

Navigating food web prediction; assumptions, rationale, and methods

Tanya Strydom ¹; Jennifer A. Dunne ²; Timothée Poisot ^{3,4}; Andrew P. Beckerman ¹

Abstract: TODO

Keywords: food web, network construction, scientific ignorance

1 key ‘aim’ is to highlight how we need to think of interactions at scales and the same goes fro
2 predicting. this is important becasue it will influence the we way in which we are actually able
3 to ude a network

4 At the heart of modern biodiversity science are a set of concepts about biodiversity, community structure,
5 productivity, and asynchrony, and how they define the stability, resilience, and dynamics of complex com-
6 munities. The use of species interaction networks provides a powerful abstraction that one can use to help
7 quantify, conceptualise, and understand these concepts. However, network ecology has its own nuance and
8 idiosyncrasies that not only provide a barrier to entry but causes dissonance even within the field (Dormann,
9 2023). This is perhaps particularly pervasive within the space of network prediction...

10 One of the fundamental challenges that we are faced when working with and studying interaction networks
11 (and, within the context of this manuscript, specifically food webs) is that there is a scarcity of ‘real world’
12 interaction data (Hortal et al., 2015; Poisot et al., 2021). The difficulty of recording interactions in the field
13 (Jordano, 2016a, 2016b) has necessitated that researchers find and develop alternative means to construct
14 and build food webs using **models** (Morales-Castilla et al., 2015; Strydom et al., 2021). Over the past
15 decade, there has been a proliferation of tools and processes for characterising food webs, these models
16 span a wide range of philosophies that rely on different approaches, data, and definitions, which ultimately
17 determine how the food web is constructed and coded. Although the development of these different models
18 have carved out the path for constructing either synthetic, ecologically plausible networks (Poisot, Gravel,
19 et al., 2016), or providing ‘first draft’ networks that can be utilised in real world settings (Strydom et al.,
20 2022) we are still lacking in discussions that are explicitly comparing and contrasting how the way one
21 chooses to approach the task of constructing a food web is introducing (and ultimately embedding) specific
22 assumptions and hypotheses (Petchey et al., 2008). Most attempts that focus on comparing and contrasting
23 models are focused on the same group of **model families** (Pichler et al., 2020; Williams & Martinez, 2008)
24 and only benchmark the different models as opposed to contextualising them within the bigger framework
25 of understanding the data needs of the different models, as well as how the resulting network is defined and
26 structured. As food webs become a more integrated part of some of the broader fields of ecology (Bhatia et
27 al., 2023; Thuiller et al., 2024) it is critical that we review these different model families as a whole (not only
28 in isolation), and move away from simply benchmarking the performance of these different model families.
29 This is important because different models impose different constraints upon themselves and will not only
30 delimit and dictate the potential questions one will be able to ask (Petchey et al., 2011) but also determine
31 the appropriate research setting for which the model (and resulting network) can be used. For example
32 the use of ‘structural food webs’ are useful for developing additional theory such as re-wiring of networks

³³ (Staniczenko et al., 2010) but would be meaningless if one's intention is to produce a location-specific network
³⁴ (Dansereau et al., 2023). This will allow us to ensure the right models are being used to answer the right
³⁵ questions, particularly within the context of trying to accelerate cross-cutting research in the face of global
³⁶ change.

³⁷ When navigating the seas of using and constructing food webs the researcher needs to be able to clearly
³⁸ articulate and define the parameters that are used to define their food web(s) of interest. This will aid them
³⁹ in being able to select the correct model to help them to reach their goal. In order to be able to make
⁴⁰ informed decisions it is important that one has a strong grasp of exactly what it means to 'code'/define a
⁴¹ food web, a clear understanding of why one wants to predict a food web, and ultimately one needs to be
⁴² able to asses and evaluate which model family is going to best match up with the goal of network prediction.
⁴³ Here we specifically aim to look at not look at only the performance of the different models but also initiate
⁴⁴ a (thus far lacking) discussion around how the interplay between the language used to define networks and
⁴⁵ the underlying theory/philosophy should also be a part of the broader discussion when it comes to the task
⁴⁶ of 'model selection'.

⁴⁷ 1 Understanding the drivers of species interactions

⁴⁸ Important goal here is to introduce the idea that there are multiple facets as to what determines
⁴⁹ the interaction between species and that there is some sort of 'scale of organisation' Figure 1. We
⁵⁰ can then introduce these different scales/theories and I think some key points to highlights are the
⁵¹ *features, limits, and descriptions* of these different scales (by that I mean what rules them, what
⁵² finds them, and what binds them - sorry not sorry). I think it is also worth either in this section
⁵³ or in the one where we talk about model families to discuss the idea of 'moving between' different
⁵⁴ levels - e.g. downsampling but also feasibility - can we actually do that? ANother interesting
⁵⁵ discussion here (or maybe actually something that can make its way into the concluding remarks)
⁵⁶ is thinking about what determines interactions vs what determines structure...

- ⁵⁷ • These different theories are shown in Figure 1 and we can see there is some element of scaling (species
⁵⁸ - population - individual)

⁵⁹ [Figure 1 about here.]

⁶⁰ 1. (Co)occurrence

⁶¹ Although the outright assumption that because two species are co-occurring it must mean that they are

62 interacting is inherently flawed (Blanchet et al., 2020), it is of course impossible for two species to interact (at
63 least in terms of feeding links) if they are not co-occurring in time and space. Hence it is of course important
64 to take into consideration the co-occurrence of both the resource and the consumer. An example of this
65 would be the work from Dansereau et al. (2023), where a metaweb (feasibility network) is downsampled into
66 smaller realisations based on better data/knowledge as to which species are occurring at a specific location
67 - however arguably these are still firmly in the space of feasible interactions for the specific location but are
68 approaching a better approximation of ‘reality’...

