



How many predator guts are required to predict trophic interactions?

Anubhav Gupta^{a,*}, David Figueroa H.^b, Eoin O’Gorman^c, Iwan Jones^d, Guy Woodward^e, Owen L. Petchey^a

^a Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland

^b Departamento de Ciencias Ambientales, Universidad Católica de Temuco, Temuco, Chile

^c School of Life Sciences, University of Essex, CO4 3SQ Colchester, UK

^d School of Biological and Behavioural Sciences, Queen Mary University of London, River Laboratory, Wareham, Dorset, UK

^e Georgina Mace Centre for the Living Planet, Department of Life Sciences, Imperial College London, Ascot, Berkshire SL5 7PY, UK

ARTICLE INFO

Keywords:

Predator guts

ADBM

Food web accuracy

Food web prediction

Trophic interactions

Food web model

Undersampling

ABSTRACT

- 1) A large obstacle in food web ecology is the time and effort required to adequately describe the structure of a food web using individual predator guts. Food web models such as the allometric diet breadth model (ADBM) can be used to circumvent this problem by predicting the interactions based on easily measured characteristics, such as the size of organisms. However, diet data such as that which comes from analysis of predator guts is still required to parameterise these food web models, and collecting and analysing these data from the field is an expensive and time-consuming task. Therefore, it is important to know how many predator guts are required to parameterise food web models to obtain food web structures with high accuracy and precision.
- 2) Here, we explore seven exceptionally well-characterised food webs and determine the minimum number of predator guts needed to accurately predict their structure using the ADBM. We use Bayesian computation to parameterise the ADBM, and true skill statistics to measure the goodness of fit, and do so while varying the number of predator guts used in the parameterisation to test the effect of sampling effort.
- 3) We found that relatively few, and many fewer than were actually collected, predator guts can be used to parameterise the ADBM. The lowest number of predator guts was 27% of the number of available predator guts. The number of predator guts required to accurately characterise food webs increases by $\sim 7 \pm 2.2$ guts for 10 units increase in the number of trophic links and $\sim 9 \pm 4.7$ guts for a unit increase in the number of species.
- 4) These results suggest that one need not collect and analyse such a large quantity of predator guts in order to adequately predict the structure of a food web, thereby reducing sampling effort considerably, while having little effect on precision or accuracy of predictions.

1. Introduction

Knowledge about the trophic interactions in a food web is crucial in ecology for purposes ranging from identifying keystone species (Jordán, Ferenc, 2009), to quantifying robustness of a food web, and to predicting species extinctions (Dunne et al., 2002). This has led to the

development of numerous food web models and associated theories (Allesina et al., 2008; Cohen et al., 1985; Gravel et al., 2013; Petchey et al., 2008; Tamaddoni-Nezhad et al., 2013). Along with inferring missing links in an observed food web, such food web models are also increasingly used for ecological forecasting (Hattab et al., 2016; Lindgren et al., 2010) and for understanding the underlying mechanism

* Corresponding author.

E-mail addresses: anubhav.gupta@ieu.uzh.ch (A. Gupta), dfiguero@uct.cl (D. Figueroa H.), e.ogorman@essex.ac.uk (E. O’Gorman), j.i.jones@qmul.ac.uk (I. Jones), guy.woodward@imperial.ac.uk (G. Woodward), owen.petchey@ieu.uzh.ch (O.L. Petchey).

<https://doi.org/10.1016/j.fooweb.2022.e00269>

Received 12 September 2022; Received in revised form 16 November 2022; Accepted 8 December 2022

Available online 31 December 2022

2352-2496/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

governing trophic interactions in the wild (O’Gorman et al., 2019; Petchey et al., 2008; Valdovinos et al., 2010; Sentis et al., 2014; Arim et al., 2010).

Although food web models are constructed using prior theory and hypotheses about the factors that determine trophic interactions, empirical data about interactions are required to parameterise them. For example, Petchey et al. (2008) and Gupta et al. (2022) used presence-absence information about trophic interactions to parameterise the allometric diet breadth model and thereby predict species interactions. Such empirical information about interactions can come from a diverse set of methods such as gut content analysis (Peralta-Maraver et al., 2017), stable isotope ratio analysis of tissues (Layman et al., 2007), experimentation (Warren, 1989), DNA metabarcoding of gut contents or faeces (Roslin and Majaneva, 2016) and literature research (Gray et al., 2015; Cohen and Mulder, 2014a; Goldwasser and Roughgarden, 1993a). Nevertheless, each of these sources of information about trophic interactions has shortcomings that hinder the advancement of the field. For example: stable isotope ratio analysis of the organism’s tissue does not give direct taxonomically resolved information of the diet of that organism. Instead, stable isotope ratios provide an approximate trophic position of that species in the food web (Wada et al., 1991; Jennings and van der Molen, 2015) that can be paired with mixing models to determine what prey items are most likely fed upon by a predator. However, this results in uncertainty in the estimates (Kadoya et al., 2012; Crawford et al., 2008). Similarly, more recent approaches using DNA metabarcoding may give much higher taxonomic resolution but present other challenges, such as not providing trait information (e.g. body sizes of prey) directly and an inability to resolve secondary predation or cannibalism (Pompanon et al., 2012; Nielsen et al., 2018) which are common in nature. Of course, if metabarcoding reveals species identity then the trait information can be inferred in a number of ways (Compson et al., 2019; Djemiel et al., 2022; Andriollo et al., 2021). Furthermore, DNA approaches are prone to environmental contamination (e.g. DNA in the water swallowed along with DNA from an aquatic consumer’s prey cannot be differentiated from actual prey) (Kelly et al., 2014). Construction of food webs via literature review, a common practice in food web research, can lead to links being assigned between species that do not occur in reality. This can result if the presence of a link between two species is context dependent, i.e. occur in one environmental context but not in another (Gray et al., 2015; Cohen and Mulder, 2014b; Goldwasser and Roughgarden, 1993b). It is unsurprising, given the limitations of these proxy or inferential approaches, that gut content analysis based diets have a better match with real diets when compared to other methods (Nielsen et al., 2018).

