
OPTIMAL AMOUNT OF GUT CONTENT DATA REQUIRED TO PREDICT TROPHIC INTERACTIONS USING A FOOD WEB MODEL

A PREPRINT

Anubhav Gupta *

Department of Evolutionary Biology and Environmental Studies
University of Zurich
8057 Zurich, Switzerland
anubhav.gupta@ieu.uzh.ch

Eoin O' Gorman

School of Life Sciences
University of Essex
CO4 3SQ Colchester, UK
e.ogorman@essex.ac.uk

Owen L. Petchey

Department of Evolutionary Biology and Environmental Studies
University of Zurich
8057 Zurich, Switzerland
owen.petchey@ieu.uzh.ch

February 7, 2022

Abstract

- 1) Presence absence data of species interactions such as gut content data is used to parameterise food web models such as allometric diet breadth model (ADBM), thereby predicting interactions. And collecting and analysing these data from the field is an expensive and time consuming task. Therefore, it is crucial to know how much gut content data is required to parameterise the ADBM with high accuracy and high precision for a given food web with minimum sampling effort.
- 2) The novelty of our current approach is that we use approximate Bayesian computation to parameterise the ADBM, and true skill statistics to measure the goodness of fit for different amount of predators' gut content data.
- 3) We estimated the minimum number of gut content data required for seven different food webs that resulted in 95% of the maximum true skill statistics achieved by the ADBM for total gut content data. We found that incomplete gut content data can be used to

*Corresponding author

parameterise the ADBM with the lowest amount of gut content data being 28% of the total gut content data.

- 4) These results suggest that one need not collect a large quantity of gut content data to predict a food web structure using a food web model, thereby reducing sampling effort considerably.

Keywords gut content data · ADBM · food web accuracy · food web prediction

AG: My comment starts with “AG:”

1 Alternate titles

- Predicting trophic interactions using incomplete gut content data
- Effect of the amount of gut content data on the accuracy and precision of food web prediction

2 Introduction

Knowledge about the trophic interactions in a food web is crucial in ecology for purposes ranging from identifying keystone species (Jord’an 2009) to quantifying robustness of a food web to species extinctions (Dunne, Williams, and Martinez 2002). This has led to the development of numerous food web models (Allesina, Alonso, and Pascual 2008; Cohen, Newman, and Steele 1985; Gravel et al. 2013; Petchey et al. 2008; Tamaddoni-Nezhad et al. 2013). Along with inferring missing links in an observed food web, the food web models are also used for ecological forecasting (Hattab et al. 2016; Lindegren et al. 2010) and for understanding the underlying mechanism governing the interactions in a food web (O’NAGorman et al. 2019).

Although food web models are constructed using prior theory about the factors that determine trophic interactions, empirical data about interactions is required to parameterise a model. For example, Petchey et al. (2008) used presence-absence information about trophic interactions to parameterise the allometric diet breadth model and thereby predict species interactions. These empirical information about interactions can be inferred from diverse set of methods such as gut content analysis (Peralta-Maraver, Lopez-Rodriguez, and de Figueroa 2017), stable isotope ratio analysis of tissues (Layman et al. 2007), experimentation (Warren 1989), DNA metabarcoding (Roslin and Majaneva 2016) and literature research (Gray et al. 2015; Cohen and Mulder 2014a; Goldwasser and Roughgarden 1993a).

Food web methods mentioned above have multiple shortcomings. Stable isotope ratio analysis of the organism’s tissue does not give a direct information of the diet of that organism but rather approximate trophic position of that species in the food web. Although mixing models can be used to determine what prey items are most likely fed upon by a predator but this results in uncertainty in the estimates (Kadoya,

Osada, and Takimoto 2012; Crawford, McDonald, and Bearhop 2008). Experimentation (e.g. feeding trials) may create unrealistic conditions where a predator is confined in a small space with a potential “prey” and is effectively “forced” into eating it (citreqd). DNA metabarcoding could have many issues such as failure to relate OTUs to species, or inability to distinguish secondary predation, i.e. the DNA signature of the prey consumed by the prey consumed by the primary item found in the predator’s gut. Another approach to infer missing links is literature research which involves assuming that a link will occur because another author described it in some other system, or for related predator/prey species, e.g. from the same genus or family (Gray et al. 2015; Cohen and Mulder 2014b; Goldwasser and Roughgarden 1993b). This typically results in a food web that will be very oversampled with false positives in the presence of links (citreqd). Nielsen et al. (2018) have shown that gut content analysis method has a better match with real diet when compared to other methods. **AG: Working on this paragraph.

Acquiring food web data from gut content analysis is time consuming and expensive (reference). In part this is due to the need for many gut contents to be collected and analysed in order to be confident that the majority of possible trophic links have been observed. Furthermore, considerable time, resource investment, and expertise is required to process the collected samples. Hence, there is of considerable importance to know the minimum number of gut contents and of which consumer species that are required to parameterise a food web model with high accuracy and high precision, so that one knows when is a good time to stop collecting any further empirical data on the food web.

Therefore, the key question we are interested in answering is how much presence-absence information, in the form of gut content data, is required to infer food web structure from a food web model specifically the ADBM with high accuracy and high precision. In other words, how many samples of gut content should one collect from the field to parameterise a food web model? To answer this question, we use the presence-absence information from gut content data to parameterise the allometric diet breadth model (ADBM) thereby predicting trophic interactions in seven different food webs using rejection approximate Bayesian computation, where the fitness of the model was measured using true skill statistics. We computed the true skill statistics of the predicted food webs for different number of predator guts to calculate the minimum number of gut content data to infer food web structure. Our study provides a guideline on how many gut content data is required to predict food web structure using a food web model.

3 Materials and Methods

In the upcoming sections, we present the food webs, the allometric diet breadth model (ADBM) and the gut content data used to infer the trophic interactions in the presented food webs using the ADBM. We also give a detailed account of using partial gut content data to parameterise the ADBM using rejection

approximate Bayesian computation (ABC). We assessed model predictions using the true skill statistic and various structural food web properties for comparison across food webs.

3.1 The Empirical Food Webs

In our study, we used food webs for which gut content data is available at an individual level and, is FAIR. Presence and absence interactions of food web data can be aggregated in different ways (Gilljam et al. 2011). A common way of aggregating food web data is the taxonomic approach, in which nodes in the food web are taxa (e.g. species) and one species is said to feed on another species if at least one individual of the latter is found in at least one gut of the former. Another approach is based on size, rather than taxonomy, and in this the nodes in the food web are size classes. Here, a feeding link occurs between two size classes if at least one prey item within a size class was found in the gut of another size class, irrespective of their taxonomy.

In our study, we aggregate the gut content data on the basis of size class because the food webs constructed using this approach were better predicted by the ADBM compared to the taxonomic approach (Woodward et al. 2010a), and this is because the ADBM is based on foraging rules defined using the body sizes of organisms (Petchey et al. 2008).