69 **2. Feasibility**

70 This is based on the idea of forbidden links introduced by Jordano (2016b), specifically that there must
71 be some degree of *trait complementarity* that allows a predator to chase, capture, kill, and consume, its
72 prey. This is probably the level that the idea of a metaweb (Dunne, 2006) is most applicable to. Within
73 the network prediction ‘field’ this is perhaps the most developed space. Predictive models run the gamut
74 including mechanistic models (Morales-Castilla et al., 2015), binary classifiers (Pichler et al., 2020), and
75 graph embedding (Strydom et al., 2023) and use either traits (or phylogeny as a proxy for the conservation
76 thereof) as a means to ‘evaluate’ if an interaction is *possible* between two species (again not the likelihood
77 of it happening but the likelihood of its feasibility). It is probably worth having a brief interlude here to be
78 really clear that just because an interaction is probabilistic it does not make it weighted (at least not in the
79 traditional sense of weighted interactions, *e.g.*, J. T. Wootton & Emmerson (2005)) - it is still ‘binary’, it
80 just happens to be defined by a binomial distribution (*sensu* Banville, in prep).

81 **3. Mass effect**

82 Not sure if there are models that ‘only’ consider abundance (barring the neutral model) and that it is rather
83 more of a building block in some of the models that are more relevant to the next steps. Maybe there is an
84 argument that this ‘rule’ is ‘irrelevant’ in the context of how I am presenting network prediction and more
85 so a data parameter one needs... maybe...

86 This is probably the point where we start to shift from a *potential* (presence/absence) way of defining
87 interactions and start moving into the ‘qualitative’/weighted interaction space - we are not ‘determining’ if
88 the interaction is feasible but rather making an assumption on prey selection based on the species’ likelihood of
89 ‘meeting’, although Banville (in prep) presents a compelling case that this could still be considered something
90 that falls under the ‘feasibility’ and not ‘reality’ side of the spectrum... (well at least past Tanya seemed to
91 think so)

92 **4. Energetics**

93 This is where we begin to move into the foraging ecology space - specifically consumption rate and how that
94 pertains to energy acquisition *i.e.*, optimal foraging theory. In the loosest sense I think this is the ‘prey choice’
95 space - but specifically in the context of how prey choice as informed by energetic cost (not just purely based
96 on *e.g.*, the most abundant species). If we think about ways that people have approached this there are the
97 diet models of (Beckerman et al., 2006) and (Petchey et al., 2008) as well as the ‘trait’ framework developed
98 by K. L. Wootton et al. (2023) that moves the ‘energy’ into different ‘modules’ related to the process of
99 the consumer acquiring energy from the resource (however there is a disregard for the ‘Rule 1’ requirement
100 of forbidden links, again not bad just pointing it out). The idea of the consumer search space developed
101 by Pawar et al. (2012) is also an interesting consideration. Finally the environment itself is also imposing
102 energy costs on the predator. Basically the ideas presented in Cherif et al. (2024), which is essentially a
103 take on movement ecology? What it boils down to is being able to quantify the cost of movement *i.e.*, the
104 physical constraints that the environment imposes on a species... Maybe we can also think of it more in terms
105 of metabolic rate?

i Box 1 -The anatomy of a food web

Important goal of this box is to highlight the different terminology that is used to describe a food web but especially in the context of the ‘feasibility’ vs ‘realised’ food webs. Energy v ‘feasibility’ being the underlying driver but also determinant of the ‘final’ network - this is what is going to influence the ways in which we are defining the network as well as how we can use them. i.e this may seem very fundamental but is something that is overlooked...

Defining a food web seems simple, it is the representation of the interactions (edges) between species (nodes), however the definition of ‘edges’ and ‘nodes’, as well as the scale at which they are aggregated can take many forms. As highlighted in Poisot, Stouffer, et al. (2016) networks can be constructed at the population (the links among individuals), community (the links between species), or metacommunity (fluxes between locations) level. Even if one were to limit their scope to thinking of interaction networks only in terms of food webs at the community-level there are still many ways to define the various components of the network Panel A of 2, one needs to understand the different intentions/assumptions that are made when a food web is constructed. Although the main intention of a food web is to capture and represent the feeding links between species there are many ways to define the nodes (*e.g.*, species or taxonomic group), edges (*e.g.*, **potential or realised feeding links**), the magnitude of the edges (*e.g.*, binary vs probabilistic), and even how the network itself is delimited (does it represent an aggregation of interactions over time?).

[Figure 2 about here.]

1.0.1 How do we define a node?

Although this may seem an elementary question in the context of food webs — a node *should* represent a (taxonomic) species, the reality is that nodes can often represent an aggregation of different species - so called ‘trophic species’ or segregation of species by life stages. Representing nodes as non-taxonomic species can be useful in certain contexts (Williams & Martinez, 2000) and in cases where the adult and larval stages of a species have different diets it may make ecological sense (Clegg et al., 2018) meaning that it is not uncommon that networks often have nodes that have different definitions of a ‘species’ *e.g.* consisting of both taxonomic and trophic species. Practical implications of how we are aggregating the nodes is that the resolution may not always be ‘pixel perfect’ *i.e.*, we may be unable to assess the co-extinction risk of a species pair, however there is value in having nodes that represent an aggregation of species, as these convey a much more general overview of how the links are distributed within the community.

1.0.2 What is meant by an edge?

As discussed earlier there are many ways to define the links between species — even feeding links. At its core links within food webs can be thought of as a representation of either the flow of a resource [ref], realised (Pringle, 2020) or potential (Dunne, 2006) feeding links, or energy transfer and material flow (Lindeman, 1942). How we specify links will influence the resulting structure of the network - and the inferences we will make thereof. For example taking a food web that consists of links representing *potential* feeding links between species will be meaningless if you are interested in understanding *e.g.*, the flow of energy through the system as the links within the network are over overrepresented. In addition to the various ways of defining the links between species pairs there are also a myriad of ways in which the links themselves can be quantified. Links between species are often treated as being present or absent (*i.e.*, binary) but it is also possible to use probabilities (which quantifies how likely an interaction is to occur, Poisot, Cirtwill, et al., 2016) or continuous measurements (which quantifies the strength of an interaction, Berlow et al., 2004). Moving away from a purely binary way of representing allows us to quantify a level of (un)certainty of our knowledge of interactions (*i.e.*, moving from being able to ask if they are occurring to quantifying how likely they are to occur) does add an additional level of ‘complexity’ to the construction and interpretation of networks, but ultimately it allows us to capture more information at different scales (Banville, in prep).

