Navigating food web prediction; assumptions, rationale, and methods

Tanya Strydom

Jennifer A. Dunne

Timothée Poisot

Andrew P. Beckerman

2024-08-21

Abstract

New perspective on key aim is to showcase how the SuperTriangle represents a synthesis across scales and mechanisms, how this underpins differences in concepts and methods, and how it ultimately defines how to answer fundamental questions and at what scale+type of modelling we need to focus on to advance networks into valuable predictive ecology.

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

At the heart of modern biodiversity science are a set of concepts and theories about biodiversity, stability and function. These relate to the abundance, distribution and services that biodiversity provides, and how biodiversity – as an interconnected set of species – responds to multiple stressors. The interaction between species (or individuals) is one of the fundamental building blocks of ecological communities. Collecting species interactions among many species (e.g. biodiversity) results in the formation of networks which are long established as a canon of ecology since the penning of the ‘entangled bank’ [1].

The use of species interaction networks provides a powerful abstraction that can help quantify, conceptualise, and understand biodiversity dynamics, and ultimately, one hopes, make prediction, mitigate change and manage services. Such network representations of biodiversity (including within species diversity) are increasingly argued to be an asset to predictive ecology, climate change mitigation and resource management. Here, it is argued that characterising biodiversity in a network will allow deeper capacity to understand and predict the abundance, distribution, dynamics and services provided by multiple species facing multiple stressors.

However, the way that a network is coded (constructed) defines an epistemology of the network concept which, we argue, can influence the resulting observations and conclusions about pattern and mechanisms that are made [2,3]. This process of coding/constructing networks has two major pillars: data and theory, the latter representing an expression of mechanism and process giving rise to patterns that emerge from collating interactions among species. Each of these carries with it a set of practical, semantic and conceptual constraints that not only influence progress in making network ecology more valuable and potentially predictive, but help define the spatial, temporal and evolutionary scale of assumptions we make and predictions we might generate from the networks.

With respect to data, it is extremely challenging to actually record species interactions in the field [4,5]. Despite notable herculean efforts (**Woodward? Benguela?**), actual coverage of ‘real world’ interaction data is sparse [6]. Against this practical challenge, there is additionally high variance in the terminology we use to define networks. Finally, the mathematical and statistical tools we use to construct, conceptualise, analyse and predict with these networks are also highly variable.

1. what are the underlying assumptions about nodes, edges, scale and process that are made when we attempt to delimit and describe a food webs;
2. are there families of commonly used tools that map onto assumptions about scales and processes;

The provision of this detail ultimately leads to a set of insights and conclusions about whether, when and under what conditions network representations of biodiversity can contribute to the advancement of ecological theory and generate value in predictive ecology. Specifically, we finish this perspective with an overview of fundamental questions in ecology that we think can benefit from network thinking and a proposal that such thinking can accelerate our capacity to predict the impact of multiple stressors on biodiverse communities.

# 1. Setting the Scene: The Not So Basics of Nodes and Edges

Important goal here is to introduce the idea that there are multiple facets as to what determines the interaction between species and that there is some sort of ‘scale of organisation’ [Figure 2](#fig-feasibility). We can then introduce these different scales/theories and I think some key points to highlights are the *features*, *limits*, and *descriptions* of these different scales (by that I mean what rules them, what finds them, and what binds them - sorry not sorry). I think it is also worth either in this section or in the one where we talk about model families to discuss the idea of ‘moving between’ different levels - e.g. downsampling but also feasibility - can we actually do that? ANother interesting discussion here (or maybe actually something that can make its way into the concluding remarks) is thinking about what determines interactions vs what determines structure…

|  |
| --- |
| Box 1 -The anatomy of a food web |
| 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 [7] 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 1](#fig-anatomy), 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 1: The many ways in which a food web can be defined and described at the node, edge, and even network level. |  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 [8] and in cases where the adult and larval stages of a species have different diets it may make ecological sense [9] 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 [10] or potential [11] feeding links, or energy transfer and material flow [12]. 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, 13] or continuous measurements [which quantifies the strength of of an interaction, 14]. 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 are they 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 [15]. 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 [2,3]. 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 2](#fig-feasibility)) 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. |

# 2. From Nodes and Edges to Scales and Processes

Armed with these basics, it is now possible to review the scales and assumptions that are made by a wide range of tools to assist in constructing networks against poor data with the hope of capturing important processes that underpin accurate prediction. Our thesis centres on a four-tier conceptualization of networks: co-occurrence, feasibility, mass effects and energetics. In the following sections we review each of these and then provide a synthesis among them.