Our study is based on the assumption that the food web constructed using the complete gut content data is the true food web. However, this might not be the case because the links are still under-sampled for many nodes (Woodward et al. 2010b). This could result in a lower TSS of the predicted food web from the ADBM which can result in a biased estimate of the number of gut content data.

Table 1: Information about the empirical food webs. **AG: Working on this.

Common food web name (Original Publication)	Predation matrix source	Body size source	General ecosystem	Number of nodes	Number of links	Connectance	Body size range (approximate)	Type of interactions
Broadstone Stream (Woodward et al. 2010)	NA	NA	Freshwater	29	185	0.24	10^{-7} to 10^2	Predation
Celtic Sea	NA	NA	Marine	48	386	0.17	10^{-2} to 10^4	NA
Tadnoll Brook	NA	NA	Freshwater	59	485	0.14	10^{-6} to 10^5	NA
Afon Hirnant	NA	NA	Freshwater	33	221	0.20	10^{-6} to 10^2	NA
Coilaco	NA	NA	Freshwater	45	123	0.06	10^{-6} to 10^2	NA
Guampoe	NA	NA	Freshwater	44	139	0.07	10^{-6} to 10^3	NA
Trancura	NA	NA	Freshwater	35	78	0.06	10^{-6} to 10^1	NA

3.2 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). We chose this model because it uses the body size of organisms as input to predict feeding interactions. 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 are: 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. Further details on the foraging rules defined in the ADBM and ADBM’s predictive power across different food webs can be found in Petchey et al. (2008).

3.3 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:

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

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 (Allouche, Tsoar, and Kadmon 2006).

3.4 Inferring food web using partial gut content data

From an empirical dataset of gut content data, we draw a set of gut contents data randomly to create a partial gut content dataset. We then fit the ADBM to this partial dataset.

To fit the ADBM to a gut content dataset (be it complete or partial), we used the rejection ABC method we previously developed in Gupta and Petchey (2021) to accept a parameter value from a prior distribution which would have resulted in the minimum distance, where distance = $1 - \text{True skill statistic}$. The true skill statistic was computed between the predators’ diet predicted from

the ADBM, and that from 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 $P : P = \{p_1, p_2, \dots, p_k\}$
- A pool of gut content data $G : G = \{g_1, g_2, \dots, g_n\}$, where g_n is a one-dimensional matrix containing ones and zeros.
- A model prediction $model(\theta) : ADBM(\theta) = \{d_{p_1}, d_{p_1}, \dots, d_{p_k}\}$, 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 $Y : Y = \{d'_{p_1}, d'_{p_1}, \dots, d'_{p_k}\}$

Sampling:

- for $i = 1, \dots, tgut$ where $tgut$ is the total number of gut content data in the pool G
- for $j = 1, \dots, nsample$ where $nsample$ is the number of independent samples to be drawn
 - Draw a set of gut content data $y = \{g_1, g_2, \dots, g_i\}$ from the pool of gut content data G
 - for $k = 1, \dots, npar$ 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 $TSS_i(x, y) = \{TSS(x_i, y) : x_i = ADBM(\theta_j), \theta_j \text{ computed from previous step}\}$ using the accepted $\theta_1, \dots, \theta_{nsample}$

Output:

The TSS between observed and predicted food webs for every i number of gut content data drawn from the pool of gut content data.

144 4 Results

145 We first present how the accuracy of the food web model i.e ADBM vary in predicting trophic
 146 interactions with increasing amount of gut content data provided to the food web model. Then,
 147 we calculate the number of gut content data for each food webs that results in accuracy equal to
 148 95% of the maximum true skill statistics predicted by the food web model. We also compute the
 149 goodness of the food web model in predicting structural properties from incomplete gut content
 150 data. We did that by computing mean standardised error in food web properties.

151 4.1 Inferring trophic interactions using ADBM and incomplete gut content data

152 The true skill statistics of the food webs predicted by the ADBM using subsets of gut content data
 153 improved quickly for lower number of predator guts (Fig. 1) (a, c, e, g, i, k, m). The prediction
 154 interval of the true skill statistics decreased with increasing number of predator guts with the mean
 155 TSS asymptoting to the maximum mean TSS achieved by the ADBM when all the gut content data
 156 was used. Although the maximum TSS varied among the food webs, the qualitative increase in the
 157 TSS was the same.

158 For Broadstone Stream food web, with only 381 number of gut content data which is 38% of
 159 the total gut content data, the ADBM predicted the food web with the TSS of 0.74 which was 95%
 160 of the mean TSS (0.78) achieved using complete gut content data by the ADBM (Fig. 1(a)). In case
 161 of the Celtic Sea food web, only 171 gut content data which is 35% of the total gut content data
 162 was required by the ADBM to predict food web with TSS equal to 95% of the mean TSS (0.68)
 163 achieved using complete gut content data by the ADBM (Fig. 1(c)).

164 4.2 TSS(Model)/TSS(Data) for different food webs

165 For a low number of predator guts, the TSS of the model predicted food web was higher than the
 166 TSS of the food web constructed from gut content data only (Fig. 1 (b, d, f, h, j, l, n)). After a
 167 certain increase in the number of predator guts, the ratio TSS(Model)/TSS(data) reduced to less
 168 than one, and reduced further gradually.