1.0.3 Putting the parts together; what does it mean?

The ingredients one uses to construct networks from nodes and edges generates a unique representation of the mechanisms (see Box 1 - Mechanisms that determine feeding links) that allow inference and reasoning about the structure, aspects of dynamics (*e.g.*, stability), and potentially the function of communities (*e.g.*, flux). It is thus beneficial to keep in mind that in the process of ‘codifying’ a network one is already embedding some sort of hypothesis as to the nature of the feeding links between species (Brimacombe et al., 2023; Proulx et al., 2005). Here it may be meaningful to contextualise the different ‘types’ of food webs within the larger research programmes (or even practical needs) that have been driving the construction of them.

Before thinking about the ways in which we can predict networks it is perhaps meaningful to take a step back and think about the different criteria that must be met in order for an interaction to be able to occur between two species, specifically thinking of this in terms of distinguishing between the feasibility versus realisation of an interaction and how these are determined (and defined by) different ‘rules’/mechanisms. If we look at this feasibility-reality continuum (Figure 1) it is clear how the different predictive approaches (methods) tend to fall within one of the broader categories identified (distinguished) in the triangle. This is not to say that this shortcoming should be viewed as a ‘bug’ but rather a ‘feature’ of the field as it allows one to engage with, as well as construct networks at different scales, which is particularly valuable if one takes into consideration the considerable ‘data cost’ of predicting well resolved, realised networks in comparison to constructing high-level metawebs. However, it is important that there is an awareness and acknowledgement of where within this feasibility-reality one is working at and how this will impact and limit the contexts in which the resulting network can be used and applied within.

108

109 2 Network prediction is scale dependant

110 The way in which we predict a network is driven by the underlying theory Figure 1 which con-
111 strains or informs the assumptions we make (this of course also has implications with regards
112 to how the resulting network is defined (Box 1)). We can then spend a moment introducing
113 the different model families Table 1. I think a clear messaging here might be that models can
114 share a similar underlying theory but use different methods to get there (*e.g.*, using ecological
115 rules (explicit), ecological expectations (pattern finders), or mathematical models (assumptions
116 on the structure of the matrix - maybe even network)). Importantly different models will also

117 have different ‘limits’ to them - this is probably a product of both where they are found within
118 the ‘theory space’ Figure 1 as well as the definition of the network (Box 1) space. Should we
119 also maybe revisit the idea of interaction vs structure predictors... I think it is still a point that
120 is owrth raising but no longer the framework on which we hanf the different model families...

- 121 • The way in which we predict networks is ‘constrained’/informed by the different theories shown in
122 Figure 1
- 123 • Need to be aware of this and be aware how/what we can use the networks - Petchy dilemma
- 124 • The ‘scale’ that a network is constructed should be a determinant of what we can learn about a system
125 e.g., can’t use a feasibility network to learn something about energy flows. This is because they are
126 capturing different processes
- 127 • Link the ‘model families’ to the different scales/theories
- 128 • Data...

129 As discussed in Box 1 there are many ways to define a food web, meaning that there are equally as many
130 reasons one might be interested in predicting a food web. However we may think of two primary drivers
131 for wanting to predict networks (Panel B Figure 2), namely an interest in generating a set of ecologically
132 plausible networks (*i.e.*, being able to describe networks using a model) or being able to recover (predict)
133 location specific, ‘realised’, interactions for a specific species community (*i.e.*, being able to predict/infer the
134 interactions between species). Of course these two categories are not distinct, mutually exclusive, groups but
135 can rather be viewed as operating on a continuum ranging from a need for generality (*i.e.*, creating a network
136 that, when taken in aggregate, the distribution of links (interactions) between nodes (species) are ecologically
137 plausible) to a need for specificity (*i.e.*, local-level predictions between specific species pairs). Although the
138 ability to predict ‘real-world’ interactions (and the resulting food webs) can have more intuitive ‘real world’
139 applications e.g., being able to ‘recover’ food webs that have since gone extinct (Dunne et al., 2008; Yeakel
140 et al., 2014), using pairwise interactions to understand species distributions (Pollock et al., 2014) or even
141 co-extinction risk (Dunn et al., 2009), a more structural approach to network construction affords one an
142 opportunity to interrogate some of the more high-level mechanisms that are structuring networks (Box 1).

143 It is perhaps more important that when one is talking about ‘why’ they want to predict networks to articulate
144 exactly what anatomical part of the food web we are interested in scrutinising.

145 **2.1 How do we predict food webs?**

146 Selecting a model for the task of network prediction should come down to two things; what *aspect* of a
147 food web one is interested in predicting, and what data are available, necessary, and sufficient. As shown
148 in panel B of Figure 2 the interest in a network is (usually) at either the ‘structural’ or ‘interaction’ level
149 and the development of models for the task of network prediction often focus on high fidelity (performance)
150 at one of these scales. With this in mind it is beneficial to think of the different model families relative
151 to these two different goals; here we refer to models that are used to predict the structure of a network as
152 **topology generators** and models developed to infer the interactions for a given species pool as **interaction**
153 **predictors**. It is meaningful to make this distinction because although it is possible to construct a food
154 web given using an *interaction predictor* the models themselves lack any sort of parametrisation of the
155 network structure and so the resulting network is a poor reflection of the actual network structure (Caron
156 et al., 2024). This is primarily because *interaction predictors* are models that evaluate the feasibility of
157 an interaction between species pairs and not in the context of feasibility at the community level. Models
158 themselves are a reflection of the different goals and intentions of the research program from which they are
159 developed and are often ‘described’ by a specific mechanism that will determine the resulting structure or
160 interactions (Box 1). Models such as the niche (Williams & Martinez, 2000) or cascade (Cohen et al., 1990)
161 were developed with the intent of being used to understand the *structural* aspects of food webs, specifically
162 how links are distributed amongst species in the community, whereas bayesian (Cirtwill et al., 2019) or trait
163 hierarchy (Shaw et al., 2024) models have been developed on the basis that the traits of a species are the
164 underlying mechanism in determining the feasibility of interactions (*i.e.*, species *a* has the capacity to eat
165 species *b*). Along with predicting different anatomical parts of a food web the different models have varying
166 degrees of data that are needed to ‘parametrise’ the network. Once these two limitations are assessed and
167 addressed it is then possible to select the model (or model family) that will best be able to capture food web
168 feature that the researcher is most interested in (see Box 2 - Assessing model outputs). It is thus clear that
169 (realistically) there will probably never be a ‘best fit’ tool that is able to construct a food web that will span
170 the entire range of needs, and rather the responsibility lies with the researcher to be aware of not only the
171 underlying philosophy of the specific toolset (as this could have knock-on effects when using those networks
172 for downstream analyses/simulations; pers. comms. Beckerman, 2024), but also how well the tool is able to
173 retrieve the specific network or interaction properties that is of interest.