Although gut content analysis is viewed as the “gold standard”, acquiring food web data from the gut contents is extremely time-consuming and expensive (Gray et al., 2015). It also requires high skill levels in taxonomic identification, often involving dissection and microscopy techniques (Hyslop, 1980). The perception that this is unavoidably laborious and costly is also in part due to the assumption that many gut contents must be collected and analysed in order to be confident that the majority of possible trophic links among species have been observed. Some studies fail to quantify the effort needed, with yield-effort curves being the exception rather than the rule and those that have been quantified often point to the apparent need for hundreds or thousands of guts to be analysed to fully capture a food web’s structure. Hence, it is of considerable importance to know the minimum number of predator guts required to parameterise a food web model with high accuracy and precision: this would enable researchers to allocate resources more effectively a priori. Moreover, a rough rule of thumb of when to stop collecting data from gut contents would provide a method to adapt the amount of sampling/analysis during the sampling/analysis process itself.

Therefore, the key question we wish to answer is how much diet information, derived from predator gut data, is required to infer food web structure from a food web model with adequate accuracy and

precision? In other words, how many samples of predator guts should one collect and analyse from the field to parameterise a food web model? To answer this question, we vary the number of predator gut data that is used to parameterise the allometric diet breadth model (ADBM) and record how well the model can then predict the food web produced when all the available predator gut data are used. We do this separately for seven different food webs, for each calculating the minimum number of predator guts required to adequately infer food web structure. To get an estimate of how many samples of predator guts need to be collected from the field to infer structure for a food web with a given number of species or number of trophic links, we also investigate how this minimum number of predator guts depends upon the number of species and the number of trophic links in a food web. We expect a positive linear relationship between this minimum number of predator guts and the number of trophic links and the number of species because a larger number of predator guts are required to characterise structure of a food web with a high number of species or a high number of trophic links. Our study provides a guideline on how many predator guts are required to predict food web structure using a food web model.

2. Materials and methods

Here we present the empirical food webs, the allometric diet breadth model (ADBM), and the predator guts used to infer the trophic interactions. We also give a detailed account of using partial predator guts to parameterise the ADBM using the rejection approximate Bayesian computation (ABC). We assessed model predictions using the true skill statistic for comparison across the food webs.

2.1. The empirical food webs

Traditionally, food webs describe the trophic links among species (i.e. each of the species is a node in the food web, and links occur when one species preys on another). However, we consider food webs where nodes are size classes, i.e. individuals are aggregated into these size classes based on their body size. 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 of predator, irrespective of the taxonomy of the individuals. We used this approach for several reasons, such as to take into account the ontogenetic shifts in the diet of a predator (Woodward et al., 2010) and to account for individual-based interactions which would not have been considered if nodes were aggregated based on taxonomy because a taxonomic group can have a large variation in the body size. This size-based approach has often been used in strongly size-structured ecosystems that also have considerable intraspecific size variation and has been instrumental in understanding the commercial exploitation of marine ecosystems (Jennings and Brander, 2010). Second, we also consider more traditional food webs where nodes are aggregated based on the taxonomy of the individuals. The number of nodes in the size-based food web and taxonomic-based food web were equal to make them comparable.

We analysed food webs for which predator guts are available at an individual level, with body size data, and with data that is or that we could make FAIR (Findable Accessible Interoperable Reusable). Our study food webs are freshwater food webs except for the Celtic Sea food web which is a marine food web. Most of the food webs are dominated by invertebrates except the Celtic Sea which is dominated by fishes and the Tadnoll Brook which is dominated by fishes as well as invertebrates. The food webs vary in the number of nodes, trophic links, connectance and body sizes (Table 1 and S1).

Invertebrates in freshwater food webs were collected using a Hess or Surber sampler, while fishes were caught with an electrofisher, and anaesthetised using 2-phenoxyethanol. In the case of the Celtic Sea, fishes were caught using trawling.

The foreguts (i.e. the stomodaeum, which is the part of the gut between the mouth and the midgut) of the collected invertebrate predators

Table 1
Information about the empirical size-based food webs.

| Common food web name (Original Publication) | Location | Predation matrix source | Body size source | General ecosystem | Number of nodes | Number of links | Connectance | Body size Range (mg) (approximate) |
|---------------------------------------------|--------------------------------------------------------|-------------------------|----------------------|-------------------|-----------------|-----------------|-------------|-------------------------------------|
| Broadstone Stream (Woodward et al., 2010) | England, UK 51°05'N 0°03'E | Woodward (2021) | Woodward (2021) | Freshwater | 28 | 185 | 0.24 | 10 ⁻⁷ to 10 ² |
| Celtic Sea (Barnes et al., 2008) | British Isles and French coastal shelf 50°50'N 08°00'W | Barnes et al. (2016) | Barnes et al. (2016) | Marine | 55 | 439 | 0.15 | 10 ⁻² to 10 ⁴ |
| Tadnoll Brook (Woodward et al., 2010) | England, UK 50°41'N 02°19'W | Jones et al. (2022) | Jones et al. (2022) | Freshwater | 59 | 485 | 0.14 | 10 ⁻⁶ to 10 ⁵ |
| Afon Hirnant (Woodward et al., 2010) | Wales, UK, 50°52'N 03°34'E | Figueroa (2022a) | Figueroa (2022a) | Freshwater | 77 | 381 | 0.06 | 10 ⁻⁶ to 10 ² |
| Coilaco (Figueroa, 2007) | Chile 39°17'S 71°44'W | Figueroa (2022b) | Figueroa (2022b) | Freshwater | 45 | 123 | 0.06 | 10 ⁻⁶ to 10 ² |
| Guampoe (Figueroa, 2007) | Chile 39°23'S 71°41'W | Figueroa (2022b) | Figueroa (2022b) | Freshwater | 44 | 139 | 0.07 | 10 ⁻⁶ to 10 ³ |
| Trancura (Figueroa, 2007) | Chile 39°26'S 71°32'W | Figueroa (2022b) | Figueroa (2022b) | Freshwater | 35 | 78 | 0.06 | 10 ⁻⁶ to 10 ¹ |

were dissected and examined under a microscope. Regression equations were used to convert predator and prey lengths to the respective body masses. In the case of highly digested prey items, previously established regressions based on the width of head capsule were used as an alternative linear dimension. A more detailed description of these food webs is present in Gilljam et al. (2011).

2.2. 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 can predict species interactions based on an easily measurable trait body size (Petchey et al., 2008). The ADBM predicts the set of prey types (e.g. species or sizes classes) a consumer should feed upon to maximise its rate of energy intake. The foraging variables in the model are: energy content of the prey, handling times of the predator on a prey, space clearance rate (also known as the attack rate—how fast a predator searches space), and prey densities. Each of these are predicted from allometric scaling relationship, thus the model requires body sizes and the allometric constants and exponents in order to make predictions. 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).

