la heterogeneidad derivada de fuentes alternativas de δ^{15}N.

3. Los métodos actualmente implementados en *'tRophicPosition'* incluyen cargar, graficar y resumir datos de isótopos estables desde múltiples sitios y/o comunidades o ensambles locales; cargar factores de discriminación trófica desde una base de datos interna o generarlos; definir e inicializar un modelo Bayesiano para la PT; muestrear parámetros *a posteriori*; analizar, comparar y graficar estimaciones *a posteriori* de la TP y otros parámetros; y calcular una estimación frecuentista (no Bayesiana) para la PT. Adicionalmente, toda la documentación incluyendo ejemplos, distintas viñetas y el código están disponibles para descarga.

Palabras clave: Métodos Bayesianos, Series de tiempo o espaciales; Tramas tróficas; Modelación (Ecología Comunitaria); Modelación (Ecología de Poblaciones); Software

Introduction

Stable isotope analysis has provided ecologists with a powerful tool to capture the complexity of trophic interactions (Peterson & Fry, 1987; Vander Zanden, Casselman, & Rasmussen, 1999; Vander Zanden & Rasmussen, 2001), to identify energy sources (Peterson, 1999; Phillips et al., 2014; Grey, 2016) and to infer consumer trophic position (Cabana & Rasmussen, 1994; Vander Zanden, Cabana, & Rasmussen, 1997; Post, 2002). Trophic position (TP) is an important concept

used to describe the ecological role of consumers in food webs (Cabana & Rasmussen, 1994; Post, 2002), and typical uses includes calculation of food chain length, maximum trophic position, degree of omnivory and bioaccumulation in food webs, among others (Cabana & Rasmussen, 1994; Vander Zanden et al., 1997; Post, Pace, & Hairston, 2000). However, current methods for estimating trophic position using stable isotopes are limited and do not fulfil the potential of the approach. For instance, users often simply compare consumer $\delta^{15}N$ with that of a baseline, ignoring the utility provided by including $\delta^{13}C$ (e.g. when estimating TP in systems with multiple energy pathways as developed by Post, 2002). Furthermore, researchers typically use point estimates for key parameters including trophic discrimination factors (TDF) (Vander Zanden et al., 1999; Post, 2002; but see Bastos, Corrêa, Winemiller, & Garcia, 2017) and isotopic baselines, and often do not explicitly include within-population variability i.e. that shown by individual consumers. A Bayesian approach to calculate TP overcomes many such limitations by explicitly including individual variability and propagating sampling error (of TDF, baselines, and higher consumers) in the modelling approach and posterior estimates of parameters (Jackson, Inger, Parnell, & Bearhop, 2011; Parnell et al., 2012; Phillips et al., 2014). Here, we detail an R package that implements a Bayesian approach to calculate TP at the population level.

Bayesian calculation of TP

tRophicPosition implements three Bayesian models, depending on the number of baselines and the inclusion of a TDF for carbon (C). The simplest model is:

$$\delta^{15}N_c = \delta^{15}N_b + \Delta N(TP - \lambda) \qquad \text{(eqn 1)}$$

where $\delta^{15}N_c$ and $\delta^{15}N_b$ refer to the $\delta^{15}N$ values of consumers and a single baseline respectively, ΔN is the TDF for nitrogen (N), TP is the trophic position of the consumer, and λ is the trophic position of the baseline (Vander Zanden et al., 1997). The dual baseline Bayesian approach includes a mixing model to discriminate among two distinct sources of C and N, e.g. pelagic vs benthic or terrestrial vs aquatic, and extends the single baseline approach using two equations:

$$\delta^{15}N_c = \Delta N(TP + \lambda) + \alpha(\delta^{15}N_{b1} + \delta^{15}N_{b2}) - \delta^{15}N_{b2} \qquad \text{(eqn 2)}$$

and

$$\delta^{13}C_c = \delta^{13}C_{b1}\,\alpha + \delta^{13}C_{b2}(1 - \alpha) \qquad \text{(eqn 3)}$$