174 In order for a model to formalise a ‘complete’ food web it is necessary to formalise two aspects
175 of the network, ‘who eats whom’ (to determine the links between nodes) as well as the structure
176 of the network (to limit the distribution of links), however most models are inclined to focus on

177 one of the two aspects panel B of 2.

178 Crucially most topology generators lack some key data on the interaction between species (this
179 can be because of how the model itself defines species or the way in which links are assigned in
180 the network) and interaction predictors lack some sort of parametrisation of network structure
181 (just because two species can interact it does not mean that they will, Poisot et al., 2015).

182 What is the purpose of generating a network? Is it an element of a bigger question we are asking,
183 e.g., I want to generate a series of networks to do some extinction simulations/bioenergetic stuff
184 OR are we looking for a ‘final product’ network that is relevant to a specific location? (this can
185 still be broad in geographic scope).

186 2.2 Model families

187 As there are many food web models to choose from it is perhaps useful to think about the models in terms of
188 model families, a summary of these families is presented in Table 1 highlights the differences and similarities
189 of the philosophies and assumptions that determine a network. A more extensive overview of the different
190 models that fall with in the different model families can be found in SuppMat 1 and for a more detailed
191 breakdown of the different ‘traits’ of the model families refer to SuppMat 2.

Table 1: A summary of the different families of tools that can be used to generate food webs, this includes a brief description of the underlying philosophy of the family as well as how the different elements (nodes and edges) of the generated network represents.

| Model family | Assumption | Theory | Nodes | | |
|--------------|---|--------|-----------|-------------|---------------|
| | | | represent | Interaction | Key reference |
| null | Links are randomly distributed within a network | | agnostic | binary | |

| Model family | Assumption | Theory | Nodes | | |
|-----------------|---|--------------------|-----------------|---------------|----------------------------|
| | | | represent | Interaction | Key reference |
| neutral | Network structure is random, but species abundance determines links between nodes | abundance | species | binary | |
| resource | Networks are interval, species can be ordered on a ‘niche axis’ | Networks are | trophic species | binary | Williams & Martinez (2008) |
| generative | Networks are determined by their structural features | Networks are | agnostic | binary | |
| energetic | Interactions are determined by energetic costs | abundance + energy | species | quantitative | |
| graph embedding | Interactions can be predicted from the latent traits of networks | feasibility | species | probabilistic | Strydom et al. (2023) |

| Model family | Assumption | Theory | Nodes | | |
|--------------------|--|-------------|-----------|-------------|---|
| | | | represent | Interaction | Key reference |
| trait matching | Interactions can be inferred by a mechanistic frame-work/relationships | feasibility | species | binary | Morales-Castilla et al. (2015) |
| binary classifiers | Interactions can be predicted by learning the relationship between interactions and ecologically relevant predictors | feasibility | species | binary | Pichler et al. (2020) |
| expert knowledge | ‘Boots on the ground’ ecological knowledge and observations | feasibility | species | binary | |
| data scavenging | Webscraping to create networks from online databases | | species | binary | Poisot, Gravel, et al. (2016) (f you squint?) |

| Model family | Assumption | Theory | Nodes | | |
|---------------|--|---------------|-----------|-------------|---------------|
| | | | represent | Interaction | Key reference |
| co-occurrence | co-occurrence patterns arise from interactions so we can use these patterns to reverse engineer the interactions | co-occurrence | species | binary | |

3 How do we use networks?

In this section I want to highlight that we don't actually have any clear guidelines as to how we can 'use' networks - which probably stems from both the fact that when I am talking about a network and when someone else is talking about a network we may actually be talking about two very different conceptualisations of 'a network' (this should actually be a selling point in the intro - may have just found my *raison d'être*) as well as that a lot of the ideas that we have about networks are not really tied to any sort of tangible function (i.e. Tim's GeoBon ms thing-y). However we can maybe at least try to present some guidelines - but I think specifically within the sort of Petchy dilemma space and clearly tied to the ideas we discuss in the ms. This includes: understanding the limits of how a network is defined and how the underlying theory impacts the use as well as data?? IDK we need to shoehorn data in here somehow... We can also use this as a gap identifying space and I think the framing can still rest under the limits concept particularly time, space, and boundaries - which will all probably fall under some aspect of biological scale... We can also raise the idea of trust - as in which methods have more support/trust than others. Also what even a 'real' network entails (and this links again back to Tim's stuff) as well as a subtle jab at Pringles notion that the most critical issue in the world of food webs is being able to identify every. single. link. even though there is no real discussion as to what is an 'opportunistic' link vs a link that represents a sustainable energy source for a population (or would it be an individual)...

211 We need to be aware of the parameter space that is possible given a specific definition of a network
212 and operate within those parameters.

213 4 Concluding remarks

214 I think the idea of time and how we are aggregating networks across that should be a prominent
215 feature here...

- 216 In certain situations structure is ‘enough’ but there may be use cases where we are really interested in
217 the node-level interactions *i.e.*, species identity is a thing we care about and need to be able to retrieve
218 specific interactions at specific nodes correctly.

