Effect of the amount of gut content data on the accuracy and precision of food web prediction

June 2, 2021

Abstract

1. 1)
2. 2)
3. 3)
4. 4)

5. **K**eywords gut content data · ADBM · food web accuracy · food web prediction

1 Alternate titles

7. • Predicting trophic interactions using incomplete gut content data
8. • Optimal amount of gut content data required to predict trophic interactions

2 Introduction

2.1 Why knowing trophic interactions is important?

11. Knowledge about the trophic interactions in a food web can be used to answer crucial ecological questions:
12. Which species is a keystone species in a food web (Jordán 2009)? How would a food web respond if the
13. top predators are removed (Knight et al. 2005)? Which species should be given priority when conservation
14. efforts are implemented (McDonald-Madden et al. 2016)? How robust is a food web to extinctions (Dunne,
15. Williams, and Martinez 2002)?

2.2 What are the methods for knowing interactions?

17. Species interactions can be inferred from different types of methods such as gut content analysis (Peralta-
18. Maraver, Lopez-Rodriguez, and de Figueroa 2017), stable isotope ratio analysis of tissues (Layman et al.
19. 2007), experimentation (Warren 1989) and DNA metabarcoding (Roslin and Majaneva 2016). The types
20. of information extracted from these methods vary. E.g. Gut content analysis of a predator results in the
21. identification of prey species fed upon by that predator, while the stable isotope ratio analysis of the organism's
22. tissue can help us infer the approximate trophic position of that species in the food web.

2.3 Gut content analysis

24. Gut content data have been used to predict the diet of a predator in a food web (Eitzinger et al. 2018;
25. O'NAMalley et al. 2017; Dixon et al. 2017; McClain-Counts, Demopoulos, and Ross 2017; Peralta-Maraver,
26. Lopez-Rodriguez, and de Figueroa 2017). A single gut content data of a predator gives the information what
27. the predator has consumed recently. The diet of that individual can be inferred from the items in the stomach.
28. There is considerable uncertainty involved in this method (Baker, Buckland, and Sheaves 2014). E.g. In fish,

there is ambiguity while separating prey items in the gut. There are loose tissues that are not identifiable and cannot be assigned to a specific prey item with certainty. There are factors such as sample size of consumers, mechanical prey handling, differential digestion and evacuation rates of different types of different prey types and volumes, and the ingestion order that in combination result in an unquantifiable error which is difficult to interpret in the predator diet (Hyslop 1980; Rindorf and Lewy 2004; Baker, Buckland, and Sheaves 2014).

2.4 Why inferring interactions is not easy?

As a food web can be quite complex with many species and trophic interactions (Williams and Martinez 2000; Berlow et al. 2009), inferring these trophic interactions with high accuracy and low uncertainty would require a large amount of food web data (citreqd). E.g. One would need to sample many individuals for gut content data, and processing these guts require a large amount of effort and time as well.

Gut content data might only be used to infer the partial diet of a predator individual because gut content data only provides the information of the prey individuals that have been foraged by the predator recently. There is also a possibility that some of the information provided by these types of food web data are redundant i.e. taking into account any further food web data does not necessarily increase the food web prediction significantly. E.g. This can happen if the guts of two predator individuals give the same information about the prey the predator has consumed.

2.5 Food web model can be used to compensate for undersampling

When interactions are difficult to observe in a food web, this results in an undersampled food web i.e there are missing links in the food web. Limitations of resources and time can also lead to an undersampled food web because collecting and analysing gut content data is tedious and expensive task (citreqd). It is crucial to have well sampled food web because food web properties are highly sensitive to varying levels of sampling effort (Goldwasser and Roughgarden 1997). It is important to predict food web properties with high accuracy and high precision. Hence, to compensate for undersampling we can use a food web model which can help us infer missing links.