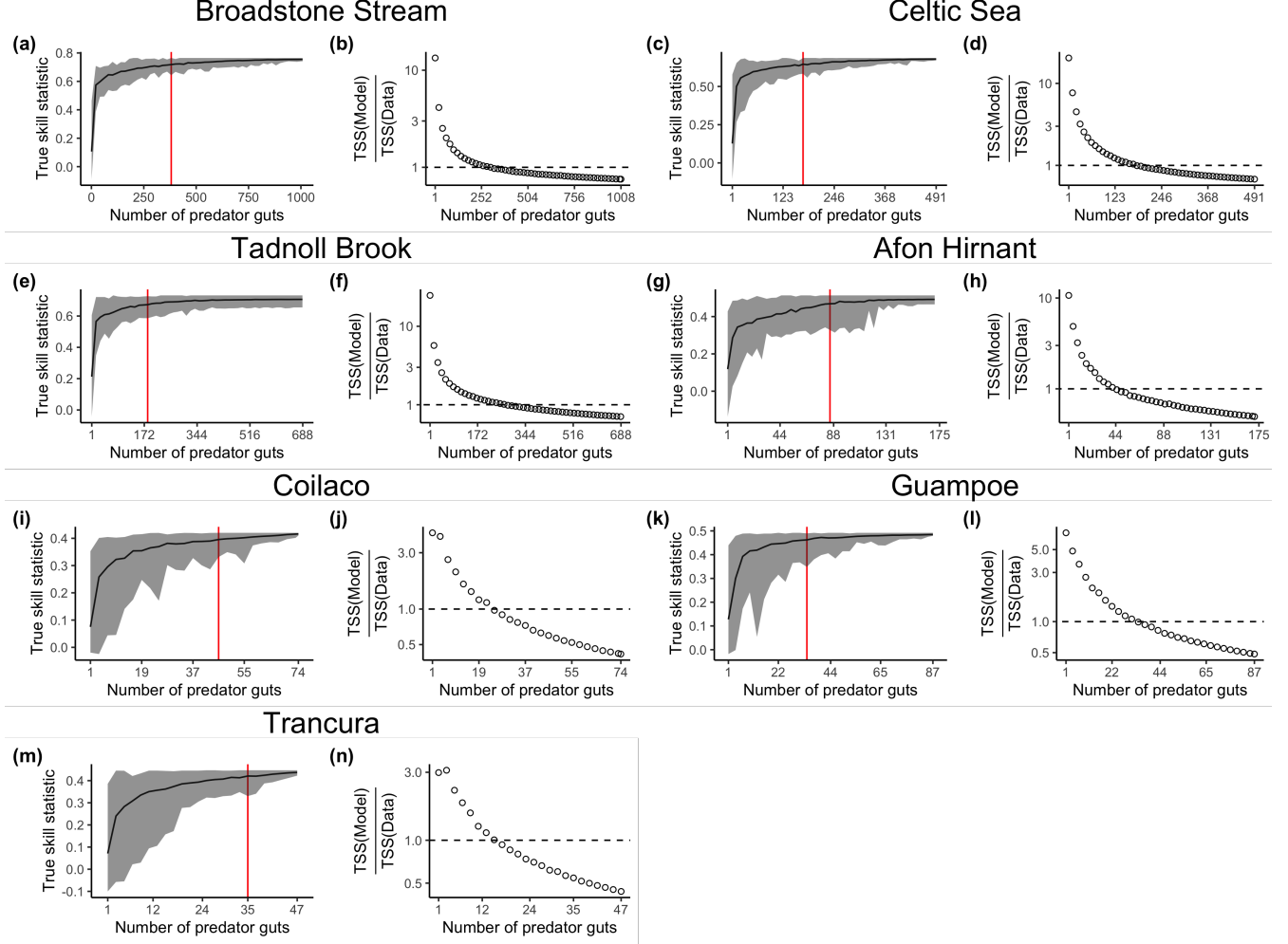


Figure 1: (a, c, e, g, i, k, m) Accuracy of the predicted food web measured using the true skill statistics, predicted by the ADBM parameterised using gut content data. Line and shaded grey region represents the mean and the prediction interval of 100 independent samples respectively. Red line represents the number of predator guts required to achieve a TSS of 95% of the maximum TSS. (b, d, f, h, j, l, n) Ratio of the true skill statistic of the predicted food web by the ADBM parameterised using gut content data to that of the predicted food web constructed using gut content data only. The dashed line shows where the ratio is one for reference.

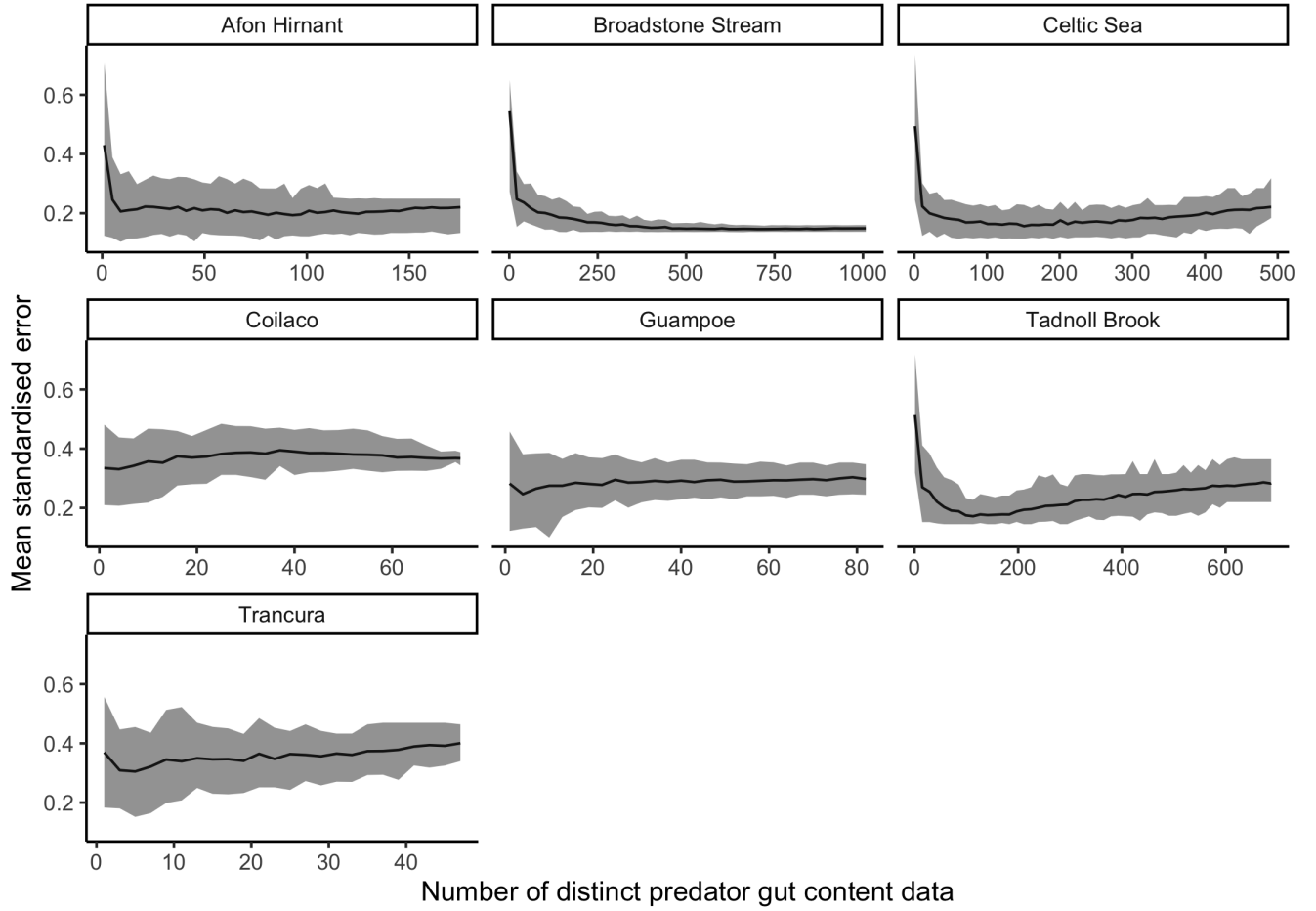


Figure 2: Mean standardised error in structural properties of different food webs with respect to the number of gut content data.

4.3 Mean standardised error in structural properties with respect to the number of gut content data

Mean standardised error in structural properties of afon hirnant, broadstone stream, celtic sea and tadnoll brook food web reduced drastically with increasing number of predator gut content data at low number of predator gut content data (Fig. 2). In case of Broadstone Stream food web, the 95% prediction interval reduced as well with increasing number of predator gut content data.