where additional variables $\delta^{15}N_{b1}$, $\delta^{13}C_{b1}$, $\delta^{15}N_{b2}$ and $\delta^{13}C_{b2}$ refer to the $\delta^{15}N$ and $\delta^{13}C$ values of baselines 1 and 2 respectively, and α is the proportion of N derived from baseline 1 (Vander Zanden & Rasmussen, 2001; Post, 2002). The full model is based on eqn 2, but α includes the TDF for carbon:

$$\alpha = \left(\left(\delta^{13}C_{b2} - (\delta^{13}C_c + \Delta C)\right)/(TP - \lambda)\right)/(\delta^{13}C_{b2} + \delta^{13}C_{b1}) \qquad \text{(eqn 4)}$$

where the additional variables $\delta^{13}C_c$ is the $\delta^{13}C$ value of the consumer, and ΔC is the TDF for C (modified from Post, 2002).

In the Bayesian approach, C and N observations of consumers, baselines and TDFs are modelled as random variables, each having a prior normal distribution on their means (by default to the vague $dnorm(0, \tau)$, with precision $\tau = 1/SD^2$) and a uniform prior distribution on their standard deviations (by default set to the vague $dunif(0, 100)$), while TP and α are both treated as random parameters. Furthermore, TP is modelled as having a uniform prior distribution and the

simple mixing model proportion α as having a Beta distribution prior, whereas λ is fixed. Each of these prior distributions or values can be changed by the user through a specific function to define the respective Bayesian model (see below). The likelihood function assumes a normal distribution with mean defined by the equations 1–4 and a precision $\tau = 1/SD^2$ for each model, on the observed δ^{15}N and δ^{13}C values of consumers.

Implementation in R

tRophicPosition is based on the approach originally coded in BUGS by Lopes *et al.* (2011). Improved coding, addition and enhancement of models (based on Vander Zanden et al., 1997; Vander Zanden & Rasmussen, 2001; and modified from Post, 2002), and implementation in R using the framework of JAGS (Plummer, 2017) is presented here. JAGS is a cross-platform program that works with R to allow the analysis of Bayesian models using Markov Chains Monte Carlo simulations. It needs to be installed along with *tRophicPosition* (see S1–S4 for details). The basic use of the package *tRophicPosition* includes: i) loading, plotting and summarising isotopic data, ii) defining and initializing a Bayesian model of TP, and iii) estimating and plotting posterior TP (Table 1; Figure 1), and iv) to perform hypothesis tests comparing two or more posterior estimations of TP and/or α. A thorough use of the package (Table 1) includes generating random data for sensitivity analysis, simulating or calling TDFs from the internal database, extracting isotope data of multiple species from a data frame, performing a multi-model calculation of TP for one or more consumer groups or species, or performing calculations of TP for multiple species, communities or consumers across time. Additionally, we have included a parametric (non-Bayesian) calculation of TP based on means, and implemented a parallel calculation of Bayesian TP to make full use of multi-core processors.

Complete use of *tRophicPosition* is explained in the vignettes included as supplementary material (S1–S4) and documentation and examples within the package.

Basic use of tRophicPosition

Data can be loaded into R according to standard procedures, following the convention of having $\delta^{15}N$ values for one consumer and one baseline at least (see S1 for details). To convert a data frame into an object *isotopeData* ready to use within the package, we developed the function *loadIsotopeData()*. By default, this function searches for the columns labelled *d13C* and *d15N* and extracts isotope values grouping them according to the *FG* (i.e. functional group) column, that should have "consumer" and "baseline1" as values. An important assumption associated with the accurate estimation of TP is the use of a relevant TDF (Vander Zanden & Rasmussen, 2001; Post, 2002; McCutchan, Lewis, Kendall, & McGrath, 2003). By default, the function *loadIsotopeData*() uses a mean ± SD TDF of 3.4 ± 0.98 for N and of 0.4 ± 1.3 for C (Post, 2002). However, the user can easily use other TDF values, including values from meta-analyses (e.g. McCutchan et al., 2003), the literature, or their own values based on experimental estimates of TDFs (e.g. Healy et al., 2017) (see S2 for details).