2.6 Explain oversampling/diminishing returns when using a food web model

A food web model used to predict trophic interactions is parameterised using food web data such as gut content data. When one has a well sampled food web constructed with information from gut content data, using a food web model might not result in any further improvement in food web prediction. It is crucial to know at what amount of gut content data this occurs.

2.7 Research question

How much gut content data is required when a food web model is used to predict trophic interactions? We are interested in maximizing the accuracy and precision of the predicted food web and simultaneously minimizing the amount of gut content data used.

2.8 What we do in our study and how?

In this paper, we answer how much gut content is optimal to predict a food web with high accuracy and high precision and until what quantity of gut content data it is advantageous to use a food web model. We predict the trophic interactions using allometric diet breadth model (ADBM) and use incomplete gut content data to parameterise the ADBM using approximate Bayesian computation (ABC).

3 Materials and Methods

In the upcoming sections, we present the gut content data and the allometric diet breadth model (ADBM) used to infer trophic interactions. We also give a detailed account of using partial gut content data to parameterise the ADBM using approximate Bayesian computation (ABC). We assessed model predictions using true skill statistic and various structural food web properties for comparison across food webs.

3.1 Broadstone Stream food web

Broadstone Stream (51° 05' N 0° 03' E; 120 m above sea-level) is a second-order tributary of the River Medway in south-east England. It is dominated by invertebrates as the acidity of the stream (pH 4.7-6.6) excludes fish. The stream consists of 25 common invertebrate species (Woodward, Speirs, and Hildrew 2005a). Among the predators, the three large species are *Cordulegaster boltonii* Donovan, *Sialis fuliginosa* Pict. and *Plectrocnemia conspersa* [Curtis] and the three small species are the larvae of the tanypod midges *Macropelopia nebulosa* [Meigen], *Trissopelopia longimana* [Staeger], and *Zavrelimyia barbatipes* [Kieffer]. Broadstone Stream food web is one of the most completely described food webs for any system (Schmid-Araya et al. 2002; Hildrew 2009; Layer et al. 2010; and Woodward, Speirs, and Hildrew 2005b).

The Broadstone Stream food web in our study was constructed by aggregating by size (Woodward et al. 2010a) i.e a node is defined as collection of individuals within a specified size range. There were a total of 1008 gut content data. The number of prey items in predator's gut varied within a predator size class (Fig. 1). A single gut did not give complete information about the diet of a predator as can be observed from the species accumulation curves (Fig. 1).

3.2 Celtic Sea food web

Description of Celtic Sea food web here

3.3 Allometric Diet Breadth Model (ADBM)

The allometric diet breadth model (ADBM) is based on optimal foraging theory, specifically the contingency foraging model (MacArthur and Pianka 1966). The ADBM predicts the set of prey species a consumer should feed upon to maximise its rate of energy intake (Petchey et al. 2008). The foraging variables in the model: energy content of the resources, handling times of the prey, space clearance rate and prey densities are allometrically scaled to the body sizes of the species.

The model correctly predicted a greater proportion of links when the trophic interactions were more strongly dependent on size (Petchey et al. 2008). Indeed, constructing a food web based only on body size (i.e. ignoring taxonomy) resulted in almost twice the number of correctly predicted links, i.e. 83%, in contrast to taxonomy (Woodward et al. 2010b). The goodness of fit of the ADBM's predictions depends on the types of interactions in the food webs (Petchey et al. 2008), because some of the interactions are more size structured than other interactions. For instance, predacious and aquatic herbivore interactions were predicted better than parasitoid and herbivory ones (Petchey et al. 2008).

3.4 Inferring food web using gut content data

From a pool of gut content data, we draw a set of gut content data randomly. Then we used the rejection ABC to accept a parameter value which would have resulted in the minimum distance, where distance = 1 - True skill statistic and the true skill statistic was computed between the predators' diet predicted from the ADBM and the sampled gut content data. We repeated this process for n (= 100) number of times for every i number of guts, where i lies between 1 and total number of gut content data in the pool.