- 219 Why do interaction models do so badly at predicting structure? Nuance of metaweb vs realisation but
220 also time? At the core of it interaction models are trained on existing interaction data; this is data
221 that are most likely closer to a metaweb than a local realisation even if they are being inventoried at a
222 small scale...

223 – We can briefly shoehorn downsampling here maybe??

- 224 It will be interesting to bring up the idea that if a model is missing a specific pairwise link but doing
225 well overall then when does it matter?

226 – The fact that *some* people are concerned about the taxonomic resolution and cascading effects
227 those might have on our understanding of network structure (Pringle, 2020; Pringle & Hutchinson,
228 2020), but that puts us in a place where we are at risk of losing our ability to distinguish the wood
229 from the tree - are we not (at least at times) concerned more with understanding ecosystem level
230 processes than with needing to understand things *perfectly* at the species level.

231 – I don’t think these ‘rare’/nuanced links (e.g. carnivorous hippos) are going to rock the boat when
232 we think about networks at the structural level.

233 “The resolution of food-web data is demonic because it can radically change network topology
234 and associated biological inferences in ways that are unknowable in the absence of better data.” -
235 Pringle & Hutchinson (2020) The counter to this is that structural models are often not working
236 at the species level and thus the structure remains ‘unchanged’ when you increase the resolution
237 - I don’t think that people are that concerned with the structure of real world networks barring
238 connectance and since that scales with species richness anyway your final proportion will probably
239 still remain the same...

240 • I think a big take home will (hopefully) be how different approaches do better in different situations
241 and so you as an end user need to take this into consideration and pick accordingly. I think Petchey
242 et al. (2011) might have (and share) some thoughts on this. I feel like I need to look at Berlow et al.
243 (2008) but maybe not exactly in this context but vaguely adjacent.

244 – I think this is sort of the crux of the argument presented in Brimacombe et al. (2024) as well.

245 “we highlight an interesting paradox: the models with the best performance measures are not
246 necessarily the models with the closest reconstructed network structure.” - Poisot (2023)

247 • Do we need network models to predict interactions and interaction models to predict structure?

248 – “Another argument for the joint prediction of networks and interactions is to reduce circularity
249 and biases in the predictions. As an example, models like linear filtering generate probabilities of
250 non-observed interactions existing, but do so based on measured network properties.” - Strydom
251 et al. (2021)

252 – Aligning (dove-tailing) with this the idea of ensemble modelling as presented by Becker et al.
253 (2022)

254 • Close out with a call to action that we have models that predict networks very well and models that
255 predict interactions very well but nothing that is doing well at predicting both - this is where we should
256 be focusing our attention when it comes to furthering model development...

257 • Do we expect there to be differences when thinking about unipartite vs bipartite networks? Is there
258 underlying ecology/theory that would assume that different mechanisms (and thus models) are relevant
259 in these two ‘systems’.

260 – The Terry & Lewis (2020) paper looks at some methods but is specifically looking at a bipartite
261 world...

262 do we bring this up? this could be a box... if we have the ‘finances’ for it... otherwise it should go to the
263 outstanding questions fur sure

264 “That being said, there is a compelling argument for the need to ‘combine’ these smaller functional units
265 with larger spatial networks (Fortin et al., 2021) and that we should also start thinking about the interplay
266 of time and space (Estay et al., 2023). Although deciding exactly what measure might actually be driving
267 differences between local networks and the regional metaweb might not be that simple (Saravia et al., 2022).”

²⁶⁸ 4.1 Time

²⁶⁹ We lack a clear agenda (and conceptualisation) as to what the appropriate level of aggregation is for a
²⁷⁰ ‘network’. Realistically most empirical networks are more aligned with ‘feasibility networks’ as opposed to
²⁷¹ ‘realised networks’ as they are often the result of some sort of aggregation of observations across time. This
²⁷² ‘problem’ is two-fold. Firstly we need to think about how this affects any sort of development of theory
²⁷³ that sits closer to the ‘realised network’ side of the spectrum - how often are we trying to ask and answer
²⁷⁴ questions about realised networks using feasible networks? The second is that this lack of ‘direction’ as to
²⁷⁵ how we should define a network is (actually) probably one of the biggest barriers that is affecting the use of
²⁷⁶ networks in applied settings...

²⁷⁷ Another time perspective question is when do we determine a link to be ‘real’... In the context of feasible
²⁷⁸ networks this is perhaps clearer - all things equal would the predator be able to consume the prey. However
²⁷⁹ in the realised space there is also the question of the long term ‘energetic feasibility’ of an interaction - just
²⁸⁰ because an interaction is possible in the now is it able to sustain a population in the long term. And what
²⁸¹ is the scale for that long term - are we thinking at the generational scale? Because ultimately when we are
²⁸² constructing a network we are aggregating not only across space but also across time.

²⁸³ Glossary

| Term | Definition |
|-----------------------|---|
| food web | a representation of feeding links between species |
| topology generator | a model that predicts a network based on assumptions of structure, this network is species agnostic in the sense that it does not necessarily contain information at the node level |
| interaction predictor | a model that predicts species interactions, these interactions can be used to construct a network but there are no <i>a priori</i> assumptions as that will constrain the network structure |

| Term | Definition |
|------------------------|---|
| model | A tool that can be used to construct food webs, where the resulting network is a representation of a real world network. Models typically only capture specific elements of real world networks and are intended to be used in specific settings |
| model family | A family of models that share an underlying philosophy when it comes to the mapping, pragmatism, and reduction of a network. Families have the same underlying philosophies and assumptions that determine the links between nodes as well as how these may be encoded |
| metaweb | A network that represents <i>all</i> the potential links between species. Importantly these links will not necessarily all be realised in a specific location for a specific time |
| realised network | A network that represents the links between species that are occurring. These networks represent a very localised network... |
| potential feeding link | links that indicate that an interaction is ecologically feasible but not realised <i>per se</i> (a metaweb would contain potential feeding links) |
| realised feeding link | links that indicate that the interaction is realised ‘in the field’. (a realised network contains realised feeding links) |
| confusion matrix | captures the number of true positives (interaction predicted as present when it is present), false negatives (interaction predicted as absent when it is present), false positives (interaction predicted as present when it is absent), and true negatives (interaction predicted as absent when it is absent) |

²⁸⁴ **Outstanding questions**

- ²⁸⁵ • non-consumptive effects
- ²⁸⁶ • how do we define the spatial and temporal ‘boundaries’ of a network?
- ²⁸⁷ • how do we define a ‘real’ network?