2.3. Assessment of prediction

The accuracy of the predicted diet of the predators was measured using a widely used accuracy measure in food web ecology, namely the true skill statistic (TSS) (Gray et al., 2015; Gravel et al., 2013; Gupta et al., 2022). We chose this metric because it takes into account the true and false predictions of both the presence and absence of links. It is 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 indicates a performance no better than random, and less than zero indicates a performance worse than random (Allouche et al., 2006).

2.4. Parameterisation using partial predator gut information

From an empirical dataset of predator guts, we take a random sample of gut contents of a specific size (see below) to create a partial predator dataset, and then fit the ADBM to this partial dataset (Fig. 1).

To fit the ADBM to the partial predator guts dataset, we used the rejection approximate Bayesian computation method we previously developed in Gupta et al. (2022). In brief, we accept a parameter value from a prior distribution which would have resulted in a predicted food web that has sufficient similarity to the observed food web. The true skill statistic was computed between the diets predicted by the ADBM, and those observed in the sampled predator guts. We repeated this process n ($=100$) times for every i number of guts, where i lies between 1 and the total number of predator guts in the pool. Pseudocode for the procedure is provided in the Supplementary information (hereafter SI).

2.5. Computing the minimum number of predator guts

Using the TSS of the model predicted food webs for different number of predator guts, we computed the number of predator guts that resulted in the mean TSS equal to the 95% of the mean TSS achieved by the model using all the predator guts available in the pool for a food web. We chose the proportion 95% because the food webs predicted using the mentioned proportion was very similar to the food web predicted using all the predator guts. We term this number of predator guts the “minimum number of predator guts”. We further tested the dependence of the minimum number of predator guts on the maximum number of predator guts, number of trophic links and number of species by fitting linear regressions in Fig. 2 (h) and Fig. 3 (a, b) separately.

2.6. Standardising sampling level of the food webs

Since the seven food webs have different levels of sampling effort, with the Broadstone Stream food web being the most sampled among all, and every other food web being considerably less well sampled (SI Fig. S3 and S6), we wished to attempt to compensate for this in our analyses. To do this we used the R *vegan* package to account for the undersampling with respect to the Broadstone Stream food web. We fitted the link accumulation curves using the *fitspecaccum* function to a set of nonlinear regression models suggested in Dengler (2009) and used the AIC criteria for model selection. We then extrapolated the link accumulation curves for all the food webs except the Broadstone Stream and computed the number of predator guts that would have resulted in the gradient of the link accumulation curve equal to the gradient of that

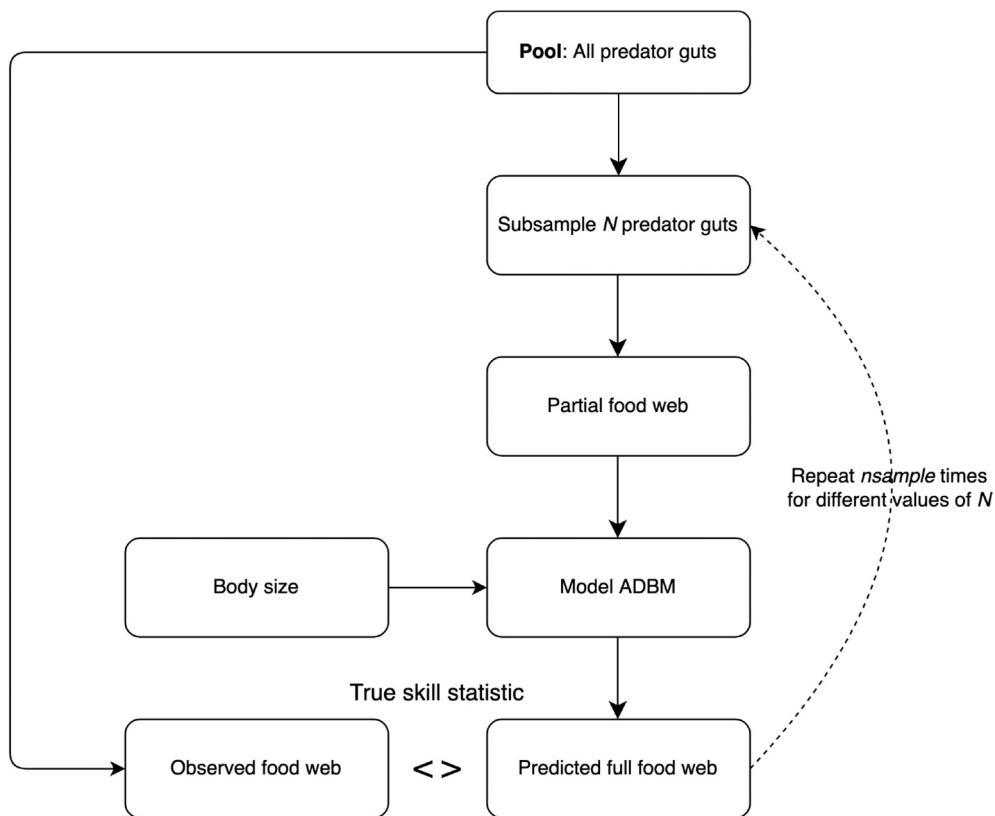


Fig. 1. Flowchart of the subsampling method implemented to predict full food web from the food web model using the predator guts.