Input:

- Predators $P : P = \{p_1, p_2, \ldots, p_k\}$

- A pool of gut content data $G : G = \{g_1, g_2, \ldots, g_n\}$, where g_n is a one-dimensional matrix containing ones and zeros.

- A model prediction $model(\theta) : ADBM(\theta) = \{d_{p_1}, d_{p_1}, \ldots, d_{p_k}\}$, where d_{p_k} is a one-dimensional diet matrix of predator k containing ones and zeros.

- A summary statistic $s(x) : s(x) \subseteq model(\theta)$

- A distance function $d(x_i, y) : d(x_i, y) = 1 - TSS(x_i, y)$

- An observed food web $Y : Y = \{d'_{p_1}, d'_{p_1}, \ldots, d'_{p_k}\}$

Sampling:

for $i = 1, \ldots, tgut$ where *tgut* is the total number of gut content data in the pool G
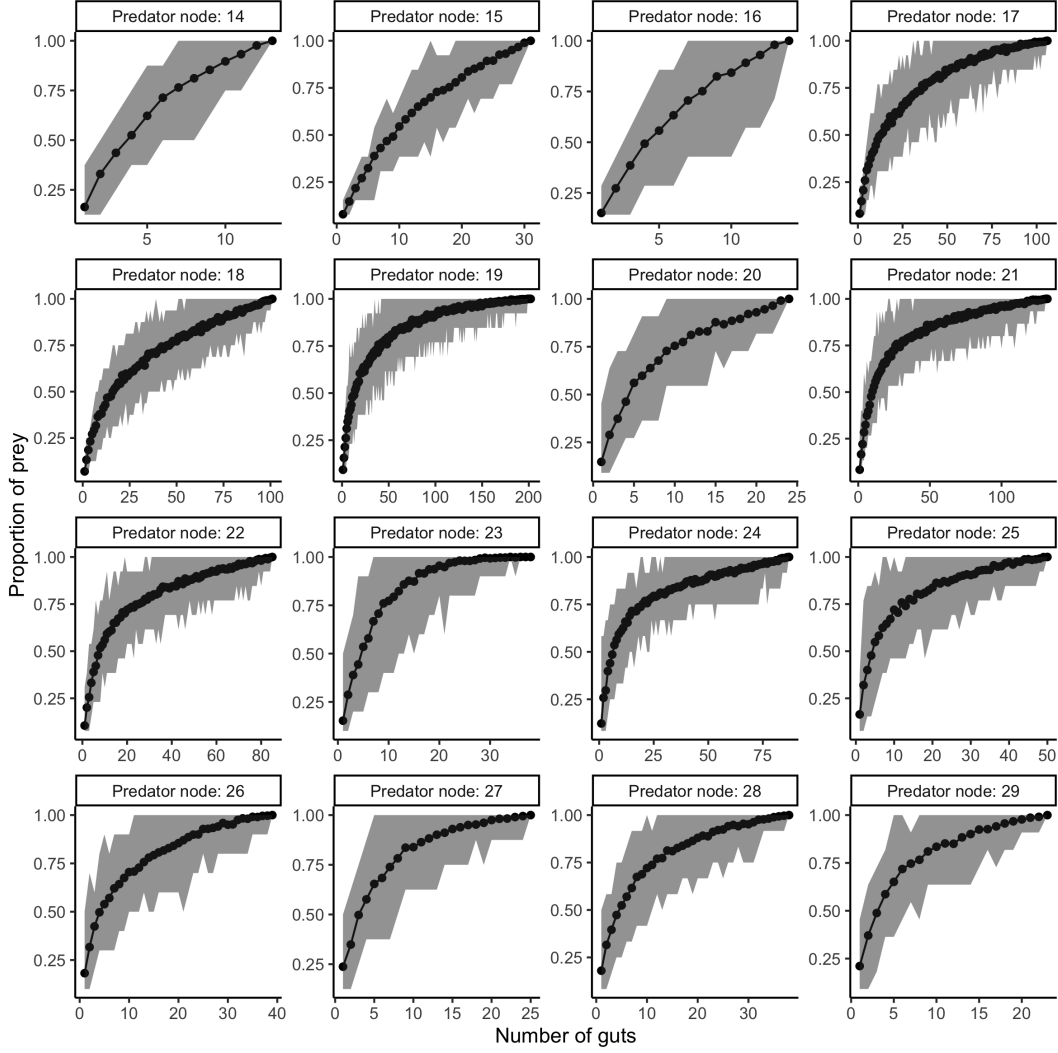
Figure 1: Species accumulation curve of diets of the predators belonging to a size class in Broadstone Stream food web. Each panel represents diet of predators belonging to different size classes. The grey region represents the prediction interval of proportion of prey for the 1000 independent samples drawn from a pool of gut content data.