²⁸⁸ **References**

- ²⁸⁹ Becker, D. J., Albery, G. F., Sjodin, A. R., Poisot, T., Bergner, L. M., Chen, B., Cohen, L. E., Dallas, T. A., Eskew, E. A., Fagre, A. C., Farrell, M. J., Guth, S., Han, B. A., Simmons, N. B., Stock, M., Teeling, E. C., & Carlson, C. J. (2022). Optimising predictive models to prioritise viral discovery in zoonotic reservoirs. *The Lancet Microbe*, 3(8), e625–e637. [https://doi.org/10.1016/S2666-5247\(21\)00245-7](https://doi.org/10.1016/S2666-5247(21)00245-7)
- ²⁹⁰ Beckerman, A. P., Petchey, O. L., & Warren, P. H. (2006). Foraging biology predicts food web complexity. *Proceedings of the National Academy of Sciences*, 103(37), 13745–13749. <https://doi.org/10.1073/pnas.0603039103>
- ²⁹¹ Berlow, E. L., Brose, U., & Martinez, N. D. (2008). The “Goldilocks factor” in food webs. *Proceedings of the National Academy of Sciences*, 105(11), 4079–4080. <https://doi.org/10.1073/pnas.0800967105>
- ²⁹² Berlow, E. L., Neutel, A.-M., Cohen, J. E., de Ruiter, P. C., Ebenman, B., Emmerson, M., Fox, J. W., Jansen, V. A. A., Iwan Jones, J., Kokkoris, G. D., Logofet, D. O., McKane, A. J., Montoya, J. M., & Petchey, O. (2004). Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73(3), 585–598. <https://doi.org/10.1111/j.0021-8790.2004.00833.x>
- ²⁹³ Bhatia, U., Dubey, S., Gouhier, T. C., & Ganguly, A. R. (2023). Network-based restoration strategies maximize ecosystem recovery. *Communications Biology*, 6(1), 1–10. <https://doi.org/10.1038/s42003-023-05622-3>
- ²⁹⁴ Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23(7), 1050–1063. <https://doi.org/10.1111/ele.13525>
- ²⁹⁵ Brimacombe, C., Bodner, K., & Fortin, M.-J. (2024). *Applying a method before its proof-of-concept: A cautionary tale using inferred food webs*. <https://doi.org/10.13140/RG.2.2.22076.65927>
- ²⁹⁶ Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T., & Fortin, M.-J. (2023). Shortcomings of reusing species interaction networks created by different sets of researchers. *PLOS Biology*, 21(4), e3002068. <https://doi.org/10.1371/journal.pbio.3002068>
- ²⁹⁷ Caron, D., Brose, U., Lurgi, M., Blanchet, F. G., Gravel, D., & Pollock, L. J. (2024). Trait-matching models predict pairwise interactions across regions, not food web properties. *Global Ecology and Biogeography*,

- 314 33(4), e13807. <https://doi.org/10.1111/geb.13807>
- 315 Cherif, M., Brose, U., Hirt, M. R., Ryser, R., Silve, V., Albert, G., Arnott, R., Berti, E., Cirtwill, A.,
316 Dyer, A., Gauzens, B., Gupta, A., Ho, H.-C., Portalier, S. M. J., Wain, D., & Wootton, K. (2024). The
317 environment to the rescue: Can physics help predict predator–prey interactions? *Biological Reviews*,
318 n/a(n/a). <https://doi.org/10.1111/brv.13105>
- 319 Cirtwill, A. R., Ekklf, A., Roslin, T., Wootton, K., & Gravel, D. (2019). A quantitative framework for
320 investigating the reliability of empirical network construction. *Methods in Ecology and Evolution*, 10(6),
321 902–911. <https://doi.org/10.1111/2041-210X.13180>
- 322 Clegg, T., Ali, M., & Beckerman, A. P. (2018). The impact of intraspecific variation on food web structure.
323 *Ecology*, 99(12), 2712–2720. <https://doi.org/10.1002/ecy.2523>
- 324 Cohen, J. E., Briand, F., & Newman, C. (1990). *Community Food Webs: Data and Theory*. Springer-Verlag.
- 325 Dansereau, G., Barros, C., & Poisot, T. (2023). *Spatially explicit predictions of food web structure from
326 regional level data*.
- 327 Dormann, C. F. (2023). The rise, and possible fall, of network ecology. In *Defining Agroecology – A Festschrift
328 for Teja Tscharntke* (pp. 143–159.). Tredition.
- 329 Dunn, R. R., Harris, N. C., Colwell, R. K., Koh, L. P., & Sodhi, N. S. (2009). The sixth mass coextinction:
330 Are most endangered species parasites and mutualists? *Proceedings. Biological Sciences*, 276(1670),
331 3037–3045. <https://doi.org/10.1098/rspb.2009.0413>
- 332 Dunne, J. A. (2006). The Network Structure of Food Webs. In J. A. Dunne & M. Pascual (Eds.), *Ecological
333 networks: Linking structure and dynamics* (pp. 27–86). Oxford University Press.
- 334 Dunne, J. A., Williams, R. J., Martinez, N. D., Wood, R. A., & Erwin, D. H. (2008). Compilation and
335 Network Analyses of Cambrian Food Webs. *PLOS Biology*, 6(4), e102. [https://doi.org/10.1371/journal.pbio.0060102](https://doi.org/10.1371/journal.
336 pbio.0060102)
- 337 Estay, S. A., Fortin, M.-J., & López, D. N. (2023). Editorial: Patterns and processes in ecological networks
338 over space. *Frontiers in Ecology and Evolution*, 11.
- 339 Fortin, M.-J., Dale, M. R. T., & Brimacombe, C. (2021). Network ecology in dynamic landscapes. *Proceedings
340 of the Royal Society B: Biological Sciences*, 288(1949), rspb.2020.1889, 20201889. [https://doi.org/10.1098/rspb.2020.1889](https://doi.org/10.1098/
341 rspb.2020.1889)
- 342 Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven
343 Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and
344 Systematics*, 46(1), 523–549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>
- 345 Jordano, P. (2016a). Chasing Ecological Interactions. *PLOS Biology*, 14(9), e1002559. [https://doi.org/10.1371/journal.pbio.1002559](https://doi.org/10.
346 1371/journal.pbio.1002559)

