

Effects of data composition and diversity on precision and accuracy of network inference

01 March 2021

Contents

1	Abstract	1
2	Introduction	1
2.1	Why is knowing trophic interactions important?	1
2.2	What are the methods for knowing interactions?	1
2.3	Gut content analysis	1
2.4	Stable isotope analysis	1
2.5	Why inferring interactions is not easy?	2
2.6	How much food web data do we need to estimate the food web structure?	2
2.7	Why do we need to combine these data?	2
2.8	What we do in our study and how?	2
3	Materials and Methods	2
3.1	Overview	2
3.2	Synthesizing gut content data	3
3.3	Synthesizing stable isotope ratios data	3
3.4	Food web data: Broadstone Stream	3
3.5	The Rejection ABC method	3
3.6	True skill statistics	3
4	Results	3
4.1	Prediction using simulated gut content data from a simulated food web	3
4.2	Prediction using real gut content data from the Broadstone Stream food web	5
4.3	Prediction using simulated stable isotope ratio data from a simulated food web	7
4.4	Combining both gut content data and stable isotope ratios data	7
5	Discussion	7
5.1	Include digestion in the gut content simulation	7
5.2	On generalising ADBM's predictions over all food webs	7
5.3	Include other food web data type	7
5.4	Extension to other food web models	7
6	Conclusion	7
7	Appendix	7
	References	7

1 Abstract

2 Introduction

2.1 Why is knowing trophic interactions important?

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

2.2 What are the methods for knowing interactions?

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

2.3 Gut content analysis

Gut content analysis is performed to know what an individual has eaten. The items in the stomach can hence be used to infer the diet of that species. 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).

2.4 Stable isotope analysis

Stable isotope analysis (SIA) of tissues from a single individual can be used to quantify its diet over varying temporal and spatial scales from one sample (Crawford, McDonald, and Bearhop 2008). Using SIA, several different sources of information can be extracted from a single sampling event as different tissues integrate diet over different temporal scales as well as spatial scales. However, there is uncertainty involved in this method which is due to variability in the isotopic fractionation values across multiple combinations of diets and tissues/species, unquantified temporal or spatial variation in prey isotopic values and variation caused by routing of particular dietary nutrients into particular tissues (Crawford, McDonald, and Bearhop 2008).

2.5 Why inferring interactions is not easy?

Since a food web can be quite complex with many species and trophic interactions, inferring these trophic interactions with high accuracy and less uncertainty would require too much food web data (citreqd). Also, it might not be feasible to use a single method to know the structure of the complete food web. E.g.: Stable isotope analysis from tissues can only be used to infer the approximate trophic position of the species which might not be sufficient to predict the food web structure. Gut content data might only result in the partial diet of a predator individual as gut content data gives us the information of the prey species that have been foraged by the predator recently.

- *Note:* Needs references

2.6 How much food web data do we need to estimate the food web structure?

Do we need the information of the gut of all the species to infer the food web structure? One could ask a similar question with stable isotope ratios data as well. The question we would like to address here is how should the sampling effort be distributed among the species for a given food web method.

- *Note:* The amount of sampling effort affects how well the food web properties are predicted (Goldwasser and Roughgarden 1997).

- *Note:* Needs references

2.7 Why do we need to combine these data?

These methods of inferring prey predator interactions require a lot of sampling effort (citreqd). E.g. One needs to know the information of the gut of all the species to know the feeding links between species in a food web. It might not be possible to know the complete structure of a food web by using a single method only. E.g. Stable isotope ratios of tissues can only give the approximate trophic position of species which might not be sufficient enough to know the complete food web structure.

- *Note:* This paragraph is a bit redundant with the paragraph “Why inferring interactions is not easy?”

2.8 What we do in our study and how?

These questions are addressed *in silico* by simulating food web networks whose true parameters are known. We used the allometric diet breadth model (ADBM) to simulate the food webs and approximate Bayesian computation (ABC) to parameterise the ADBM. We simulated the food web data which were used to parameterise the ADBM using ABC. We also used real gut content data from Broadstone Stream food web to address those questions.

- *Note:* Why did we choose gut content analysis and stable isotope analysis only?
- *Note:* Why ADBM and ABC?

3 Materials and Methods

In the upcoming sections, we present ...

3.1 Overview

First, we addressed these questions using *in silico* (Fig. 1) by simulating food web structure from the allometric diet breadth model (ADBM) whose true parameters were known. We then synthesized gut content data and stable isotope ratios data from the simulated food web. In the next step, we computed different quantities and composition of food web data and used them to predict the food web using the ADBM which was parameterised using the approximate Bayesian computation (ABC). We also used a real gut content data and a real stable isotope ratios data from the Broadstone Stream food web to predict the food web.