- for $j = 1, \ldots, nsample$ where $nsample$ is the number of independent samples to be drawn
  - Draw a set of gut content data $y = \{g_1, g_2, \ldots, g_i\}$ from the pool of gut content data G
  - for $k = 1, \ldots, npar$ where $npar$ is the number of parameter values to be sampled
    * Draw a set of parameter values θ_k from the prior distribution $\pi(\theta)$
    * Compute the model result $x_i = model(\theta_k)$
    * Compute $s(x_i)$ and $d(s(x_i), y)$
  - Accept θ_j, which results in the $min_i\{d(s(x_i), y)\}$

- Compute $TSS_i(x, y) = \{TSS(x_i, y) : x_i = ADBM(\theta_j), \theta_j$ computed from previous step$\}$ using the accepted $\theta_1, \ldots, \theta_{nsample}$

Output:

A prediction interval containing TSS between observed and predicted food webs for every i number of gut content data drawn from the pool of gut content data.
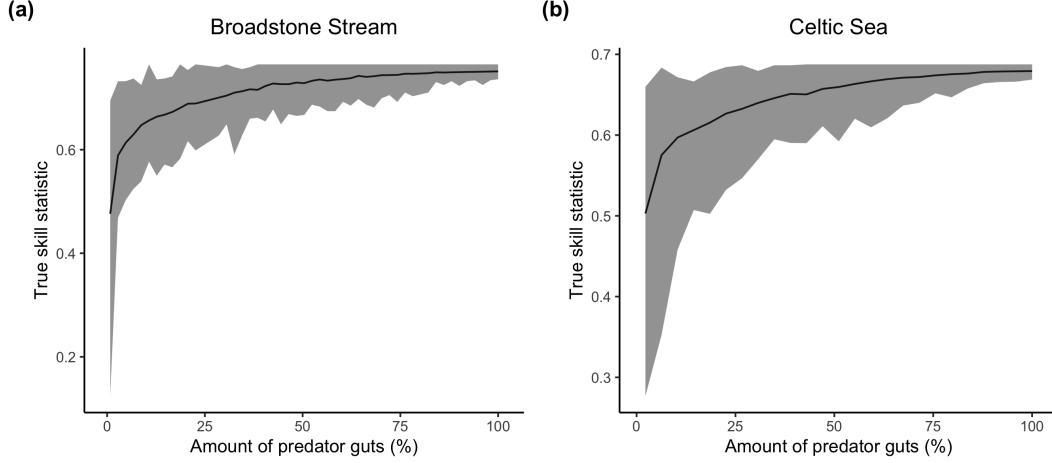
4

(a)

Broadstone Stream

(b)

Celtic Sea

Figure 2: True skill statistics between the predicted and observed predation matrices for Broadstone Stream and Celtic Sea food web computed for different amount of predator guts. Line and shaded region represents the mean and the prediction interval of 100 independent samples respectively.