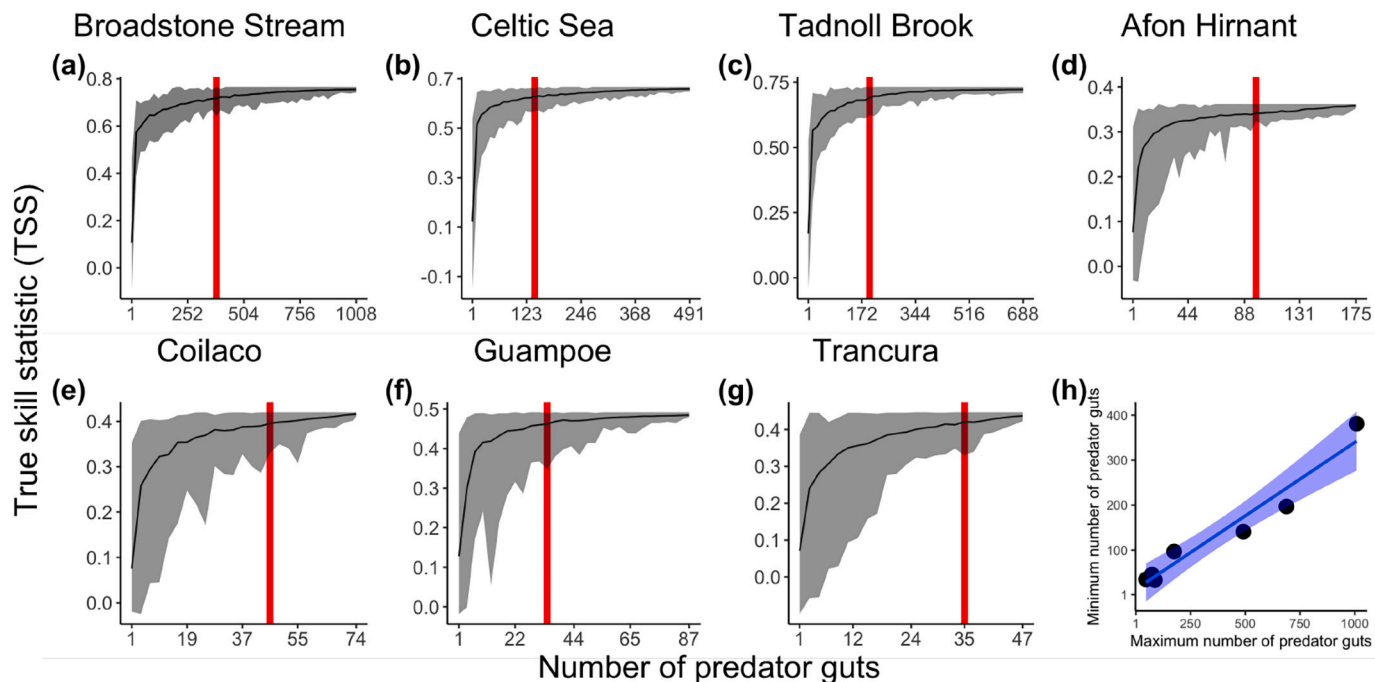


Fig. 2. (a, b, c, d, e, f, g) Accuracy of the predicted size-based food webs measured using the true skill statistic, predicted by the ADBM parameterised using predator guts. The line and the shaded gray region represent the mean and the prediction interval corresponding to 100 independent samples respectively. The red lines represent the minimum number of predator guts required to achieve a TSS of 95% of the maximum TSS. (h) Minimum number of predator guts plotted against the maximum number of predator guts. Solid blue line is a linear regression ($t = 9.327$, $df = 5$, $P = 0.0002$) and light blue region represents 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

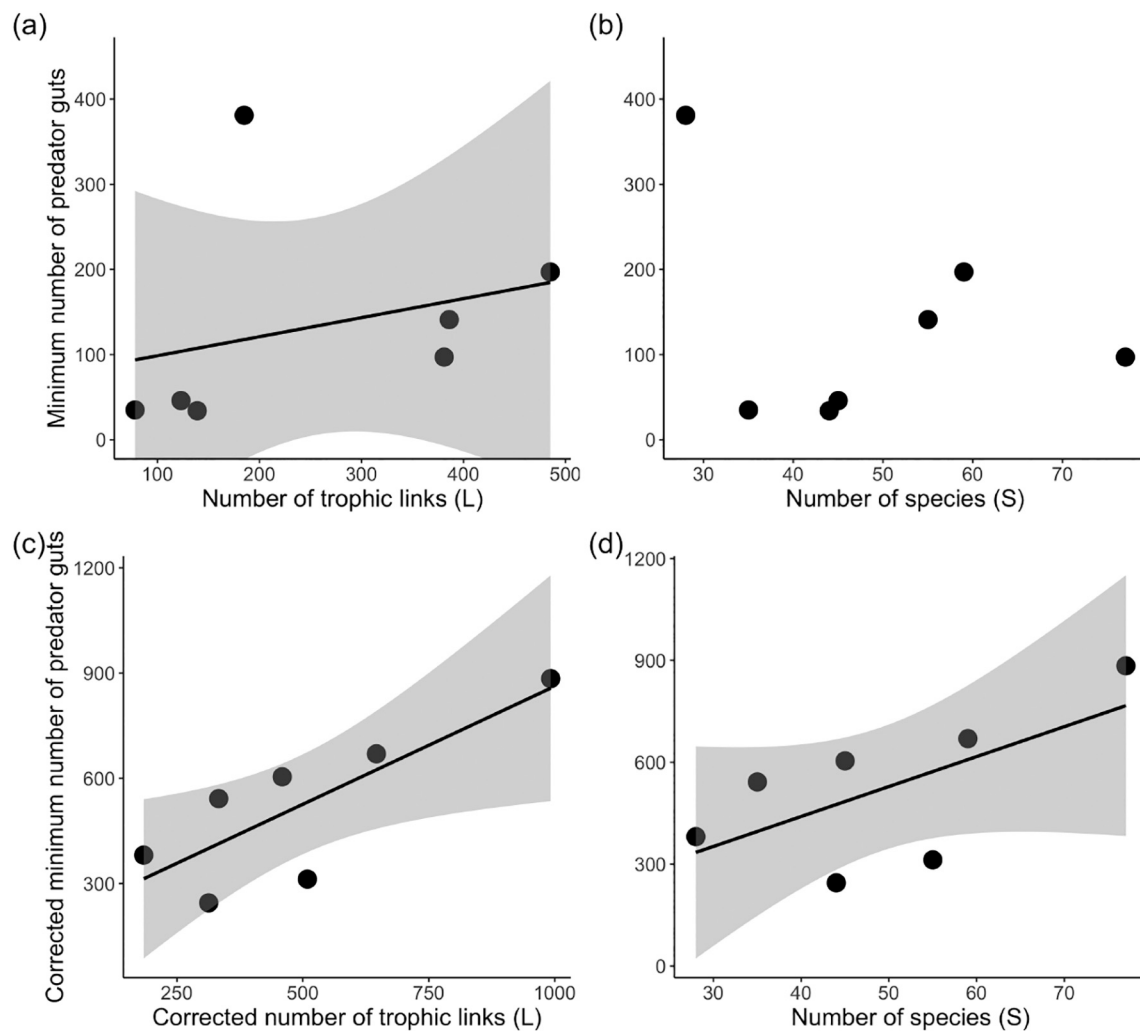


Fig. 3. (a, b) Minimum number of predator guts (i.e. the amount of predator guts used in order to ensure 95% of the maximum TSS) plotted against number of trophic links (L) and number of species respectively. (c, d) Corrected minimum number of predator guts (i.e. the minimum number of predator guts which takes into account the undersampling level of the food webs) plotted against corrected number of trophic links (L) and number of species (S) respectively. Solid lines are linear regression ((a) $t = 0.665$, $df = 5$, $P = 0.535$; (c) $t = 3.001$, $df = 5$, $P = 0.03$; (d) $t = 1.879$, $df = 5$, $P = 0.119$) and gray region represents 95% confidence intervals.