- 347 Jordano, P. (2016b). Sampling networks of ecological interactions. *Functional Ecology*. <https://doi.org/10.1111/1365-2435.12763>
- 349 Lindeman, R. L. (1942). The Trophic-Dynamic Aspect of Ecology. *Ecology*, 23(4), 399–417. <https://doi.org/10.2307/1930126>
- 351 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from
352 proxies. *Trends in Ecology & Evolution*, 30(6), 347–356. <https://doi.org/10.1016/j.tree.2015.03.014>
- 353 Pawar, S., Dell, A. I., & Savage, V. M. (2012). Dimensionality of consumer search space drives trophic
354 interaction strengths. *Nature*, 486(7404), 485–489. <https://doi.org/10.1038/nature11131>
- 355 Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2008). Size, foraging, and food web
356 structure. *Proceedings of the National Academy of Sciences*, 105(11), 4191–4196. <https://doi.org/10.1073/pnas.0710672105>
- 358 Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2011). Fit, efficiency, and biology: Some
359 thoughts on judging food web models. *Journal of Theoretical Biology*, 279(1), 169–171. <https://doi.org/10.1016/j.jtbi.2011.03.019>
- 361 Pichler, M., Boreux, V., Klein, A.-M., Schleuning, M., & Hartig, F. (2020). Machine learning algorithms
362 to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and
363 Evolution*, 11(2), 281–293. <https://doi.org/10.1111/2041-210X.13329>
- 364 Poisot, T. (2023). Guidelines for the prediction of species interactions through binary classification. *Methods
365 in Ecology and Evolution*, 14(5), 1333–1345. <https://doi.org/10.1111/2041-210X.14071>
- 366 Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., Mercier, B., Violet, C., &
367 Vissault, S. (2021). Global knowledge gaps in species interaction networks data. *Journal of Biogeography*,
368 48(7), 1552–1563. <https://doi.org/10.1111/jbi.14127>
- 369 Poisot, T., Cirtwill, A., Cazelles, K., Gravel, D., Fortin, M.-J., & Stouffer, D. (2016). The structure of
370 probabilistic networks. *Methods in Ecology and Evolution*, 7(3), 303–312. <https://doi.org/10.1111/2041-210X.12501>
- 371 Poisot, T., Gravel, D., Leroux, S., Wood, S. A., Fortin, M.-J., Baiser, B., Cirtwill, A. R., Araújo, M. B.,
372 & Stouffer, D. B. (2016). Synthetic datasets and community tools for the rapid testing of ecological
373 hypotheses. *Ecography*, 39(4), 402–408. <https://doi.org/10.1111/ecog.01941>
- 374 Poisot, T., Stouffer, D. B., & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary
375 through space and time. *Oikos*, 124(3), 243–251. <https://doi.org/10.1111/oik.01719>
- 376 Poisot, T., Stouffer, D. B., & Kéfi, S. (2016). Describe, understand and predict: Why do we need networks
377 in ecology? *Functional Ecology*, 30(12), 1878–1882. <https://www.jstor.org/stable/48582345>
- 378 Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O’Hara, R. B., Parris, K. M., Vesk, P. A., &
379 McCarthy, M. A. (2014). Understanding co-occurrence by modelling species simultaneously with a Joint

- 380 Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, 5(5), 397–406. <https://doi.org/10.1111/2041-210X.12180>
- 381
- 382 Pringle, R. M. (2020). Untangling Food Webs. In *Unsolved Problems in Ecology* (pp. 225–238). Princeton University Press. <https://doi.org/10.1515/9780691195322-020>
- 383
- 384 Pringle, R. M., & Hutchinson, M. C. (2020). Resolving Food-Web Structure. *Annual Review of Ecology, Evolution and Systematics*, 51(Volume 51, 2020), 55–80. <https://doi.org/10.1146/annurev-ecolsys-110218-024908>
- 385
- 386
- 387 Proulx, S. R., Promislow, D. E. L., & Phillips, P. C. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, 20(6), 345–353. <https://doi.org/10.1016/j.tree.2005.04.004>
- 388
- 389 Saravia, L. A., Marina, T. I., Kristensen, N. P., De Troch, M., & Momo, F. R. (2022). Ecological network assembly: How the regional metaweb influences local food webs. *Journal of Animal Ecology*, 91(3), 630–642. <https://doi.org/10.1111/1365-2656.13652>
- 390
- 391
- 392 Shaw, J. O., Dunhill, A. M., Beckerman, A. P., Dunne, J. A., & Hull, P. M. (2024). *A framework for reconstructing ancient food webs using functional trait data* (p. 2024.01.30.578036). bioRxiv. <https://doi.org/10.1101/2024.01.30.578036>
- 393
- 394
- 395 Staniczenko, P. P. A., Lewis, O. T., Jones, N. S., & Reed-Tsochas, F. (2010). Structural dynamics and robustness of food webs. *Ecology Letters*, 13(7), 891–899. <https://doi.org/10.1111/j.1461-0248.2010.01485.x>
- 396
- 397
- 398 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M. J., Fortin, M.-J., Hemming, V., Mercier, B., Pollock, L. J., Runghen, R., Dalla Riva, G. V., & Poisot, T. (2022). Food web reconstruction through phylogenetic transfer of low-rank network representation. *Methods in Ecology and Evolution*, 13(12), 2838–2849. <https://doi.org/10.1111/2041-210X.13835>
- 399
- 400
- 401
- 402 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M. J., Fortin, M.-J., Mercier, B., Pollock, L. J., Runghen, R., Dalla Riva, G. V., & Poisot, T. (2023). Graph embedding and transfer learning can help predict potential species interaction networks despite data limitations. *Methods in Ecology and Evolution*, 14(12), 2917–2930. <https://doi.org/10.1111/2041-210X.14228>
- 403
- 404
- 405
- 406 Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz, N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap towards predicting species interaction networks (across space and time). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1837), 20210063. <https://doi.org/10.1098/rstb.2021.0063>
- 407
- 408
- 409
- 410 Terry, J. C. D., & Lewis, O. T. (2020). Finding missing links in interaction networks. *Ecology*, 101(7), e03047. <https://doi.org/10.1002/ecy.3047>
- 411
- 412 Thuiller, W., Calderón-Sanou, I., Chalmandrier, L., Gaiüzère, P., O'Connor, L. M. J., Ohlmann, M., Poggiato,