3.5 Assessment of prediction

The accuracy of the predicted diet of the predators was defined using true skill statistic (TSS) which takes into account the true and false predictions of both the presence and absence of links defined as:

$$\text{TSS} = \frac{ad - bc}{(a + c)(b + d)}$$

where a is the number of observed links that are predicted by the model (true positives), d is the number of observed absences of links that are correctly predicted (true negatives), b is the number of false positives, and c is the number of false negatives.

The TSS ranges from -1 to 1, where $+1$ indicates a perfect prediction. A TSS value of zero or less indicates a performance no better than random.

4 Results

We first present the accuracy of the ADBM in predicting food web when we parameterised the ADBM using incomplete gut content data. Then, we present the upper threshold of the amount of gut content data where it is profitable/advantageous to use the ADBM. We also compute the structural properties and compare the predicted structural properties from the model with that obtained using only gut content data.

4.1 Inferring trophic interactions using ADBM and incomplete gut content data

The value of true skill statistics between the predicted food web using the ADBM and the observed food web saturated quickly with the increasing number of predator guts (Fig. 2). The uncertainty in the TSS reduced with the increasing number of predator guts. With only 23% of the total gut content data, the ADBM was able to achieve 0.69 TSS which was 91% of the mean TSS (0.75) achieved using complete gut content data by the ADBM for Broadstone Stream food web (Fig. 2(a)). In case of the Celtic Sea food web using only 23% of the complete gut content data, the ADBM achieved a TSS of 0.63 which was 92% of the mean TSS (0.68) obtained by the ADBM when we used the complete gut content data (Fig. 2(b)).

4.2 Accuracy of food web constructed using model and data vs accuracy of food web constructed using data only

It was advantageous to use the ADBM for predicting trophic interactions when we have a low amount of gut content data (Fig. 3) i.e less than 30-40% of the total gut content data because we achieve a higher TSS
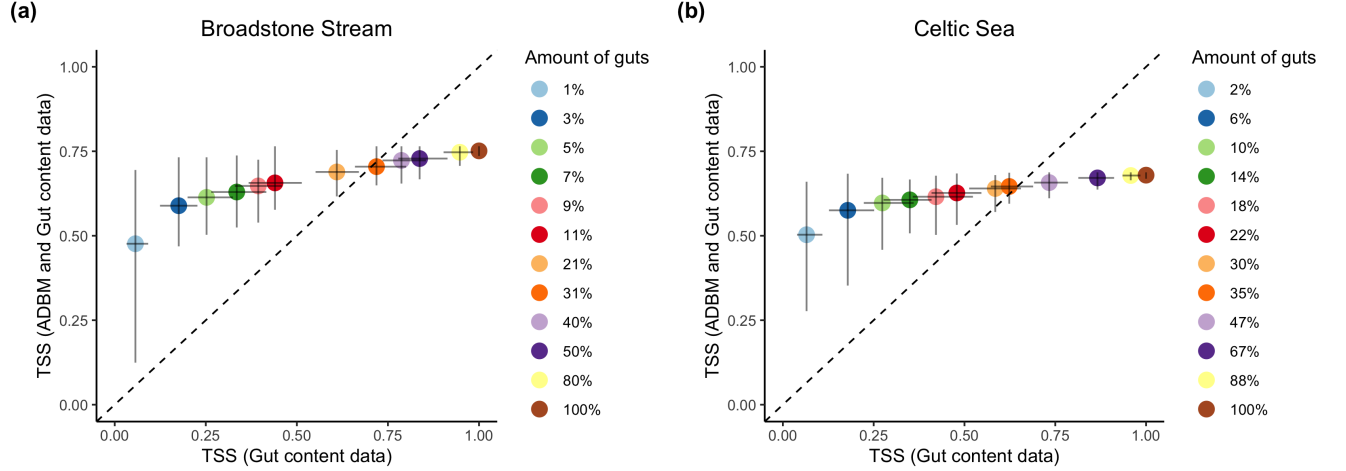
5