**AG: Working on this paragraph.

4.4 Dependence of minimum gut content data on number of species, links and connectance

There was a positive trend between number of minimum gut content data and number of species (S) (Fig. 3 (a)), links (L) (Fig. 3 (c)), connectance (C) (Fig. 3 (d)) and number of predator nodes (Fig. 3 (e)), with Broadstone Stream being an outlier in Fig. 3 (a, c, e). The number of minimum gut content data increases with number of maximum gut content data sampled (Fig. 3 (b)). There seems to be a negative relationship between proportion of minimum gut (= Number of minimum gut / Number of maximum guts sampled) and number of species (S) (Fig. 3 (f)), links (L) (Fig. 3 (g)), connectance (C) (Fig. 3 (h)).

5 Discussion

We have demonstrated how a food web model can be used to predict food web in the availability of incomplete data. At lower amount of gut content data, ADBM is quite better at predicting trophic interactions. We have computed the upper limit of amount of gut content data till when it can be used to compensate for undersampling. This can help us decide how much gut content data to actually collect when we are using a food web model to infer trophic interactions. A future development could be to use this our approach on other food web models, and to also use food web data other than gut content data to parameterise those models.

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

In all of the seven food webs, the ADBM was able to infer the trophic interactions using incomplete gut content data because the presence absence information from the gut content data was used to narrow down the possible model parameter values of the ADBM that best explained the predators' diets. Although in theory the ADBM can predict trophic interactions using only body sizes of organisms as it is based on set of foraging rule, it still requires some presence absence data to constrain the posterior parameter space thereby making more accurate predictions.

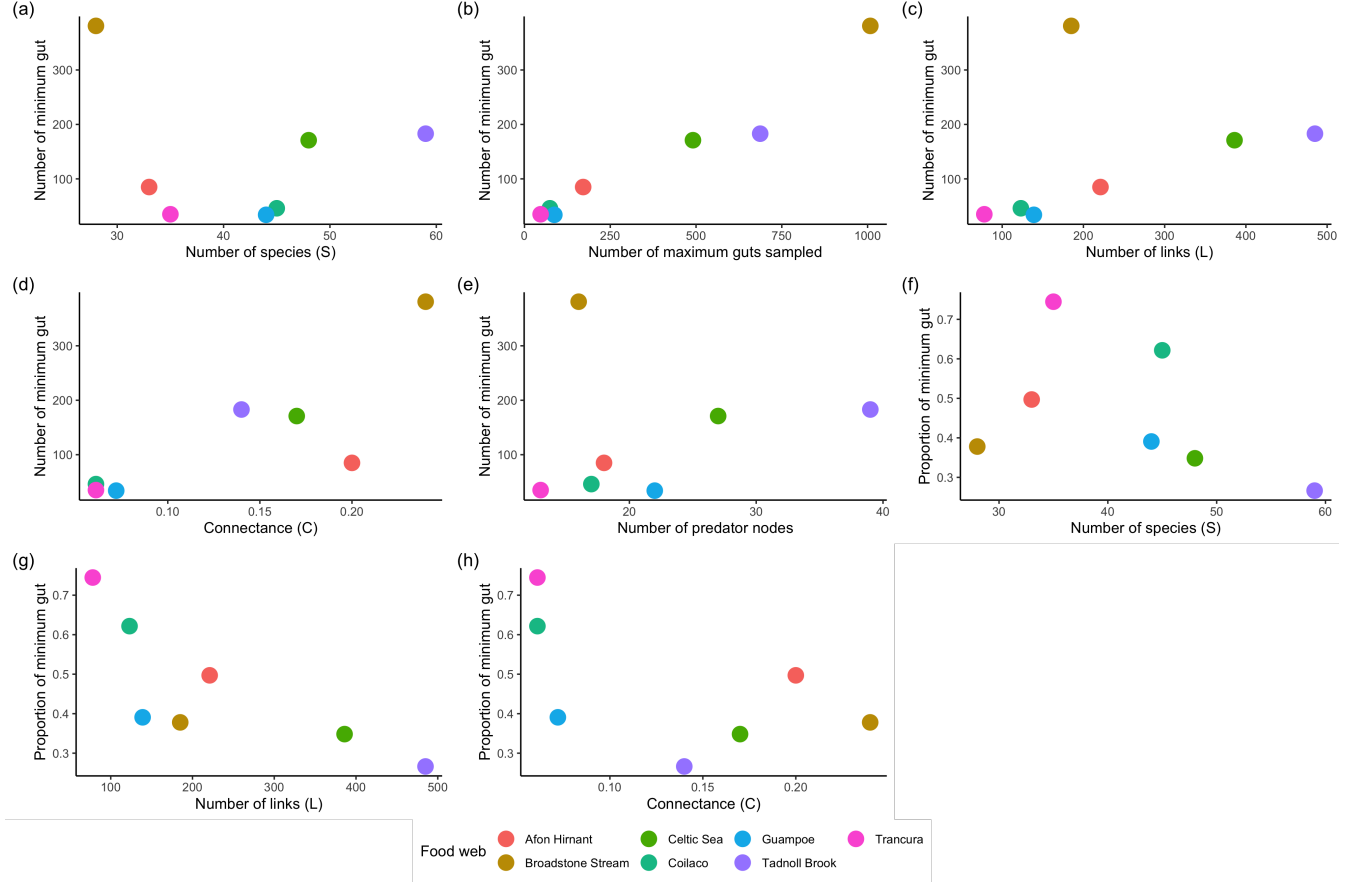


Figure 3: (a, b, c, d, e) Number of minimum gut content data plotted against number of species (S), number of maximum guts sampled, number of links (L), connectance (C) and number of predator nodes respectively. (f, g, h) Proportion of minimum gut content data plotted against number of species (S), number of links (L) and connectance (C) respectively.

5.2 Why does the minimum gut data differ among food webs?

We found the minimum number of gut content data to vary across food webs and this could be due to the possibility that the food webs are sampled differently i.e. some food webs are well sampled than others (Fig. 4). Since some food webs are more size-structured than others, they are well predicted by the ADBM and therefore might require a lower number of gut content data for model parameterisation. **:AG: I would need to add information about the interactions in these food webs.