of the Broadstone Stream when all the predator guts were used (Fig. S4, S5, S7 and S8). We also calculated the corrected number of trophic links corresponding to the corrected number of predator guts. For each food web, we then calculated an undersampling factor, equal to the ratio of the corrected number of predator guts to the number of predator guts in the pool. Using the undersampling factor, we further calculated the corrected minimum number of predator guts, which was the product of the undersampling factor and the minimum number of predator guts. We further tested the dependence of the corrected minimum number of predator guts on the number of trophic links and number of species by fitting linear regressions in Fig. 3 (c, d) separately.

3. Results

We first present how the accuracy of the food web model in predicting trophic interactions in size-based food webs varies with an increasing number of predator guts provided to the food web model. We then show how the minimum number of predator guts varies with the number of trophic links and the number of species. The results for a similar analysis performed with taxonomic-based food webs are presented in the SI Section S1 (hereafter SI S1).

The true skill statistics of the food webs predicted by the ADBM using incomplete predator guts improved quickly for lower number of

predator guts (Fig. 2 (a-g), S1 (a-g)). Furthermore, the width of the prediction interval of the true skill statistics decreased with increasing number of predator guts, with the mean TSS asymptoting to the maximum mean TSS achieved by the ADBM when all the predator guts were used. Although the maximum TSS varied among the food webs, the qualitative increase in the TSS was the same.

For the Broadstone Stream food web, with only 381 predator guts, which is 38% of the total predator guts, the ADBM predicted the food web's structure with the mean TSS of 0.74. This was equivalent to 95% of the mean TSS (0.78) achieved using complete predator guts (Fig. 2 (a)). I.e. the main characteristics of the food web could be captured with about 1/3 of the effort used in the original study. In case of the Tadnoll Brook food web, only 27% of the total predator guts ($n = 183$) was needed to accurately describe the food web structure (Fig. 2(c)). However, to accurately characterise the Trancura food web, 75% of the total number of predator guts were required (Fig. 2(g)) which is proportionally high as compared to other food webs. Fig. 2(h) summarises the minimum number of predator guts required to describe the structure of food web given the maximum number of predator guts that were actually processed to describe the structure.

The minimum number of predator guts was not significantly related to the number of trophic links (Fig. 3 (a), S2 (a)) and the number of species (Fig. 3 (b), S2 (b)) for size-based as well as taxonomic-based food

webs. The corrected minimum number of predator guts was significantly related to the corrected number of trophic links (Fig. 3 (c), S2 (c)) but not significantly related to the number of species (Fig. 3 (d)) in size-based food webs. In case of taxonomic-based food webs, there was a significant relationship between the corrected minimum number of predator guts and the number of species (Fig. S2 (d)). Correcting for the undersampling in the food webs improved the fit between the minimum number of predator guts and the number of trophic links from $R^2 = 0.08$ (Fig. 3 (a)) to $R^2 = 0.64$ (Fig. 3 (c)) in the size-based food webs and from $R^2 = 0.566$ (Fig. S2 (a)) to $R^2 = 0.9099$ (Fig. S2 (c)) in the taxonomic-based food webs.

4. Discussion

We have demonstrated how a food web model can be used to predict the full structure of a food web when incomplete data about its trophic interactions is available, which is true in most real food webs. This can help inform how many predator guts to collect and analyse to infer trophic interactions for an ecosystem with a given number of species. A future development could be to make the same assessment using other food web models, and to also use food web data other than predator guts to parameterise those models.

Our study provides a ballpark figure of the minimum number of predator guts that need to be sampled to predict the structure of a food web with a given number of species using a food web model. For instance, Fig. 3 (d) and S2 (d) can be used as a rough estimate of how many predator guts need to be collected to predict food web structure using a food web model for a given number of species. This could lead to a reduction in the number of predator guts that would be collected (Ings et al., 2009; Woodward et al., 2010), thereby saving considerable time and resources. For instance, characterisation of the Broadstone Stream food web required around 1000 predator guts in Woodward et al. (2010), however the model can accurately predict the food web's structure with only around one third of the predator guts (Fig. 2 (a)).

We expected positive relationships between the minimum number of predator guts and the number of trophic links and between the minimum number of predator guts and the number of species respectively. However, we did not observe such relationships (Fig. 3 (a, b)). We suspect this is due to the possibility that the seven food webs have very different amounts of predator gut samples (Fig. S3), with the Broadstone Stream food web being the most sampled among all the food webs. Taking into account the undersampling resulted in a better fit between the corrected minimum number of predator guts and the corrected number of trophic links and the number of species respectively (Fig. 3 (c, d)). It improved the linear model fit but did not result in a statistically significant relationship. This could be due to heterogeneity in the predator guts across the food webs and heterogeneity among the food webs. First, if fewer number of prey items are present in a predator gut, then more number of predator guts would need to be collected on an average to quantify the diet of that predator. Second, a food web which has a high proportion of generalist species would require a high number of predator guts on average to characterise the food web structure as compared to characterising the structure of a food web which has a high proportion of specialist species. Therefore, to average out the effect of these confounding variables and thereby have a much better understanding of how the minimum number of predator guts varies with the number of species, future studies should replicate our analysis with other food webs where individual based predator gut data are available.

How general is our rule of thumb across different ecosystem types? We suspect our rule will better apply to food webs where trophic interactions are size-structured as compared to otherwise. For example: our rule of thumb might better apply to aquatic ecosystem when compared to terrestrial ecosystem as Potapov et al. (2019) have shown that the relationship between the body size and trophic level of consumers was present in aquatic consumers and absent in terrestrial consumers.

In our study, we have worked only with the ADBM, which is a model based on size rules. We expect to get a similar result (i.e. minimum number of predator guts for a food web) for different food web models based on size rules such as those by Gravel et al. (2013) and Vagnon et al. (2021). For a given food web, however, some food web models might be better at predicting structure compared to others, so we suggest extending our approach to a range of models and their performance when parameterised with limited data (Williams and Martinez, 2000; Gravel et al., 2013; Allesina et al., 2008; Cattin et al., 2004). For example, future research could study how well different food web models' prediction accuracy vary with different amount of predator guts. This could also help in making decisions about which food web model to choose from for a given set of predator guts. We suspect the relationship (i.e. shape of the curve) between the TSS of the predicted food web and the number of predator guts might vary within food web models because of the difference on the set of rules used to define those models and how well those rules explain the food web structure. For example: a food web model based on body size trait would require less amount of data to predict a size-structured food web as compared to a food web model based on trait other than body size.