Figure 3: True skill statistic between predicted food web using ADBM and incomplete gut content data, and observed food web against the true skill statistic between food web constructed using incomplete gut content data, and observed food web. Error bars represent prediction intervals of 100 independent samples. Dashed line is 1:1 line for reference.
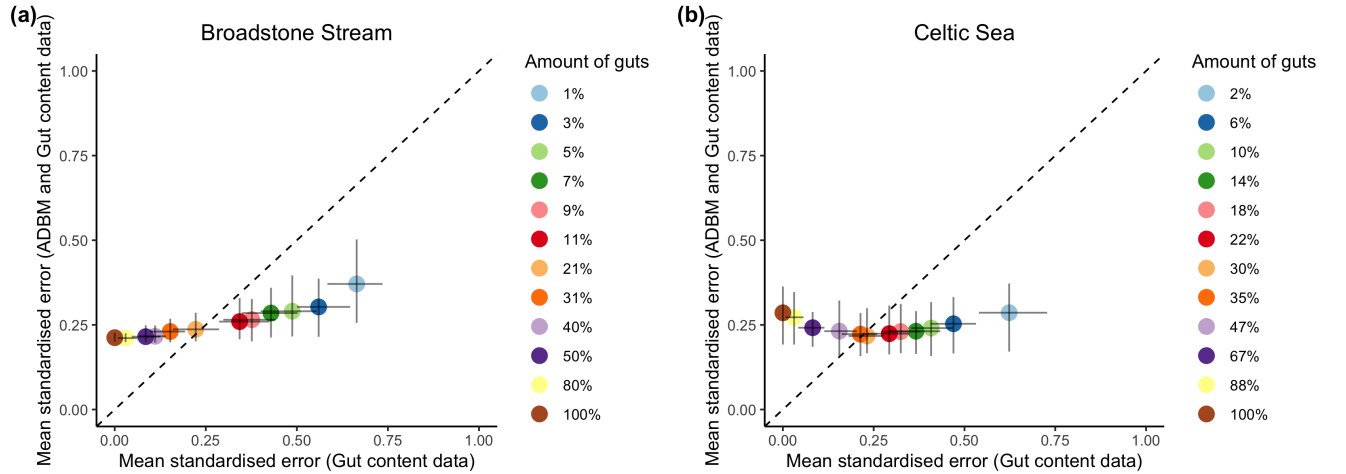

Figure 4: Mean standardised error in structural properties in the food web predicted using ADBM and incomplete gut content data against structural properties in the predicted food web constructed using incomplete gut content data. Error bars represent prediction intervals of 100 independent samples. Dashed line is 1:1 line for reference.

when we use a model as compared to when we would have only used the gut content data. However, at a higher gut content data using the ADBM did not result in better prediction when compared to predictions from only the gut content data. For Broadstone Stream food web, it was better to use the ADBM when we have gut content data less than around 30% (Fig. 3(a)).

4.3 Mean standardised error in food web properties

The ADBM's prediction of structural properties of the food web compared to prediction using only gut content data was better when we used a lower amount of gut content data to parameterise the ADBM (Fig. 4). While at a higher amount of gut content data, the (Fig. 3) structural properties of food web constructed using only gut content data was predicted better when compared to prediction from the ADBM parameterised using the gut content data.

5 Discussion

5.1 Why incomplete gut content data can be used to infer trophic interactions?

- Model fills up the missing interactions
- Partial data helps to constrain the parameter space

6 How much generalisable are our results?

- Dependent on how size structured a food web is, how the food web is aggregated, the type of interactions

7 Assumption that the food web constructed with the total gut content data is the true food web

8 How can we include other food web data type?

- experimentation (feeding trials)
- DNA Metabarcoding

8.1 Extending the approach to other food web models

- Current approach could be implemented with other food web models

9 Conclusion

10 Acknowledgements

11 Author contributions

References

Baker, Ronald, Amanda Buckland, and Marcus Sheaves. 2014. "Fish Gut Content Analysis: Robust Measures of Diet Composition." *Fish and Fisheries* 15 (1): 170–77. `https://doi.org/10.1111/faf.12026`.