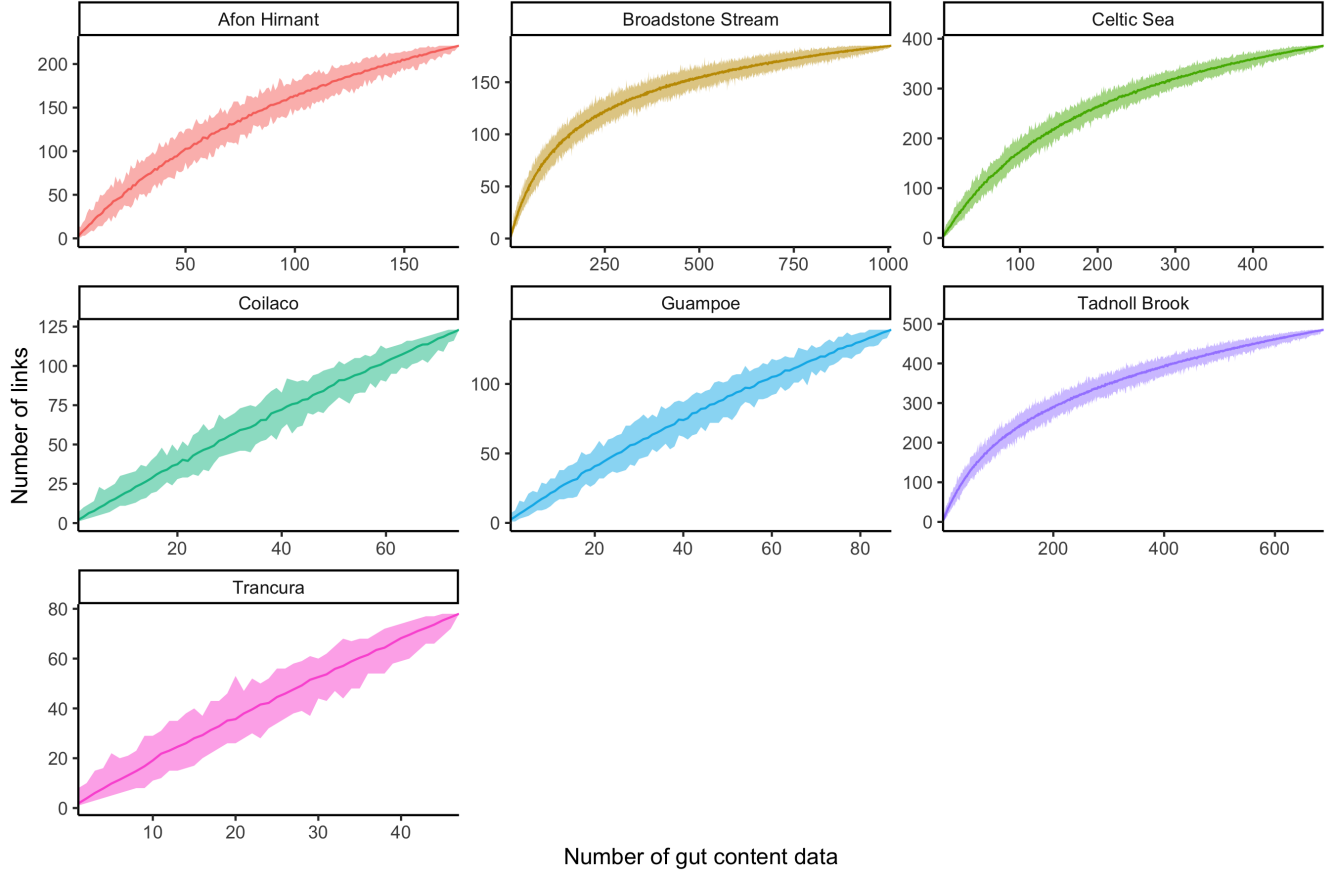


Figure 4: Link accumulation curve of the food web constructed using gut content data plotted against number of gut content data.

5.3 Model fills up missing interactions and guide sampling direction

When only few presence absence data is available, model can be used to fill up the missing interactions that has not been observed in the gut content data. Also model's predictions can be used as a guide to sample predator guts that might be significant in the food web.

5.4 Why is model's TSS less than one?

Due to limitations of the ADBM as to not able to completely explain the interactions in a food webs using the foraging rules based on the body size and also possible undersampling in the observed

214 data because of missing rare links, the ADBM's maximum predicted true skill statistics is always
 215 less than one.

216 The maximum predictive power of a model is constrained by the rules it is based on. In the
 217 case of ADBM, that limits the predictive power of the ADBM is that it can only predict diets that
 218 are contiguous with respect to the size of prey. I.e. it cannot predict that a predator with consume
 219 prey of size 1 and 3, and not prey of size 2. Also, it is important to note that if the observed diets
 220 are not contiguous when prey are ordered by their size, the estimation process could lead to a lower
 221 value of the TSS.

222 **5.5 TSS(Model)/TSS(Data) for different food webs**

223 **AG: Working on this.

224 **5.6 Why Broadstone Stream is an outlier?**

225 Broadstone Stream is the most sampled food web compared to all the other food webs (Fig. 4),
 226 which means the observed Broadstone Stream food web might be more accurate compared to others.

227 **AG: Working on this, not sure if this paragraph is needed.

228 **5.7 Propagating uncertainty from gut content data to model's predictions**

229 **OP Focus on this, at least in the context of further work that could be done. ## There is
 230 uncertainty in gut content data and needs to be propagated into model's predictions

231 There is considerable uncertainty involved in gut content analysis (Baker, Buckland, and
 232 Sheaves 2014) such as in fish's guts, there are sometimes loose tissues that are not identifiable and
 233 cannot be assigned to a specific prey item with certainty. There are factors such as sample size of
 234 consumers, mechanical prey handling, differential digestion and evacuation rates of different prey
 235 types and volumes, and the ingestion order that in combination result in an unquantifiable error
 236 which is difficult to interpret in the predator diet (Hyslop 1980; Rindorf and Lewy 2004; Baker,
 237 Buckland, and Sheaves 2014). Therefore, a future development would be to consider how this
 238 uncertainty is propagated into food web prediction.

5.8 Comparing our study with other studies of predicting food webs using incomplete data

**OP: Focus on this, and other studies that have looked at the importance of quantity and quality of food web data ## Comparing our study with other studies of predicting food webs using incomplete data

Some studies have presented how the accuracy of food web prediction change when we vary the amount of food web data (Gray et al. 2015). **AG: Working on this paragraph.

5.9 What is the generalisability of our results?

**OP: Focus on this.

In our study, we have implemented the approach to only the ADBM. However, in principle, this approach could be extended to other food web models. A future prospect could be to study how good different food web models' prediction accuracy vary for different amount of gut content data. This can also help in making decision as to which food web model to chose from given set of gut content data.

A future study could be to use our approach of estimating the minimum gut content data for different food web models. An efficient model would be the one that uses minimum available gut content data to predict species interactions with high accuracy and high precision.

5.10 Assumption that the food web constructed with the total gut content data is the true food web and how can this be improved

Some of the food webs are not very well predicted using the ADBM. And with the assumption that the food web constructed from the complete gut content data is the true food web might result in not so accurate estimates of the minimum gut content data. As it might be the case that the links are still under-sampled for many nodes (Woodward et al. 2010b).

5.11 How can we include other types of food web data?

**OP: Focus on this ## How can we include other food web data type?

A future prospect would be to include other types of food web data along with gut content simultaneously to parameterise the ADBM. For example, one could use the approximate trophic position inferred from stable isotope ratio data from a tissue of an individual and gut content data of a different predator simultaneously to parameterise the ADBM. In this case, trophic position would be an additional summary statistics required in the parameterisation method. This is useful when only limited amount of diet information can be extracted by using a single method.