Once stable isotope data are loaded into R, the user can undertake some initial graphical analysis using the method *plot()* (Figure 2a), to ensure that the data processed are correctly loaded and follow the best practices of Phillips *et al*. (2014). After this initial analysis, the user can decide to use a single or dual baseline (i.e. using only TDF for N) instead of the full (i.e. dual baseline model including TDF for C) Bayesian model, using the function *jagsBayesianModel(model = "oneBaseline")* and changing *"oneBaseline"* with one of the other models (*"twoBaselines"* or *"twoBaselinesFull"*). In this step, the Bayesian approach is implemented, defining priors for any

of the parameters and variables involved in the respective model, such as the value for λ (by default is 2, see below for details).

When the Bayesian model has been chosen and appropriate priors set, the model must be initialised using the function *TPmodel()* with the stable isotope data, the Bayesian model, and the number of chains as arguments. The number of chains must be ≥ 2 when summary statistics such as Gelman-Rubin (Gelman & Rubin, 1992) are calculated within *tRophicPosition* (or e.g. using the package coda; Plummer, Best, Cowles, & Vines, 2006). Markov Chain Monte Carlo sampling is then performed by the function *posteriorTP()*, selecting the variables to monitor and the number of iterations (e.g. *posteriorTP(model = model, n.iter = 20000, variable.names = c("TP", "muDeltaN")))*. This generates distributions of posterior estimates of *TP*, α and/or other parameters which are then available for the user to save, plot and analyse (Figure 2b). For multiple models, multiple species and parallel computation of TP please refer to Appendixes S2–4.

Checking the model: customizing parameters and priors

Although the default settings will be useful in most cases, users that want to take full advantage of the Bayesian approach may want to customize parameters and priors for the model. The most relevant parameters are i) λ, which defines the trophic level of the baseline used; ii) the number of chains, needed for checking convergence of independent replicates; and iii) the number of adaptive, burn-in and actual sampling iterations, that are the key for getting reliable posterior estimates of *TP* and α and to perform model checking. λ may be changed from the default value of two, to any trophic level that coincides to the baseline used, during the call to the *jagsBayesianModel* function (e.g. *jagsBayesianModel(lambda = 1)* for a baseline equal to 1, or even using a non-integer baseline with *jagsBayesianModel(lambda = 2.5)* if desired). A number

of chains ≥ 2 is required both to calculate summary statistics of convergence and to check the reliability of posterior inference of parameters. The function *coda::gelman.diag(samples)* will calculate a potential scale reduction factor and its upper limits (provided that the coda package is installed, and the object *samples* has the posterior estimates of *TP*, α and/or other parameters inferred with the *tRophicPosition* function *posteriorTP*). Gelman & Rubin (1992) indicated that values close to 1 for the potential scale reduction factor and its upper limit indicate that convergence was achieved and more iterations will not improve estimations of posterior distributions of parameters. Also, a well-mixed posterior distribution (such as the one shown in Figure 2b) is another clear and simple way of checking the convergence of parameters of interest. If chains are not converging it may suggest that the model was badly specified, the TDF values or the data of consumers or baselines are not normally distributed, or that the model requires more iterations to achieve convergence. Longer simulations can be achieved by extending both the burn-in or the actual sampling. While the adaptive phase is not used in the posterior inference of the model, burn-in discards an initial amount of iterations and extends the actual sampling, thus, increases the likelihood of achieving convergence of posterior distributions of parameters. By default there are 10 000 adaptive samplings, 10 000 iterations and no burn-in, but they can be changed within the call to the functions *TPmodel* (e.g. *TPmodel(n.adapt = 20000)* for increasing adaptive sampling), *posteriorTP* (e.g. *posteriorTP(n.init = 20000, burnin = 20000*) for increasing both actual sampling and burn-in) *multiModelTP()* and *multiSpeciesTP()* (and using the same arguments n.adapt, n.init and burnin for the two last functions). Preliminary tests show that changing priors does not have a large effect on posterior distributions for the models currently implemented in *tRophicPosition*, which suggests that using non-skewed, normally distributed data with enough observations of consumer