In all the seven food webs, the ADBM was able to infer the trophic interactions using incomplete predator guts. Why was the ADBM able to do so? Because the ADBM was constructed using set of foraging rules based on body sizes, partial diet information from the predator guts was sufficient to constrain 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 a set of foraging rules, it still requires some diet data to constrain the posterior parameter space thereby making more accurate predictions (Petchey et al., 2008). Also, to characterise trophic interactions which are rare in nature one would require more predator guts to observe those interactions as compared to characterising trophic interactions which are more frequent in nature. The model is able to predict these rare interactions using a relatively lower number of predator guts which one might have inferred directly from the predator guts only after collecting a large number of gut content samples.

Most observed food webs constructed using available data are not completely characterised due to undersampling (Jordano, 2016) which can result in several biased parameters and network patterns (Chacoff et al., 2012). In such cases, food web model such as the ADBM can be used to compensate for those missing links, as our study has depicted that the full food web can be predicted using relatively fewer empirical data such as predator guts. It would be useful to estimate the extent of undersampling of other published food webs relative to our well characterised food webs. To do this one could collect a number of predator guts that were sampled in as many published food webs as possible and assess how many more or less were sampled compared to the number in our well characterised food webs. This will inform about how many studies have been likely underreporting links and when a food web model could be efficiently used to capture the food web properties that would have not been captured otherwise.

Like any food web model, the food web model used (the ADBM) cannot explain all the interactions in any observed food web. The foraging rules it encodes are based on body size and have particular structure and assumptions; not all of these are met by all observed interactions (Petchey et al., 2008). For example, the ADBM can only predict diets that are contiguous with respect to the size of prey. I.e. it cannot predict that a predator will consume an organism of size 1 and 3, and not organism of size 2. Hence, if the observed diets are not contiguous when prey are ordered by their size, the estimation process could lead to a lower value of the TSS (Gupta et al., 2022).

Furthermore, the observed data may be missing links, e.g. links that rarely occur. Some of these food webs are undersampled (Fig. S3, S6) suggesting those food webs might be missing these rare trophic links, and the false positives from a model might be a correctly predicted link. A future prospect could be to incorporate other sources of presence-

absence data such as stable isotope ratio (Layman et al., 2007), DNA metabarcoding (Roslin and Majaneva, 2016), literature review (Gray et al., 2015; Cohen and Mulder, 2014a; Goldwasser and Roughgarden, 1993a) and experimentation (Warren, 1989) to complement trophic links that may have been missed by the gut content method.

In our study, we have not considered any uncertainty involved in analysing the predator guts (Baker et al., 2014). For example, there are sometimes 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 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 et al., 2014). Therefore, the next step could be to incorporate these different factors of uncertainty in parameterising the model and to understand how these affect the accuracy of the predicted food webs.

We have provided a rule of thumb on how many predator guts one should collect from the field to predict the structure of a food web for a given number of species. This could lead to a reduction in the number of predator guts that would have been collected otherwise thereby saving considerable time and resources. We have also demonstrated how a food web model can be used to predict the full structure of a food web when only partial information of trophic interactions is available.

Data accessibility statement

All the data used in this study was collected in other studies and is openly available. We list those studies and the open access source in Table 1. The complete code used in the analysis is available in the repository doi:<https://doi.org/10.5281/zenodo.7071190>.

CRedit authorship contribution statement

Anubhav Gupta: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Writing – original draft, Writing – review & editing. **David Figueroa H:** Data curation, Writing – review & editing. **Eoin O’Gorman:** Conceptualization, Methodology, Writing – review & editing. **Iwan Jones:** Data curation. **Guy Woodward:** Conceptualization, Methodology, Writing – review & editing. **Owen L. Petchey:** Conceptualization, Resources, Supervision, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

None declared.

Acknowledgements

This work was supported by the University Research Priority Program Global Change and Biodiversity (Grant number: U-704-04-11) of the University of Zurich. We thank the Petchey group members for their valuable suggestions in the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.fooweb.2022.e00269>.

References

- Allesina, Stefano, Alonso, David, Pascual, Mercedes, 2008. A general model for food web structure. *Science* 320 (5876), 658–661. <https://doi.org/10.1126/science.1156269>.
 Allouche, Omri, Tsoar, Asaf, Kadmon, Ronen, 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43 (6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>.