6 Conclusion

**AG: Working on this.

7 Supplementary Information

7.1 Gut content data simulation

**OP: It needs to be very clear in the main text, and here, why this is being described. It is currently unclear, not least because the main text does not, I think, refer to the information here.

We simulated a food web using the ADBM for a given set of parameters. For a given set of predators, we subset the diet from the simulated food web. Then using a probability mass function (distribution), we sampled the gut content data from predators' diet thereby incorporating the uncertainty in the gut content data. We repeated this process multiple number of times for every predator in the food web.

Input:

- Predators whose diet are to be simulated $P = \{p_1, \dots, p_n\}$
- A simulated food web $ADBM(\theta_i) = \{d_{p_1}, d_{p_1}, \dots, d_{p_k}\}$, where d_{p_k} is a one-dimensional diet matrix of predator k containing ones and zeros.
- A function which describes uncertainty in the diet $U(d)$
- Number of independent guts to be simulated for a predator $p_i : ngut$

Sampling:

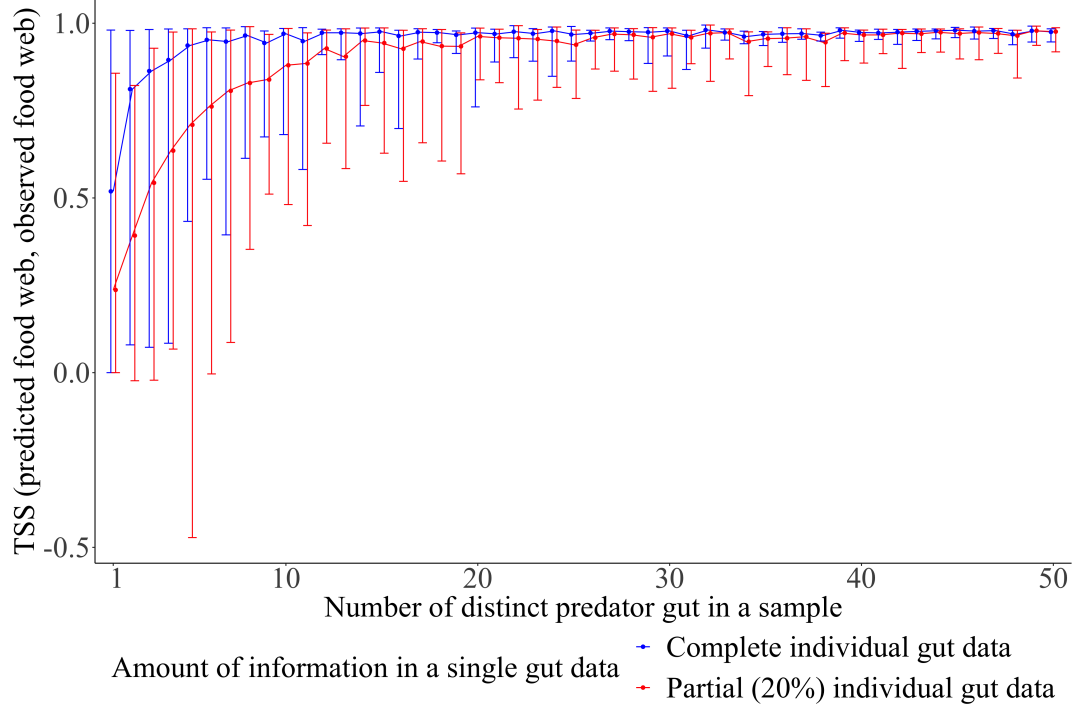


Figure 5: True skill statistics between predicted food web and observed food web for a simulated small reef food web estimated for distinct predator guts in a sample. The observed simulated food web consists of 50 species and ... links. The vertical bars correspond to the prediction intervals of the 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.

288 • for $p_i \in P$

289 – for $j = 1, \dots, ngut$, where $ngut$ is the number of guts to be simulated

290 * Simulate a single gut of a predator $p_i : g(p_i) = d_{p_i} * U(d_{p_i})$

291 – Set of gut of a predator $p_i : G(p_i) = \{g(p_i) : g(p_i) = d_{p_i} * U(d_{p_i})\}$

292 *Output:*

293 • We simulated a pool of gut content data which contains simulated gut content data $G(p_i)$
 294 for every predator p_i

295 7.2 Prediction using simulated gut content data from a simulated food web

296 The true skill statistics (TSS) between the predicted food web and observed food web saturated with
 297 an increasing number of distinct predator guts in a sample (Fig. 5). The TSS of the predicted food
 298 webs estimated using the complete individual gut data had narrower prediction intervals resulting
 299 in less uncertainty, and higher mean TSS than that using the partial individual gut data. The
 300 maximum limit of the prediction interval of TSS estimated using the complete gut data and the
 301 partial gut data were almost equal, with the minimum limit of the prediction interval of TSS using
 302 partial gut data being lower than that from the complete gut data. Eventually, the gap between the
 303 mean TSS using the partial gut data and the complete gut data reduced with an increasing number
 304 of distinct predator guts suggesting when we have enough predator species' gut data, the achieved
 305 TSS was almost constant and hence independent of the amount of gut data.

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

312 8 Acknowledgements

313 9 Author contributions

314 10 References

- 315 Allesina, Stefano, David Alonso, and Mercedes Pascual. 2008. "A General Model for Food Web
 316 Structure." *Science* 320 (5876): 658–61. <https://doi.org/10.1126/science.1156269>.
 317 Allouche, Omri, Asaf Tsoar, and Ronen Kadmon. 2006. "Assessing the Accuracy of Species
 318 Distribution Models: Prevalence, Kappa and the True Skill Statistic (TSS)." *Journal of Applied
 319 Ecology* 43 (6): 1223–32. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>.

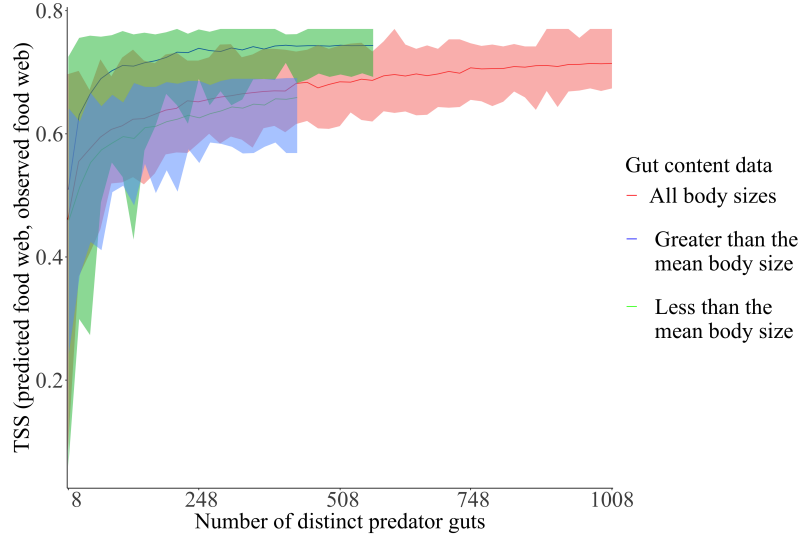


Figure 6: True skill statistics between predicted food web and observed food web estimated for different number of distinct predator guts. The estimation is done for three sets of gut data: gut content data of predators whose body sizes are smaller than the mean body size, larger than the mean body size, and all the gut content data.