- *Note:* Uncertainty

- *Note:* ‘in silico’ to be in italics in the caption
- *Note:* Remove direct observation from the flowchart

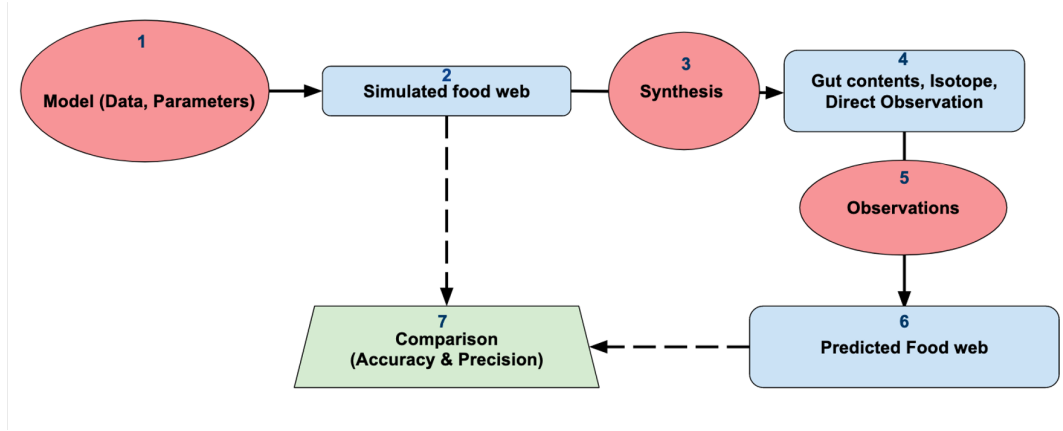


Figure 1: Flowchart depicting the in silico approach implemented for investigating the effect of quantity and composition of food web data in predicting food web structure using the ADBM.

3.2 Synthesizing gut content data

3.2.1 An example of synthesized gut content data

3.3 Synthesizing stable isotope ratios data

3.3.1 An example of synthesized stable isotope ratio data

3.4 Food web data: Broadstone Stream

- *Note:* Description of the Broadstone food web

3.5 The Rejection ABC method

3.6 True skill statistics

4 Results

We first present the ...

4.1 Prediction using simulated gut content data from a simulated food web

The true skill statistics (TSS) between the predicted food web and observed food web saturated with an increasing number of distinct predator guts in a sample (Fig. 2). The TSS of the predicted food webs estimated using the complete individual gut data had shorter prediction intervals resulting in less uncertainty, and higher mean TSS than that using the partial individual gut data. The maximum limit of the prediction interval of TSS estimated using the complete gut data and the partial gut data were almost equal, with the minimum limit of the prediction interval of TSS using partial gut data being lower than that from the complete gut data. Eventually, the gap between the mean TSS using the partial gut data and the complete gut data reduced with an increasing number of distinct predator guts suggesting when we have enough predator species' gut data, the achieved TSS was almost constant and hence independent of the amount of gut data.

The maximum TSS estimated using the complete gut data was very close to one and almost remained constant with an increasing number of different predator species sampled. With the gut data sample of only five distinct predator species, 95% of the maximum mean TSS was achieved when complete individual gut data was used, while the same was achieved with 15 predator species for partial gut data. This shows that one does not need to know the gut data of all the species to predict the food web and the accuracy is dependent on the completeness of an individual gut data.