- Andriollo, Tommy, Michaux, Johan R., Ruedi, Manuel, 2021. Food for everyone: differential feeding habits of cryptic bat species inferred from DNA Metabarcoding. *Mol. Ecol.* 30 (18), 4584–4600.
 Arim, Matias, Abades, Sebastian R., Laufer, Gabriel, Loureiro, Marcelo, Marquet, Pablo A., 2010. Food web structure and body size: trophic position and resource acquisition. *Oikos* 119 (1), 147–153. <https://doi.org/10.1111/j.1600-0706.2009.17768.x>.
 Baker, Ronald, Buckland, Amanda, Sheaves, Marcus, 2014. Fish gut content analysis: robust measures of diet composition. *Fish Fish.* 15 (1), 170–177. <https://doi.org/10.1111/faf.12026>.
 Barnes, Carolyn, Bethea, D.M., Brodeur, R.D., Spitz, J., Ridoux, V., Pusineri, C., Chase, B. C., et al., 2008. Predator and prey body sizes in marine food webs: ecological archives E089-051. *Ecology* 89 (3), 881–81.
 Barnes, Carolyn, Bethea, D.M., Brodeur, R.D., Spitz, J., Ridoux, V., Pusineri, C., Chase, B. C., et al., 2016. PREDATOR AND PREY BODY SIZES IN MARINE FOOD WEBS. <https://doi.org/10.6084/m9.figshare.c.3300257.v1>.
 Cattin, Marie-France, Bersier, Louis-Felix, Banašek-Richter, Carolin, Baltensperger, Richard, Gabriel, Jean-Pierre, 2004. Phylogenetic constraints and adaptation explain food-web structure. *Nature* 427 (6977, 6977), 835–839. <https://doi.org/10.1038/nature02327>.
 Chacoff, Natacha P., Vazquez, Diego P., Lomascolo, Silvia B., Stevani, Erica L., Dorado, Jimena, Padron, Benigno, 2012. Evaluating Sampling Completeness in a Desert Plant–Pollinator Network. *J. Anim. Ecol.* 81 (1), 190–200. <https://doi.org/10.1111/j.1365-2656.2011.01883.x>.
 Cohen, Joel E., Mulder, Christian, 2014a. Soil invertebrates, chemistry, weather, human management, and edaphic food webs at 135 sites in the Netherlands: SIZEWEB. *Ecology* 95 (2). <https://doi.org/10.1890/13-1337.1>, 578–78.
 Cohen, Joel E., Mulder, Christian, 2014b. Soil invertebrates, chemistry, weather, human management, and edaphic food webs at 135 sites in the Netherlands: SIZEWEB. *Ecology* 95 (2). <https://doi.org/10.1890/13-1337.1>, 578–78.
 Cohen, Joel E., Newman, C.M., Steele, John Hyslop, 1985. A stochastic theory of community food webs I. models and aggregated data. *Proceed. R. Soc. Lond. Ser. B Biol. Sci.* 224 (1237), 421–448. <https://doi.org/10.1098/rspb.1985.0042>.
 Compson, Zacchaeus G., Monk, Wendy A., Hayden, Brian, Bush, Alex, O’Malley, Zoë, Hajibabaei, Mehrdad, Porter, Teresita M., et al., 2019. Network-based biomonitoring: exploring freshwater food webs with stable isotope analysis and DNA Metabarcoding. *Front. Ecol. Evol.* 7, 395.
 Crawford, Kerry, McDonald, Robbie A., Bearhop, Stuart, 2008. Applications of stable isotope techniques to the ecology of mammals. *Mammal Rev.* 38 (1), 87–107. <https://doi.org/10.1111/j.1365-2907.2008.00120.x>.
 Dengler, Jürgen, 2009. Which function describes the species–area relationship best? a review and empirical evaluation. *J. Biogeogr.* 36 (4), 728–744. <https://doi.org/10.1111/j.1365-2699.2008.02038.x>.
 Djemiel, Christophe, Maron, Pierre-Alain, Terrat, Sébastien, Dequiedt, Samuel, Cottin, Aurélien, Ranjard, Lionel, 2022. Inferring microbiota functions from taxonomic genes: A review. *GigaScience* 11.
 Dunne, Jennifer A., Williams, Richard J., Martinez, Neo D., 2002. Network structure and biodiversity loss in food webs: robustness increases with Connectance. *Ecol. Lett.* 5 (4), 558–567. <https://doi.org/10.1046/j.1461-0248.2002.00354.x>.
 Figueroa, David, 2007. Food Web Dynamics: New Patterns from Southern South America and North Wales UK, and the Role of Basal Species Structuring Food Webs. PhD thesis, University of London.
 Figueroa, David H., 2022a. Afon Hirnant gut content dataset [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.6655759>.
 Figueroa, David H., 2022b. Coilaco, Guampoe and Trancura gut content dataset [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.6655771>.
 Gilljam, David, Thierry, Aaron, Edwards, Francois K., Figueroa, David, Ibbotson, Anton T., Iwan Jones, J., Lauridsen, Rasmus B., Petchey, Owen L., Woodward, Guy, Ebenman, Bo, 2011. Seeing double. In: *Advances in Ecological Research*, 45. Elsevier, pp. 67–133. <https://doi.org/10.1016/B978-0-12-386475-8.00003-4>.
 Goldwasser, Lloyd, Roughgarden, Jonathan, 1993a. Construction and analysis of a large Caribbean food web. *Ecology* 74 (4), 1216–1233. <https://doi.org/10.2307/1940492>.
 Goldwasser, Lloyd, Roughgarden, Jonathan, 1993b. Construction and analysis of a large Caribbean food web. *Ecology* 74 (4), 1216–1233. <https://doi.org/10.2307/1940492>.
 Gravel, Dominique, Poisot, Timothée, Albouy, Camille, Velez, Laure, Mouillot, David, 2013. Inferring Food Web Structure from Predator–Prey Body Size Relationships. Edited by Robert Freckleton *Methods Ecol. Evol.* 4 (11), 1083–1090. <https://doi.org/10.1111/2041-210X.12103>.
 Gray, Clare, Figueroa, David H., Hudson, Lawrence N., Ma, Athen, Perkins, Dan, Woodward, Guy, 2015. Joining the dots: an automated method for constructing food webs from compendia of published interactions. *Food Webs* 5 (December), 11–20. <https://doi.org/10.1016/j.fooweb.2015.09.001>.
 Gupta, Anubhav, Furrer, Reinhard, Petchey, Owen L., 2022. Simultaneously estimating food web Connectance and structure with uncertainty. *Ecol. Evol.* 12 (3), e8643. <https://doi.org/10.1002/ece3.8643>.
 Hattab, Tarek, Leprieur, Fabien, Lasram, Frida Ben Rais, Gravel, Dominique, Le, François, Loc’h, and Camille Albouy., 2016. Forecasting fine-scale changes in the food-web structure of coastal marine communities under climate change. *Ecography* 39 (12), 1227–1237. <https://doi.org/10.1111/ecog.01937>.
 Hyslop, E.J., 1980. Stomach contents analysis—a review of methods and their application. *J. Fish Biol.* 17 (4), 411–429. <https://doi.org/10.1111/j.1095-8649.1980.tb02775.x>.
 Ings, Thomas C., Montoya, Jose M., Bascompte, Jordi, Blüthgen, Nico, Brown, Lee, Dormann, Carsten F., Edwards, François, et al., 2009. Ecological Networks—Beyond