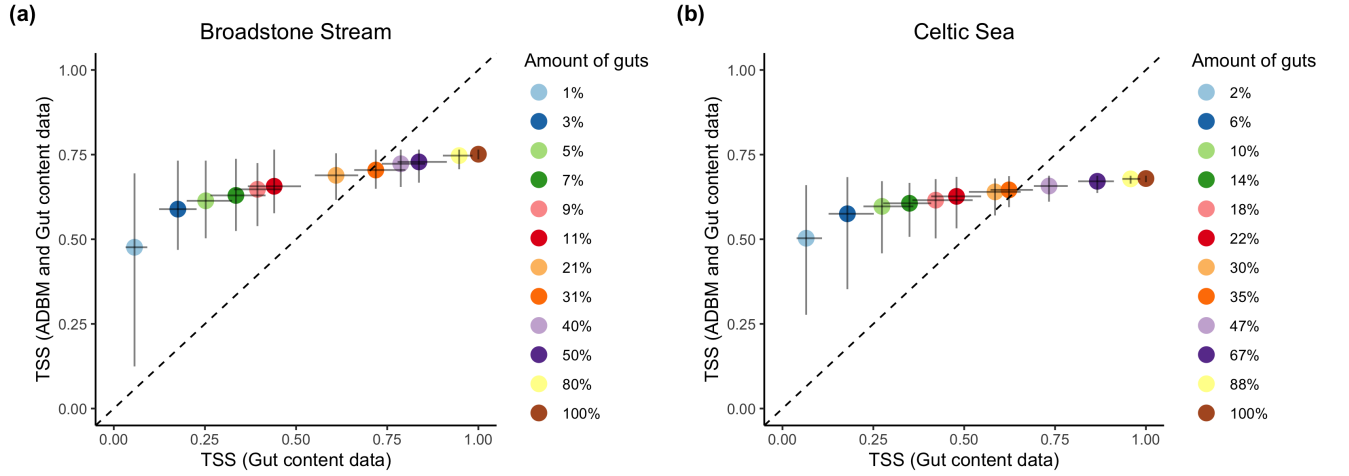


Figure 7: True skill statistic between predicted food web using ADBM and incomplete gut content data, and observed food web against the true skill statistic between food web constructed using incomplete gut content data, and observed food web. Error bars represent prediction intervals of 100 independent samples. Dashed line is 1:1 line for reference.

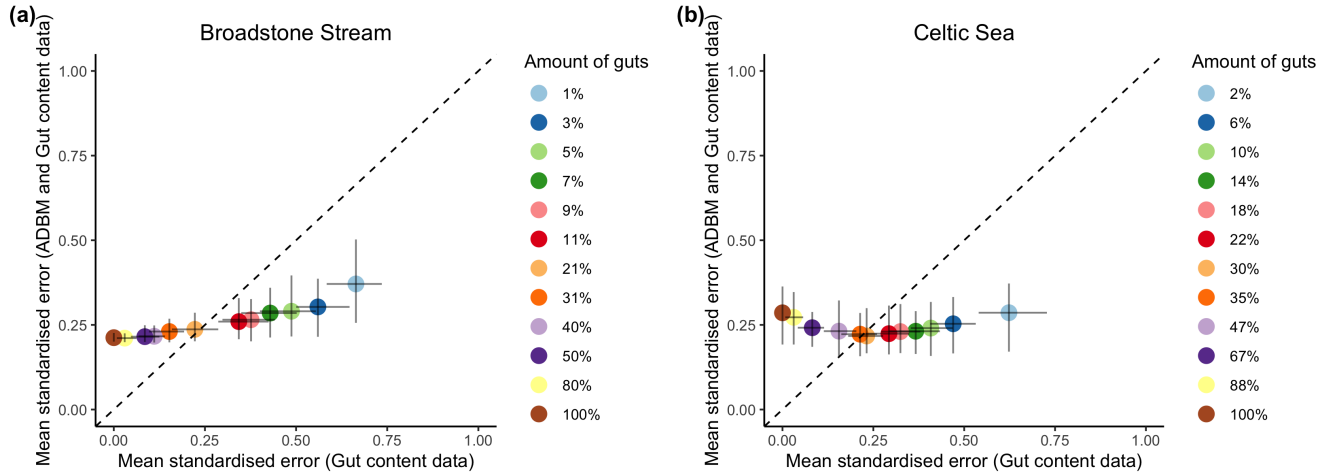


Figure 8: Mean standardised error in structural properties in the food web predicted using ADBM and incomplete gut content data against structural properties in the predicted food web constructed using incomplete gut content data. Error bars represent prediction intervals of 100 independent samples. Dashed line is 1:1 line for reference.

- 320 Baker, Ronald, Amanda Buckland, and Marcus Sheaves. 2014. "Fish Gut Content Analysis: Robust
 321 Measures of Diet Composition." *Fish and Fisheries* 15 (1): 170–77. <https://doi.org/10.1111/faf.12026>.
 322
- 323 Cohen, Joel E., and Christian Mulder. 2014a. "Soil Invertebrates, Chemistry, Weather, Human
 324 Management, and Edaphic Food Webs at 135 Sites in The Netherlands: SIZEWEB." *Ecology* 95
 325 (2): 578–78. <https://doi.org/10.1890/13-1337.1>.
 326 ———. 2014b. "Soil Invertebrates, Chemistry, Weather, Human Management, and Edaphic Food
 327 Webs at 135 Sites in The Netherlands: SIZEWEB." *Ecology* 95 (2): 578–78. <https://doi.org/10.1890/13-1337.1>.
 328
- 329 Cohen, Joel E., C. M. Newman, and John Hyslop Steele. 1985. "A Stochastic Theory of Community
 330 Food Webs I. Models and Aggregated Data." *Proceedings of the Royal Society of London. Series
 331 B. Biological Sciences* 224 (1237): 421–48. <https://doi.org/10.1098/rspb.1985.0042>.
 332 Crawford, Kerry, Robbie A. McDonald, and Stuart Bearhop. 2008. "Applications of Stable
 333 Isotope Techniques to the Ecology of Mammals." *Mammal Review* 38 (1): 87–107. <https://doi.org/10.1111/j.1365-2907.2008.00120.x>.
 334