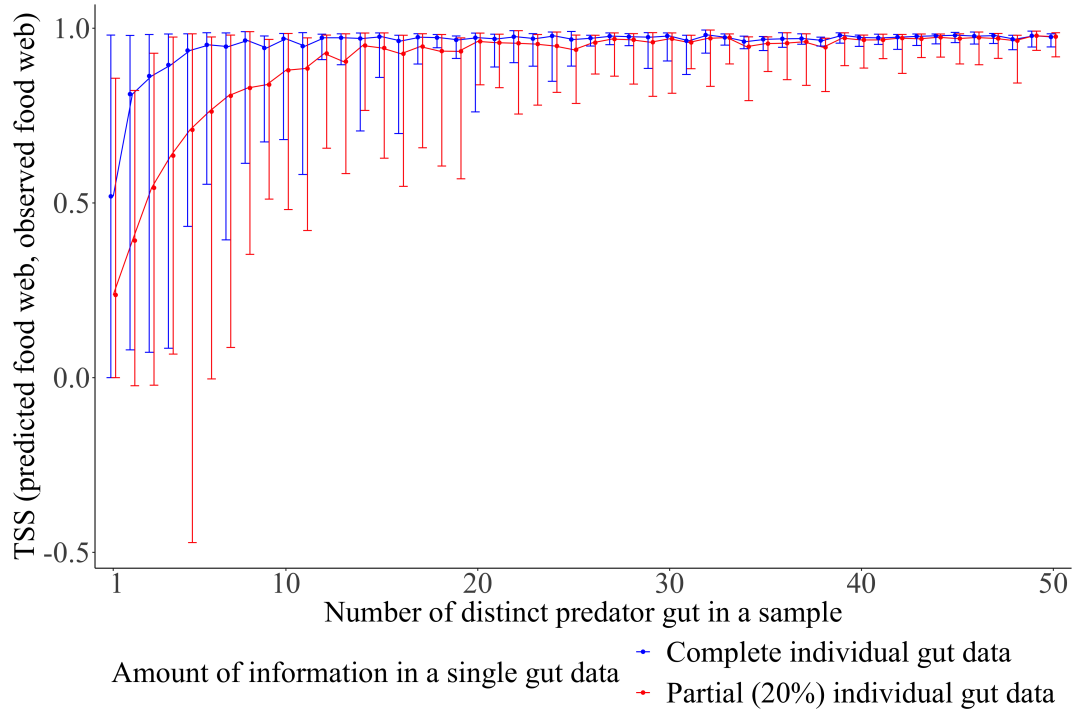


Figure 2: True skill statistics between predicted food web and observed food web for a simulated small reef food web estimated for distinct predator guts in a sample. The observed simulated food web consists of 50 species and ... links. The vertical bars correspond to the prediction intervals of the true skill statistics with filled circles representing the corresponding mean. A prediction interval of the TSS is formed using a set of 100 accepted TSS values using the ABC method.

4.2 Prediction using real gut content data from the Broadstone Stream food web

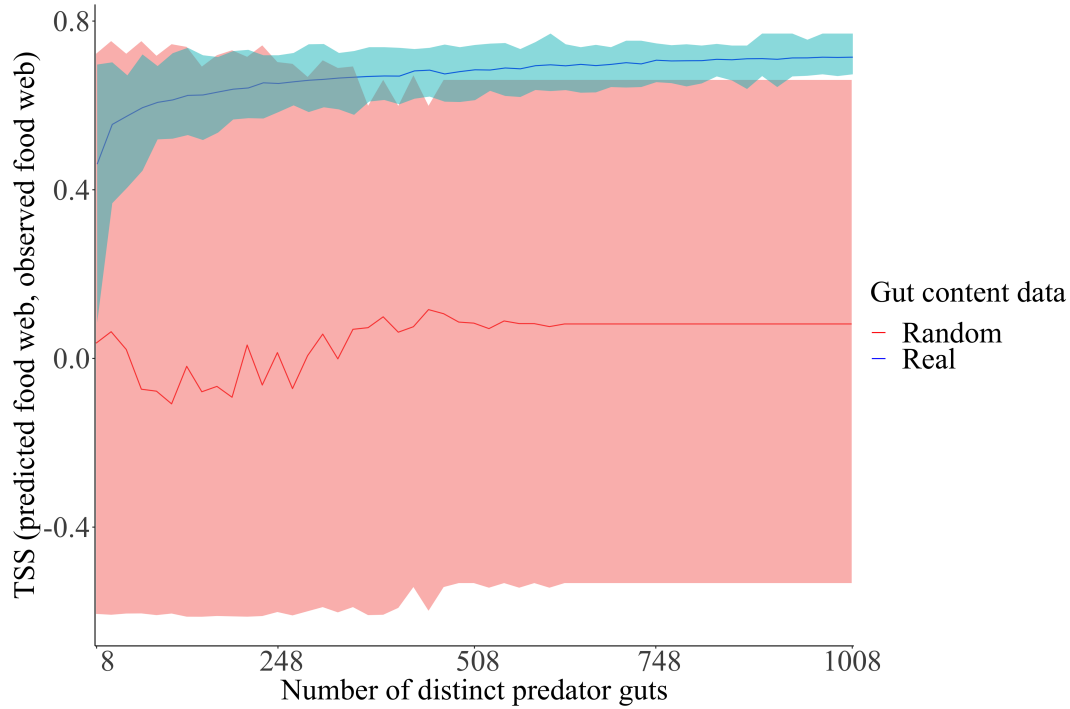


Figure 3: True skill statistics between the predicted and observed Broadstone Stream food web computed for number of different predator guts sampled. There were a total of 1008 guts from 29 species. Line represents the mean and shaded region represents the range/uncertainty. Blue and red curve/region correspond to real and randomly generated gut content data respectively.

The value of true skill statistics between the predicted food web and observed food web saturated quickly with the increasing number of distinct real predator guts (Fig. 3). With only 228 real predator guts, the ADBM parameterised with ABC predicted the food web with an accuracy of 90% of the maximum mean TSS (0.71). The uncertainty (red shaded region) in the TSS reduced with the increasing number of predators' guts. Although at a low number of predators' guts the maximum TSS was quite high the uncertainty was high as well. This was quite possibly just by chance which was suggested by the overlap of the TSS region with that of randomly generated gut content data.