- Food Webs. *J. Anim. Ecol.* 78 (1), 253–269. <https://doi.org/10.1111/j.1365-2656.2008.01460.x>.
- Jennings, Simon, Brander, Keith, 2010. Predicting the effects of climate change on marine communities and the consequences for fisheries. *J. Mar. Syst.* 79 (3–4), 418–426. <https://doi.org/10.1016/j.jmarsys.2008.12.016>.
- Jennings, Simon, van der Molen, Johan, 2015. Trophic levels of marine consumers from nitrogen stable isotope analysis: estimation and uncertainty. *ICES J. Mar. Sci.* 72 (8), 2289–2300. <https://doi.org/10.1093/icesjms/fsv120>.
- Jones, J. Iwan, Edwards, Francois K., Lauridsen, Rasmus B., 2022. Tadnoll brook gut content dataset. Zenodo. <https://doi.org/10.5281/zenodo.6911932>.
- Jord'an, Ferenc., 2009. Keystone species and food webs. *Philos. Trans. R. Soc. B: Biol. Sci.* 364 (1524), 1733–1741. <https://doi.org/10.1098/rstb.2008.0335>.
- Jordano, Pedro, 2016. Sampling networks of ecological interactions. *Funct. Ecol.* 30 (12), 1883–1893. <https://doi.org/10.1111/1365-2435.12763>.
- Kadoya, Taku, Osada, Yutaka, Takimoto, Gaku, 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>.
- Kelly, Ryan P., Port, Jesse A., Yamahara, Kevan M., Crowder, Larry B., 2014. Using environmental DNA to census marine fishes in a large Mesocosm. *PLoS One* 9 (1), e86175. <https://doi.org/10.1371/journal.pone.0086175>.
- Layman, Craig A., Albrey Arrington, D., Montaña, Carmen G., Post, David M., 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).
- Lindgren, Martin, Möllmann, Christian, Nielsen, Anders, Brander, Keith, MacKenzie, Brian R., Chr, Nils, Stenseth., 2010. Ecological forecasting under climate change: the case of Baltic Cod. *Proc. R. Soc. B Biol. Sci.* 277 (1691), 2121–2130. <https://doi.org/10.1098/rspb.2010.0353>.
- MacArthur, Robert H., Pianka, Eric R., 1966. On optimal use of a patchy environment. *Am. Nat.* 100 (916), 603–609. <https://www.jstor.org/stable/2459298>.
- Nielsen, Jens M., Clare, Elizabeth L., Hayden, Brian, Brett, Michael T., Kratina, Pavel, 2018. Diet tracing in ecology: method comparison and selection. Edited by M. Thomas P. Gilbert *Methods Ecol. Evol.* 9 (2), 278–291. <https://doi.org/10.1111/2041-210X.12869>.
- O'Gorman, Eoin J., Petchey, Owen L., Faulkner, Katy J., Gallo, Bruno, Gordon, Timothy A.C., Neto-Cerejeira, Joana, Olafsson, Jon S., Pichler, Doris E., Thompson, Murray S. A., Woodward, Guy, 2019. A simple model predicts how warming simplifies wild food webs. *Nature Climate Change* 9 (8, 8), 611–616. <https://doi.org/10.1038/s41558-019-0513-x>.
- Peralta-Maraver, I., Lopez-Rodriguez, M.J., Tierno de Figueroa, J.M., 2017. Structure, dynamics and stability of a Mediterranean River food web. *Mar. Freshw. Res.* 68 (3), 484–495. <https://doi.org/10.1071/MF15154>.
- Petchey, Owen L., Beckerman, Andrew P., Riede, Jens O., Warren, Philip H., 2008. Size, foraging, and food web structure. *Proc. Natl. Acad. Sci.* 105, 4191–4196. <https://doi.org/10.1073/pnas.0710672105>.
- Pompanon, Francois, Deagle, Bruce E., Symondson, William O.C., Brown, David S., Jarman, Simon N., Taberlet, Pierre, 2012. Who is eating what: diet assessment using next generation sequencing. *Mol. Ecol.* 21 (8), 1931–1950. <https://doi.org/10.1111/j.1365-294X.2011.05403.x>.
- Potapov, Anton M., Brose, Ulrich, Scheu, Stefan, Tiunov, Alexei V., 2019. Trophic Position of Consumers and Size Structure of Food Webs across Aquatic and Terrestrial Ecosystems. *The American Naturalist*. <https://doi.org/10.1086/705811>. August, 000–000.
- Rindorf, A., Lewy, P., 2004. Bias in Estimating Food Consumption of Fish by Stomach-Content Analysis, 61, p. 12.
- Roslin, Tomas, Majaneva, Sanna, 2016. The use of DNA barcodes in food web construction—terrestrial and aquatic ecologists unite! Edited by Elizabeth Clare Geinome 59 (9), 603–628. <https://doi.org/10.1139/gen-2015-0229>.
- Sentis, Arnaud, Hemptinne, Jean-Louis, Brodeur, Jacques, 2014. Towards a mechanistic understanding of temperature and enrichment effects on species interaction strength, omnivory and food-web structure. Edited by Micky Eubanks *Ecol. Lett.* 17 (7), 785–793. <https://doi.org/10.1111/ele.12281>.
- Tamaddon-Nezhad, Alireza, Milani, Ghazal Afroozi, Raybould, Alan, Muggleton, Stephen, Bohan, David A., 2013. Construction and validation of food webs using logic-based machine learning and text mining. In: *Advances in Ecological Research*, 49. Elsevier. <https://doi.org/10.1016/B978-0-12-420002-9.00004-4>, 225–89.
- Vagnon, Chloe, Cattaneo, Franck, Goulon, Chloe, Grimardias, David, Guillard, Jean, Frossard, Victor, 2021. An allometric niche model for species interactions in temperate freshwater ecosystems. *Ecosphere* 12 (3), e03420. <https://doi.org/10.1002/ecs2.3420>.
- Valdovinos, Fernanda S., Ramos-Jiliberto, Rodrigo, Garay-Narvaez, Leslie, Urbani, Pasquinell, Dunne, Jennifer A., 2010. Consequences of adaptive behaviour for the structure and dynamics of food webs: adaptive behaviour in food webs. *Ecol. Lett.* 13 (12), 1546–1559. <https://doi.org/10.1111/j.1461-0248.2010.01535.x>.
- Wada, E., Mizutani, H., Minagawa, M., 1991. The use of stable isotopes for food web analysis. *Crit. Rev. Food Sci. Nutr.* 30 (4), 361–371. <https://doi.org/10.1080/10408399109527547>.
- 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., Martinez, Neo D., 2000. Simple rules yield complex food webs. *Nature* 404 (6774, 6774), 180–183. <https://doi.org/10.1038/35004572>.
- Woodward, Guy, 2021. Broadstone stream gut content data. Zenodo. <https://doi.org/10.5281/zenodo.5575040>.
- Woodward, Guy, Blanchard, Julia, Lauridsen, Rasmus B., Edwards, Francois K., Jones, J. Iwan, Figueroa, David, Warren, Philip H., Petchey, Owen L., 2010. Individual-based food webs, 43. Elsevier. <https://doi.org/10.1016/B978-0-12-385005-8.00006-X>, 211–66.