- 413 G., & Münkemüller, T. (2024). Navigating the integration of biotic interactions in biogeography. *Journal*
414 *of Biogeography*, 51(4), 550–559. <https://doi.org/10.1111/jbi.14734>
- 415 Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774), 180–183.
416 <https://doi.org/10.1038/35004572>
- 417 Williams, R. J., & Martinez, N. D. (2008). Success and its limits among structural models of complex food
418 webs. *Journal of Animal Ecology*, 77(3), 512–519. <https://doi.org/10.1111/j.1365-2656.2008.01362.x>
- 419 Wootton, J. T., & Emmerson, M. (2005). Measurement of Interaction Strength in Nature. *Annual Review of*
420 *Ecology, Evolution, and Systematics*, 36(1), 419–444. <https://doi.org/10.1146/annurev.ecolsys.36.091704.175535>
- 421 Wootton, K. L., Curtsdotter, A., Roslin, T., Bommarco, R., & Jonsson, T. (2023). Towards a modular theory
422 of trophic interactions. *Functional Ecology*, 37(1), 26–43. <https://doi.org/10.1111/1365-2435.13954>
- 423 Yeakel, J. D., Pires, M. M., Rudolf, L., Dominy, N. J., Koch, P. L., Guimarães, P. R., & Gross, T. (2014).
424 Collapse of an ecological network in Ancient Egypt. *PNAS*, 111(40), 14472–14477. <https://doi.org/10.1073/pnas.1408471111>
- 425
- 426

| JUN | 27 | THU | ○ | |
|---------------|---------|---|-----------------------|---|
| | | 木 | 179 | |
| level of obj. | | (mechanisms) what determines direct interaction | | the "driver" of the interaction ≠ the underlying theory (l:1) & the way we get to the answer is diff e.g. co-occurrence is about sharing space/time but we use niche theory env. filtering as the means to determine co-occurrence. |
| 6 | Species | fund. niches (inv. filter) | co- occurrence | |
| 15 | | buckbones (mora) truit-match | "capacity" | |
| 21 | Pop. | neutral. funct. resp. | pop size/ dynamics | |
| 0 | indiv. | "physics" 2D 3D prey choice | energy | |
| 3 | | ?? allometric scaling?? | | |
| | | : THEORY OF WHAT DETERMINES FEEDING LINKS (INTERACTIONS) BETWEEN SPECIES | | |

Figure 1: TODO.

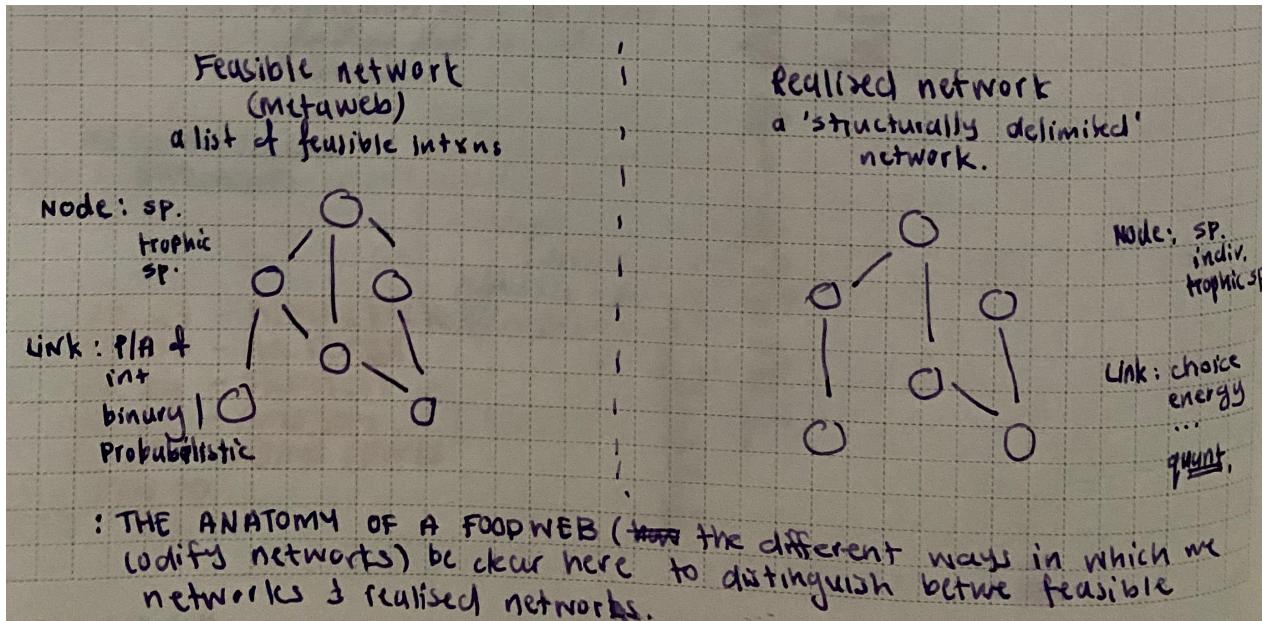


Figure 2: The many ways in which a food web can be defined and described at the node, edge, and even network level.