True skill statistics between the predicted food web and the observed food web had a saturating curve with the increasing number of distinct predator guts (Fig. 4) for both sets of gut predator guts: First set of gut data consisted of predator body sizes which were smaller than the mean body size and second set of gut data consisted of predator body sizes which were larger than the mean body size. However, the mean TSS corresponding to the guts of predators smaller than the mean body size was higher when compared to that from the guts of predators larger than the mean body size.

- *Note:* Compare the uncertainty (prediction intervals) between both set of gut data in Fig. 4.

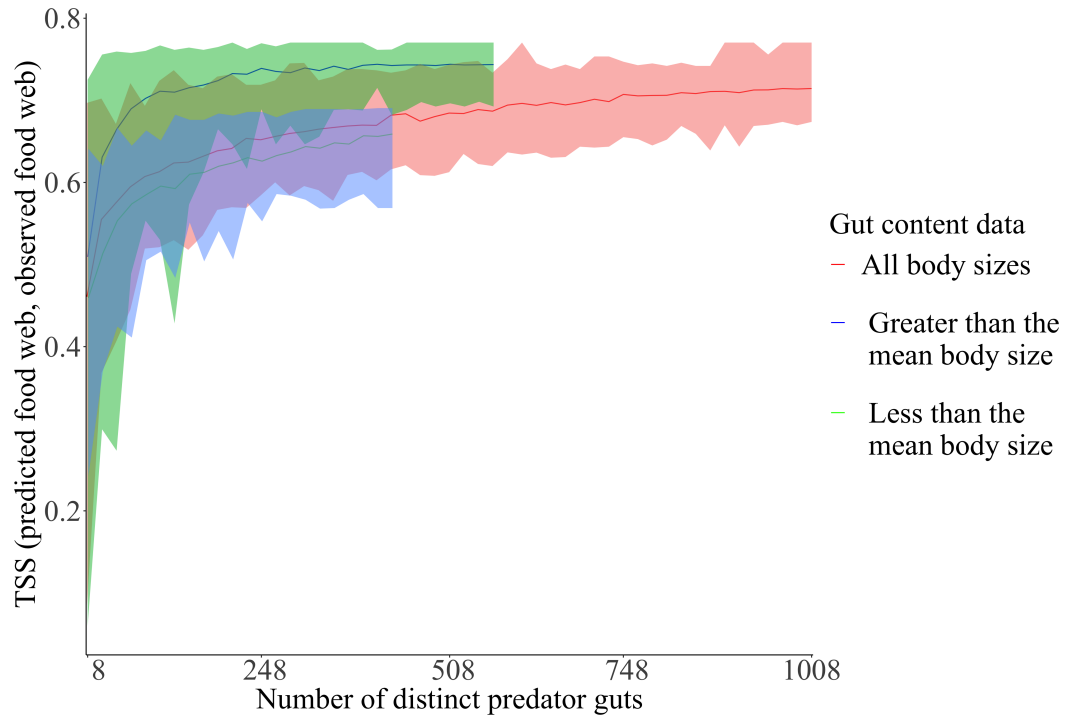


Figure 4: True skill statistics between predicted food web and observed food web estimated for different number of distinct predator guts. The estimation is done for three sets of gut data: gut content data of predators whose body sizes are smaller than the mean body size, larger than the mean body size, and all the gut content data.

4.3 Prediction using simulated stable isotope ratio data from a simulated food web

4.4 Combining both gut content data and stable isotope ratios data

5 Discussion

5.1 Include digestion in the gut content simulation

5.2 On generalising ADBM’s predictions over all food webs

5.3 Include other food web data type

- *Note:* experimentation (feeding trials)
- *Note:* DNA Metabarcoding

5.4 Extension to other food web models

6 Conclusion

7 Appendix

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>.
- Crawford, Kerry, Robbie A. McDonald, and Stuart Bearhop. 2008. “Applications of Stable Isotope Techniques to the Ecology of Mammals.” *Mammal Review* 38 (1): 87–107. <https://doi.org/10.1111/j.1365-2907.2008.00120.x>.
- 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>.
- 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](https://doi.org/10.1890/0012-9658(1997)078[0041:SEATEO]2.0.CO;2).
- 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>.
- Layman, Craig A., D. Albrecht 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](https://doi.org/10.1890/0012-9658(2007)88[42:CSIRPF]2.0.CO;2).
- 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>.
- 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>.
- 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>.

¹⁹⁷ Warren, Philip H. 1989. "Spatial and Temporal Variation in the Structure of a Freshwater Food Web."
¹⁹⁸ *Oikos* 55 (3): 299. <https://doi.org/10.2307/3565588>.