and baselines has greater influence on the modelling approach than modifying priors. Still, users can change any prior distribution within the call to *jagsBayesianModel* using the distributions available in the bugs module (a complete list is found in Plummer, 2017:43). However, users will typically use as priors a normal distribution (e.g. *jagsBayesianModel(TP = dnorm(4, 0.5)),* for a normally distributed prior for *TP* with mean 4 and SD 0.5), a beta distribution (e.g. *jagsBayesianModel(alpha = dbeta (1, 10))*, for a beta distributed prior for α, with parameters a = 1 and b = 10) or a uniform prior (e.g. *jagsBayesianModel(TP = dunif(2, 10))*, for a uniform prior for *TP*, with limits within [2, 10]). We recall that the distributions used as arguments in *jagsBayesianModel()* not necessarily are the same as those used in R, so users should check Plummer (2017) for details and using other custom distributions as priors.

Once the Bayesian model has been built and the posterior distribution of all parameters has been computed, the important step of model-checking is a crucial task to include in any statistical analysis (Gelman et al., 2014). One way of performing model-checking is applying predictive posterior checks. The idea behind this is that if the model fits, then replicated data generated under the model should look similar to observed data. If the observed data differ from the simulated data, then the Bayesian model has failed in some way to model the true data generating process, which may imply that the model-based estimates of *TP* or α might be inaccurate. In the three Bayesian models implemented in *tRophicPosition* users can generate, and then extract the δ^{13}C and δ^{15}N stable isotope data generated from the posterior predictive distribution for each baseline and consumer, and compare them with the observed data (see S1 for details). As the raw data from posterior predictive distributions are available using the function *extractPredictiveData(posterior)* (where *posterior* is the object that has the posterior distribution of predicted parameters), users can calculate correlations, the root mean squared

differences between observed and predicted data, or any other summary or statistic from them, such as those shown in Figure 3 or specialized software such as the R package DHARMa (Hartig, 2017).

Assumptions and limitations

The three Bayesian models included in the current version of the package were not coded as hierarchical models, thus the multiple models, multiple species and parallel calculations of TP and α are fitted independently for each consumer. With this in mind, users can change the freely available code to their own needs, for example including non-normal distributions for consumers and baselines in the likelihood functions of any of the models included in *tRophicPosition*, or modify them into a hierarchical version. An important aspect of these models is that the TDFs are incorporated into the models as raw data, instead of using a mean and SD as seen in other software. This approach was implemented to propagate uncertainty in TDF knowledge of the system under study into the model, and users must be aware that using a reduced number of TDF observations will have a great impact on the posterior inference of parameters (see S2 for details). The selection of a representative TDF is a key aspect of any isotopic analysis (Phillips et al., 2014).

Case studies showing the capacities of the package

We illustrate the functionality of the package using three case studies from freshwater and marine ecosystems (see Appendix S4).

In the first example (Fig. 4a) , we use data from two relatively large and deep oligotrophic Finnish lakes Inarijärvi and Kilpisjärvi to calculate TP among 20 freshwater fish species and maximum TP (MTP) per lake (Post et al., 2000; Thomas, Harrod, Hayden, Malinen, &

Kahilainen, 2017). This case study represents a typical analysis used by researchers when they compare TP differences among species or groups of consumers (e.g. fish morphs) within or even between ecosystems. Figure 4a shows TP calculated with a one baseline Bayesian model (baseline δ^{15}N was similar for benthic and pelagic baselines in both lakes, see Appendix S4), showing differences between some common species, and that MTP was similar in the two lakes.

The second example (Fig 4b) represents a common situation where TP is compared within a species, e.g. across space or time. Here we use data from a benthic rockfish common to the coastal kelp forests of Northern Chile, bilagay *Cheilodactylus variegatus*, which, along with filter feeding and grazing molluscs (representing pelagic and benthic baselines respectively) were sampled from a series of different locations (see Appendix S4 and Docmac, Araya, Hinojosa, Dorador, & Harrod, 2017). In this example our question is: given the potential influences of local differences in diet stomach contents and upwelling intensity (which increases along the S-N gradient), are there measurable differences in bilagay TP between the different locations along a S-N latitudinal gradient? The results show that bilagay δ^{15}N values differ considerably according to location, but posterior estimates of TP were largely similar.