Berlow, E. L., J. A. Dunne, N. D. Martinez, P. B. Stark, R. J. Williams, and U. Brose. 2009. "Simple Prediction of Interaction Strengths in Complex Food Webs." *Proceedings of the National Academy of Sciences* 106 (1): 187–91. `https://doi.org/10.1073/pnas.0806823106`.

Dixon, Heather J., J. Brian Dempson, Timothy F. Sheehan, Mark D. Renkawitz, and Michael Power. 2017. "Assessing the Diet of North American Atlantic Salmon (Salmo Salar L.) Off the West Greenland Coast Using Gut Content and Stable Isotope Analyses." *Fisheries Oceanography* 26 (5): 555–68. `https://doi.org/10.1111/fog.12216`.

Dunne, Jennifer A., Richard J. Williams, and Neo D. Martinez. 2002. "Network Structure and Biodiversity Loss in Food Webs: Robustness Increases with Connectance." *Ecology Letters* 5 (4): 558–67. `https://doi.org/10.1046/j.1461-0248.2002.00354.x`.

Eitzinger, Bernhard, Björn C. Rall, Michael Traugott, and Stefan Scheu. 2018. "Testing the Validity of Functional Response Models Using Molecular Gut Content Analysis for Prey Choice in Soil Predators." *Oikos* 127 (7): 915–26. `https://doi.org/10.1111/oik.04885`.

Goldwasser, Lloyd, and Jonathan Roughgarden. 1997. "Sampling Effects and the Estimation of Food-Web Properties." *Ecology* 78 (1): 41–54. https://doi.org/10.1890/0012-9658(1997)078[0041:SEATEO]2.0.CO;2.

Hildrew, Alan G. 2009. "Chapter 4 Sustained Research on Stream Communities." In *Advances in Ecological Research*, 41:175–312. Elsevier. `https://doi.org/10.1016/S0065-2504(09)00404-8`.

Hyslop, E. J. 1980. "Stomach Contents Analysis—a Review of Methods and Their Application." *Journal of Fish Biology* 17 (4): 411–29. `https://doi.org/10.1111/j.1095-8649.1980.tb02775.x`.

Jordán, Ferenc. 2009. "Keystone Species and Food Webs." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1524): 1733–41. `https://doi.org/10.1098/rstb.2008.0335`.

Knight, Tiffany M., Michael W. McCoy, Jonathan M. Chase, Krista A. McCoy, and Robert D. Holt. 2005. "Trophic Cascades Across Ecosystems." *Nature* 437 (7060): 880–83. `https://doi.org/10.1038/nature03962`.

Layer, Katrin, Jens O. Riede, Alan G. Hildrew, and Guy Woodward. 2010. "Food Web Structure and Stability in 20 Streams Across a Wide pH Gradient." In *Advances in Ecological Research*, 42:265–99. Elsevier. `https://doi.org/10.1016/B978-0-12-381363-3.00005-8`.

Layman, Craig A., D. Albrey Arrington, Carmen G. Montaña, and David M. Post. 2007. "Can Stable Isotope Ratios Provide for Community-Wide Measures of Trophic Structure?" *Ecology* 88 (1): 42–48. https://doi.org/10.1890/0012-9658(2007)88[42:CSIRPF]2.0.CO;2.

MacArthur, Robert H., and Eric R. Pianka. 1966. "On Optimal Use of a Patchy Environment." *The American Naturalist* 100 (916): 603–9.

McClain-Counts, Jennifer P., Amanda W.J. Demopoulos, and Steve W. Ross. 2017. "Trophic Structure of Mesopelagic Fishes in the Gulf of Mexico Revealed by Gut Content and Stable Isotope Analyses." Journal Article. *Marine Ecology* 38 (4). `https://doi.org/10.1111/maec.12449`.