- 335 Dunne, Jennifer A., Richard J. Williams, and Neo D. Martinez. 2002. "Network Structure and
336 Biodiversity Loss in Food Webs: Robustness Increases with Connectance." *Ecology Letters* 5 (4):
337 558–67. <https://doi.org/10.1046/j.1461-0248.2002.00354.x>.
- 338 Gilljam, David, Aaron Thierry, Francois K. Edwards, David Figueroa, Anton T. Ibbotson, J.
339 Iwan Jones, Rasmus B. Lauridsen, Owen L. Petchey, Guy Woodward, and Bo Ebenman.
340 2011. "Seeing Double:" In *Advances in Ecological Research*, 45:67–133. Elsevier. <https://doi.org/10.1016/B978-0-12-386475-8.00003-4>.
- 342 Goldwasser, Lloyd, and Jonathan Roughgarden. 1993a. "Construction and Analysis of a Large
343 Caribbean Food Web." *Ecology* 74 (4): 1216–33. <https://doi.org/10.2307/1940492>.
- 344 ———. 1993b. "Construction and Analysis of a Large Caribbean Food Web." *Ecology* 74 (4): 1216–33.
345 <https://doi.org/10.2307/1940492>.
- 346 Gravel, Dominique, Timoth'ee Poisot, Camille Albouy, Laure Velez, and David Mouillot. 2013.
347 "Inferring Food Web Structure from Predator-Prey Body Size Relationships." Edited by Robert
348 Freckleton. *Methods in Ecology and Evolution* 4 (11): 1083–90. [https://doi.org/10.1111/](https://doi.org/10.1111/2041-210X.12103)
349 [2041-210X.12103](https://doi.org/10.1111/2041-210X.12103).
- 350 Gray, Clare, David H. Figueroa, Lawrence N. Hudson, Athen Ma, Dan Perkins, and Guy Woodward.
351 2015. "Joining the Dots: An Automated Method for Constructing Food Webs from Compendia
352 of Published Interactions." *Food Webs* 5 (December): 11–20. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.fooweb.2015.09.001)
353 [fooweb.2015.09.001](https://doi.org/10.1016/j.fooweb.2015.09.001).
- 354 Gupta, Anubhav, and Owen Petchey. 2021. "Simultaneously Estimating Food Web Complexity and
355 Structure with Uncertain." <https://doi.org/10.22541/au.162176218.82114885/v1>.
- 356 Hattab, Tarek, Fabien Leprieur, Frida Ben Rais Lasram, Dominique Gravel, François Le Loc'h,
357 and Camille Albouy. 2016. "Forecasting Fine-Scale Changes in the Food-Web Structure of
358 Coastal Marine Communities Under Climate Change." *Ecography* 39 (12): 1227–37. <https://doi.org/10.1111/ecog.01937>.
- 360 Hyslop, E. J. 1980. "Stomach Contents Analysis—a Review of Methods and Their Application." *Jour-*
361 *nal of Fish Biology* 17 (4): 411–29. <https://doi.org/10.1111/j.1095-8649.1980.tb02775.x>.
- 362 Jord'an, Ferenc. 2009. "Keystone Species and Food Webs." *Philosophical Transactions of the Royal*
363 *Society B: Biological Sciences* 364 (1524): 1733–41. [https://doi.org/10.1098/rstb.2008.](https://doi.org/10.1098/rstb.2008.0335)
364 [0335](https://doi.org/10.1098/rstb.2008.0335).

- Kadoya, Taku, Yutaka Osada, and Gaku Takimoto. 2012. “IsoWeb: A Bayesian Isotope Mixing Model for Diet Analysis of the Whole Food Web.” Edited by Simon Thrush. *PLoS ONE* 7 (7): e41057. <https://doi.org/10.1371/journal.pone.0041057>.
- 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%5B42:CSIRPF%5D2.0.CO;2](https://doi.org/10.1890/0012-9658(2007)88%5B42:CSIRPF%5D2.0.CO;2).
- Lindegren, Martin, Christian Möllmann, Anders Nielsen, Keith Brander, Brian R. MacKenzie, and Nils Chr. Stenseth. 2010. “Ecological Forecasting Under Climate Change: The Case of Baltic Cod.” *Proceedings of the Royal Society B: Biological Sciences* 277 (1691): 2121–30. <https://doi.org/10.1098/rspb.2010.0353>.
- MacArthur, Robert H., and Eric R. Pianka. 1966. “On Optimal Use of a Patchy Environment.” *The American Naturalist* 100 (916): 603–9. <https://www.jstor.org/stable/2459298>.
- Nielsen, Jens M., Elizabeth L. Clare, Brian Hayden, Michael T. Brett, and Pavel Kratina. 2018. “Diet Tracing in Ecology: Method Comparison and Selection.” Edited by M. Thomas P. Gilbert. *Methods in Ecology and Evolution* 9 (2): 278–91. <https://doi.org/10.1111/2041-210X.12869>.
- O’NAGorman, Eoin J., Owen L. Petchey, Katy J. Faulkner, Bruno Gallo, Timothy A. C. Gordon, Joana Neto-Cerejeira, J’on S. ’Olafsson, Doris E. Pichler, Murray S. A. Thompson, and Guy Woodward. 2019. “A Simple Model Predicts How Warming Simplifies Wild Food Webs.” *Nature Climate Change* 9 (8, 8): 611–16. <https://doi.org/10.1038/s41558-019-0513-x>.
- 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): 484–95. <https://doi.org/10.1071/MF15154>.
- Petchey, Owen L., Andrew P. Beckerman, Jens O. Riede, and Philip H. Warren. 2008. “Size, Foraging, and Food Web Structure.” *Proceedings of the National Academy of Sciences* 105: 4191–96. <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>.

- 395 Tamaddoni-Nezhad, Alireza, Ghazal Afroozi Milani, Alan Raybould, Stephen Muggleton, and David
 396 A. Bohan. 2013. “Construction and Validation of Food Webs Using Logic-Based Machine
 397 Learning and Text Mining.” In *Advances in Ecological Research*, 49:225–89. Elsevier. <https://doi.org/10.1016/B978-0-12-420002-9.00004-4>.
 398
- 399 Warren, Philip H. 1989. “Spatial and Temporal Variation in the Structure of a Freshwater Food
 400 Web.” *Oikos* 55 (3): 299. <https://doi.org/10.2307/3565588>.
- 401 Woodward, Guy, Julia Blanchard, Rasmus B. Lauridsen, Francois K. Edwards, J. Iwan Jones,
 402 David Figueroa, Philip H. Warren, and Owen L. Petchey. 2010b. “Individual-Based Food
 403 Webs.” In *Advances in Ecological Research*, 43:211–66. Elsevier. [https://doi.org/10.1016/](https://doi.org/10.1016/B978-0-12-385005-8.00006-X)
 404 [B978-0-12-385005-8.00006-X](https://doi.org/10.1016/B978-0-12-385005-8.00006-X).
- 405 ———. 2010a. “Individual-Based Food Webs.” In, 43:211–66. Elsevier. [https://doi.org/10.1016/](https://doi.org/10.1016/B978-0-12-385005-8.00006-X)
 406 [B978-0-12-385005-8.00006-X](https://doi.org/10.1016/B978-0-12-385005-8.00006-X).