The third example examines if a consumer shows an ontogenetic shift in TP. *tRophicPosition* provides TP estimates at a population level, and here we examine how TP varies across different size classes of the omnivorous fish the roach, *Rutilus rutilus*, captured from Lough Neagh, Northern Ireland (Fig 4c). Fish were placed into deciles based on their fork length (mm) and we compared TP across different size classes. The results clearly showed an ontogenetic shift and that juvenile fish (less than 70 mm) had a higher TP than larger individuals. This fits well with observations of gut contents that show a switch from zooplanktivory to omnivory at this size in this system (C.H., unpublished data).

Documentation and data accessibility

Package documentation includes a short guide to the use, multiple models and multiple species calculation of TP and can be found in the supplementary material (S1–S3), and vignettes included within the package (*browseVignettes("tRophicPosition")*). An expanded description and R code of the examples and figures generated in this manuscript can be found in supplementary material S4 and S5 respectively.

tRophicPosition is under active development. Future versions of the package will include different methods and hierarchical models including the use of scaled TDFs (Hussey et al., 2014) and the estimation of consumer TP from amino acid δ^{15}N (Bradley et al., 2015) that include variation in TDF and β. The development version is freely available in GitHub (https://github.com/clquezada/tRophicPosition), while the stable version (currently 0.7.5) can be found in CRAN (https://cran.r-project.org/package=tRophicPosition).

Acknowledgments

Author contributions

C.Q.-R., C.L., C.H. and A.J. conceived the ideas. C.Q.-R developed the model, with inputs from C.H. and A.J. C.H., B.H. and K.K.K. collected data, C.Q.-R. and C.H. analysed the data, C.Q.R. and C.H. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

References

Bastos, R. F., Corrêa, F., Winemiller, K. O., & Garcia, A. M. (2017). Are you what you eat ? Effects of trophic discrimination factors on estimates of food assimilation and trophic position with a new estimation method. *Ecological Indicators*, *75*, 234–241. doi:10.1016/j.ecolind.2016.12.007

Bradley, C. J., Wallsgrove, N. J., Choy, C. A., Drazen, J. C., Hetherington, E. D., Hoen, D. K., & Popp, B. N. (2015). Trophic position estimates of marine teleosts using amino acid compound specific isotopic analysis. *Limnology and Oceanography: Methods*, *13*(9), 476–493. doi:10.1002/lom3.10041

Cabana, G., & Rasmussen, J. B. (1994). Modelling food chain structure and contaminant bioaccumulation using stable nitrogen isotopes. *Nature*. doi:10.1038/372255a0

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*. doi:10.1002/ecy.1936

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (Third Edit). Boca Ratón: CRC Press.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4), 457–511.

Grey, J. (2016). The Incredible Lightness of Being Methane-Fuelled: Stable Isotopes Reveal Alternative Energy Pathways in Aquatic Ecosystems and Beyond. *Frontiers in Ecology and Evolution*, *4*(February), 1–14. doi:10.3389/fevo.2016.00008

Hartig, F. (2017). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level Mixed) Regression Models. R package version 0.1.5.

Healy, K., Guillerme, T., Kelly, S. B. A., Inger, R., Bearhop, S., & Jackson, A. L. (2017). SIDER: an R package for predicting trophic discrimination factors of consumers based on their ecology and phylogenetic relatedness. *Ecography*. doi:10.1111/ecog.03371

Hussey, N. E., MacNeil, M. A., McMeans, B. C., Olin, J. A., Dudley, S. F. J., Cliff, G., … Fisk, A. T. (2014). Rescaling the trophic structure of marine food webs. *Ecology Letters*, *17*(2), 239–250. doi:10.1111/ele.12226