McDonald-Madden, E., R. Sabbadin, E. T. Game, P. W. J. Baxter, I. Chadès, and H. P. Possingham. 2016. "Using Food-Web Theory to Conserve Ecosystems." *Nature Communications* 7 (1, 1). Nature Publishing Group: 10245. `https://doi.org/10.1038/ncomms10245`.

O'NAMalley, Brian P., Lars G. Rudstam, James M. Watkins, Toby J. Holda, and Brian C. Weide. 2017. "Effects of Food Web Changes on Mysis Diluviana Diet in Lake Ontario." *Journal of Great Lakes Research*, Special Section on Lake Ontario CSMI 2013, 43 (5): 813–22. `https://doi.org/10.1016/j.jglr.2017.02.003`.

Peralta-Maraver, I., M. J. Lopez-Rodriguez, and J. M. Tierno de Figueroa. 2017. "Structure, Dynamics and Stability of a Mediterranean River Food Web." *Marine and Freshwater Research* 68 (3). CSIRO PUBLISHING: 484–95. `https://doi.org/10.1071/MF15154`.

Petchey, Owen L., A. P. Beckerman, J. O. Riede, and P. H. Warren. 2008. "Size, Foraging, and Food Web Structure." *Proceedings of the National Academy of Sciences* 105 (11): 4191–6. `https://doi.org/10.1073/pnas.0710672105`.

Rindorf, A, and P Lewy. 2004. "Bias in Estimating Food Consumption of Fish by Stomach-Content Analysis" 61: 12.

Roslin, Tomas, and Sanna Majaneva. 2016. "The Use of DNA Barcodes in Food Web Construction—Terrestrial and Aquatic Ecologists Unite!" Edited by Elizabeth Clare. *Genome* 59 (9): 603–28. `https://doi.org/10.1139/gen-2015-0229`.

Schmid-Araya, J. M., A. G. Hildrew, A. Robertson, P. E. Schmid, and J. Winterbottom. 2002. "The Importance of Meiofauna in Food Webs: Evidence from an Acid Stream." *Ecology* 83 (5): 1271–85. https://doi.org/10.1890/0012-9658(2002)083[1271:TIOMIF]2.0.CO;2.

Warren, Philip H. 1989. "Spatial and Temporal Variation in the Structure of a Freshwater Food Web." *Oikos* 55 (3): 299. `https://doi.org/10.2307/3565588`.

Williams, Richard J., and Neo D. Martinez. 2000. "Simple Rules Yield Complex Food Webs." *Nature* 404 (6774, 6774). Nature Publishing Group: 180–83. `https://doi.org/10.1038/35004572`.

Woodward, Guy, Julia Blanchard, Rasmus B. Lauridsen, Francois K. Edwards, J. Iwan Jones, David Figueroa, Philip H. Warren, and Owen L. Petchey. 2010a. "Individual-Based Food Webs." In *Advances in Ecological Research*, 43:211–66. Elsevier. `https://doi.org/10.1016/B978-0-12-385005-8.00006-X`.

———. 2010b. "Chapter 6 - Individual-Based Food Webs: Species Identity, Body Size and Sampling Effects." In *Advances in Ecological Research*, edited by Guy Woodward, 43:211–66. Integrative Ecology: From Molecules to Ecosystems. Academic Press. `https://doi.org/10.1016/B978-0-12-385005-8.00006-X`.

Woodward, Guy, Dougie C. Speirs, and Alan G. Hildrew. 2005a. "Quantification and Resolution of a Complex, Size-Structured Food Web." In *Advances in Ecological Research*, 36:85–135. Elsevier. `https://doi.org/10.1016/S0065-2504(05)36002-8`.

———. 2005b. "Quantification and Resolution of a Complex, Size-Structured Food Web." In *Advances in Ecological Research*, 36:85–135. Elsevier. `https://doi.org/10.1016/S0065-2504(05)36002-8`.