Jackson, A. L., Inger, R., Parnell, A. C., & Bearhop, S. (2011). Comparing isotopic niche widths among and within communities: SIBER - Stable Isotope Bayesian Ellipses in R. *Journal of Animal Ecology*, *80*(3), 595–602. doi:10.1111/j.1365-2656.2011.01806.x

Lopes, C., Perga, M.-E., Peretti, A., Roger, M.-C., Persat, H., & Babut, M. (2011). Is PCBs concentration variability between and within freshwater fish species explained by their contamination pathways? *Chemosphere*, *85*(3), 502–508. doi:10.1016/j.chemosphere.2011.08.011

McCutchan, J. H., Lewis, W. M., Kendall, C., & McGrath, C. C. (2003). Variation in trophic

shift for stable isotope ratios of carbon, nitrogen, and sulfur. *Oikos*, *102*(2), 378–390. doi:10.1034/j.1600-0706.2003.12098.x

Parnell, A. C., Phillips, D. L., Bearhop, S., Semmens, B. X., Ward, E. J., Moore, J. W., … Inger, R. (2012). Bayesian Stable Isotope Mixing Models. *Environmetrics*, *24*(6), 387–399. doi:10.1002/env.2221

Peterson, B. J. (1999). Stable isotopes as tracers of organic matter input and transfer in benthic food webs: A review. *Acta Oecologica*, *20*(4), 479–487. doi:10.1016/S1146-609X(99)00120-4

Peterson, B. J., & Fry, B. (1987). Stable Isotopes in Ecosystem Studies. *Annual Review of Ecology and Systematics*, *18*(1), 293–320. doi:10.1146/annurev.es.18.110187.001453

Phillips, D. L., Inger, R., Bearhop, S., Jackson, A. L., Moore, J. W., Parnell, A. C., … Ward, E. J. (2014). Best practices for use of stable isotope mixing models in food web studies. *Canadian Journal of Fisheries and Aquatic Sciences*, *835*, 823–835. doi:10.1139/cjz-2014-0127

Plummer, M. (2017). *JAGS version 4.3.0 user manual*. Retrieved from https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence Diagnostics and Output Analysis for MCMC. *R News*, *6*(1), 7–10. Retrieved from https://www.r-project.org/doc/Rnews/Rnews_2006-1.pdf

Post, D. M. (2002). Using Stable Isotopes to Estimate Trophic Position: Models, Methods, and Assumptions. *Ecology*, *83*(3), 703. doi:10.2307/3071875

Post, D. M., Pace, M. L., & Hairston, N. G. (2000). Ecosystem size determines food-chain length in lakes. *Nature*, *405*(6790), 1047–1049. doi:10.1038/35016565

Thomas, S. M., Harrod, C., Hayden, B., Malinen, T., & Kahilainen, K. K. (2017). Ecological speciation in a generalist consumer expands the trophic niche of a dominant predator. *Scientific Reports*, *7*(1), 8765. doi:10.1038/s41598-017-08263-9

Vander Zanden, M. J., Cabana, G., & Rasmussen, J. B. (1997). Comparing trophic position of freshwater fish calculated using stable nitrogen isotope ratios (? 15 N) and literature dietary data. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(5), 1142–1158. doi:10.1139/f97-016

Vander Zanden, M. J., Casselman, J. M., & Rasmussen, J. B. (1999). Stable isotope evidence for the food web consequences of species invasions in lakes. *Nature*, *401*(6752), 464–467. doi:10.1038/46762

Vander Zanden, M. J., & Rasmussen, J. B. (2001). Variation in delta N-15 and delta C-13 trophic fractionation: Implications for aquatic food web studies. *Limnology and Oceanography*, *46*(8), 2061–2066. doi:10.4319/lo.2001.46.8.2061

Tables

Table 1. Main functions included in *tRophicPosition* grouped by type of analysis

Function	Description
Isotope data manipulation	
loadIsotopeData()	Load isotope data for one consumer from a data frame and convert it to an isotopeData class object
extractIsotopeData()	Extract isotope data for multiple species from a data frame and convert them to isotopeData class objects
generateTPData()	Generates random stable isotope data (δ^{15}N and δ^{13}C) for trophic position calculation
TDF()	Get trophic discrimination factors (TDF) values from internal database
simulateTDF()	Simulate TDFs
summariseIsotopeData()	Summarise isotope data
parametricTP()	Parametric (non-Bayesian) calculation of trophic position, based on means
Screening and plotting data	
screenIsotopeData(), plot()	Biplot of consumer and baselines
screenFoodWeb()	Biplot of a food web, coloured by species and shaped by functional group
credibilityIntervals()	Plot of credibility intervals and median for posterior estimations
Bayesian modelling	
jagsBayesianModel()	Defines a Bayesian model for trophic position calculation
TPmodel()	Creates a JAGS model object to calculate trophic position
posteriorTP()	Generates random samples from the posterior distribution of a JAGS model
extractPredictiveData()	Extracts predictive data of baselines and consumers, if they were monitored when calling posteriorTP()
multiModelTP()	Calculation of trophic position for one consumer using one or more Bayesian models
multipleSpeciesTP()	Calculation of trophic position for multiple species dataset using one or more Bayesian models
Hypothesis testing	
compareTwoDistributions()	Compare samples from two posterior distributions and test a hypothesis

Function	Description
pairwiseComparisons()	Perform a pairwise comparison between two or more posterior distributions

Figure captions

Fig. 1. Flowchart showing basic workflow of *tRophicPosition*.

Fig. 2. Plots generated to analyse trophic position from stable isotope data. (a) Output from *plot()* method and an isotopeData class object, (b) plot generated with the CODA package (Plummer et al. 2006) using the *summary()* function applied to a posterior distribution of TP and α. Data and figures were generated using the code available in Appendix S5.

Fig. 3. Posterior predictive model-checking plots for the example shown in Figure 2. (a) Raw residuals of $\delta^{15}N$ predicted and $\delta^{15}N$ observed data. (b) Histogram of the $\delta^{15}N$ observed data superimposed with a density function of the $\delta^{15}N$ predicted data. Data and figures were generated using the code available in Appendix S5.

Fig. 4. Case studies illustrating typical uses of *tRophicPosition*: (a) Variation in estimated TP of different fish species and whitefish morphs from Lakes Inarijärvi and Kilpisjärvi, ordered by modal TP values. The horizontal dotted line shows maximum trophic position, which is similar in the two lakes. (b) Variation in $\delta^{15}N$, TP and α (relative use of pelagic-derived N) in bilagay (a benthic rockfish in N Chile) along a latitudinal gradient. The upper panel indicates mean ± SD

δ^{15}N of bilagay, the mid-panel shows posterior TP estimations and lower panel shows estimated α at each location. (c) Variation in posterior TP and α estimated for ten fork length deciles (range 45 – 295 mm) of roach from Lough Neagh (whole population estimate shown in blue). In all graphs, bars show 95% credibility interval of each posterior TP and estimates. For more details of case studies please see Appendix S4.

Supporting information

Appendix S1. The vignette 'A short guide to *tRophicPosition'* with an example on running a trophic position analysis.

Appendix S2. The vignette 'Multiple Bayesian models calculation with *tRophicPosition'*, with examples on calculating multiple Bayesian models of trophic position for a single species, and calculation of parametric trophic position.

Appendix S3. The vignette 'Multiple species calculation with *tRophicPosition'*, with examples on managing isotope data with multiple species, plotting isotope data, and performing multiple Bayesian comparison of trophic position and alpha parameters between groups.

Appendix S4. R code for the examples included in case studies.

Appendix S5. R code for generating and plotting Figures 2 and 3.

Input data	Stable Isotope Data
Conversion into an isotopeData class object	extractIsotopeData()
tRophicPosition working format	isotopeData class object
Data checking	summary() plot()
Model definition	jagsBayesianModel()
Model initialization	TPmodel()
Posterior sampling	posteriorTP()
Output data	Trophic position alpha
Model checking	extractPredictiveData()
Post analysis	pairwiseComparisons() pairwiseComparisons()

(a)

(b)