## 2.1 Understanding the drivers of species interactions

Important goal here is to introduce the idea that there are multiple facets as to what determines the interaction between species and that there is some sort of ‘scale of organisation’ [Figure 2](#fig-feasibility). We can then introduce these different scales/theories and I think some key points to highlights are the features, limits, and descriptions of these different scales (by that I mean what rules them, what finds them, and what binds them - sorry not sorry). I think it is also worth either in this section or in the one where we talk about model families to discuss the idea of ‘moving between’ different levels - e.g. downsampling but also feasibility - can we actually do that? Another interesting discussion here (or maybe actually something that can make its way into the concluding remarks) is thinking about what determines interactions vs what determines structure…

* These different theories are shown in [Figure 2](#fig-feasibility) and we can see there is some element of scaling (species - population - individual)

|  |
| --- |
| Figure 2: TODO. |

**1. (Co)occurrence**

Although the outright assumption that because two species are co-occurring it must mean that they are interacting is inherently flawed [16], it is of course impossible for two species to interact (at least in terms of feeding links) if they are not co-occurring in time and space. Hence it is of course important to take into consideration the co-occurrence of both the resource and the consumer. An example of this would be the work from [17], where a metaweb (feasibility network) is downsampled into smaller realisations based on better data/knowledge as to which species are occurring at a specific location - however arguably these are still firmly in the space of feasible interactions for the specific location but are approaching a better approximation of ‘reality’…

**2. Feasibility**

This is based on the idea of forbidden links introduced by [5], specifically that there must be some degree of *trait complementarity* that allows a predator to chase, capture, kill, and consume, its prey. This is probably the level that the idea of a metaweb [11] is most applicable to. Within the network prediction ‘field’ this is perhaps the most developed space. Predictive models run the gamut including mechanistic models [18], binary classifiers [19], and graph embedding [20] and use either traits (or phylogeny as a proxy for the conservation thereof) as a means to ‘evaluate’ if an interaction is *possible* between two species (again not the likelihood of it happening but the likelihood of its feasibility). It is probably worth having a brief interlude here to be really clear that just because an interaction is probabilistic it does not make it weighted (at least not in the traditional sense of weighted interactions, *e.g.,* [21]) - it is still ‘binary’, it just happens to be defined by a binomial distribution (*sensu* [15]).

**3. Mass effect**

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

This is probably the point where we start to shift from a *potential* (presence/absence) way of defining interactions and start moving into the ‘qualitative’/weighted interaction space - we are not ‘determining’ if the interaction is feasible but rather making an assumption on prey selection based on the species’ likelihood of ‘meeting’, although [15] presents a compelling case that this could still be considered something that falls under the ‘feasibility’ and not ‘reality’ side of the spectrum… (well at least past Tanya seemed to think so)

**4. Energetics**

This is where we begin to move into the foraging ecology space - specifically consumption rate and how that pertains to energy acquisition *i.e.,* optimal foraging theory. In the loosest sense I think this is the ‘prey choice’ space - but specifically in the context of how prey choice as informed by energetic cost (not just purely based on *e.g.,* the most abundant species). If we think about ways that people have approached this there are the diet models of [22] and [23] as well as the ‘trait’ framework developed by [24] that moves the ‘energy’ into different ‘modules’ related to the process of the consumer acquiring energy from the resource (however there is a disregard for the ‘Rule 1’ requirement of forbidden links, again not bad just pointing it out). The idea of the consumer search space developed by [25] is also an interesting consideration. Finally the environment itself is also imposing energy costs on the predator. Basically the ideas presented in [26], which is essentially a take on movement ecology? What it boils down to is being able to quantify the cost of movement *i.e.,* the physical constraints that the environment imposes on a species… Maybe we can also think of it more in terms of metabolic rate?

# 3. Network prediction is scale dependent

The way in which we predict a network is driven by the underlying theory [Figure 2](#fig-feasibility) which constrains or informs the assumptions we make (this of course also has implications with regards to how the resulting network is defined (Box 1)). We can then spend a moment introducing the different model families [Table 1](#tbl-families). I think a clear messaging here might be that models can share a similar underlying theory but use different methods to get there (*e.g.,* using ecological rules (explicit), ecological expectations (pattern finders), or mathematical models (assumptions on the structure of the matrix - maybe even network)). Importantly different models will also have different ‘limits’ to them - this is probably a product of both where they are found within the ‘theory space’ [Figure 2](#fig-feasibility) as well as the definition of the network (Box 1) space. Should we also maybe revisit the idea of interaction vs structure predictors… I think it is still a point that is worth raising but no longer the framework on which we hang the different model families…

* The way in which we predict networks is ‘constrained’/informed by the different theories shown in [Figure 2](#fig-feasibility)
* Need to be aware of this and be aware how/what we can use the networks - Petchey dilemma
* The ‘scale’ that a network is constructed should be a determinant of what we can learn about a system *e.g.,* can’t use a feasibility network to learn something about energy flows. This is because they are capturing different processes
* Link the ‘model families’ to the different scales/theories
* Data…

As discussed in Box 1 there are many ways to define a food web, meaning that there are equally as many reasons one might be interested in predicting a food web. However we may think of two primary drivers for wanting to predict networks (Panel B [Figure 1](#fig-anatomy)), namely an interest in generating a set of ecologically plausible networks (*i.e.,* being able to describe networks using a model) or being able to recover (predict) location specific, ‘realised’, interactions for a specific species community (*i.e.,* being able to predict/infer the interactions between species). Of course these two categories are not distinct, mutually exclusive, groups but can rather be viewed as operating on a continuum ranging from a need for generality (*i.e.,* creating a network that, when taken in aggregate, the distribution of links (interactions) between nodes (species) are ecologically plausible) to a need for specificity (*i.e.,* local-level predictions between specific species pairs). Although the ability to predict ‘real-world’ interactions (and the resulting food webs) can have more intuitive ‘real world’ applications *e.g.,* being able to ‘recover’ food webs that have since gone extinct [27,28], using pairwise interactions to understand species distributions [29] or even co-extinction risk [30], a more structural approach to network construction affords one an opportunity to interrogate some of the more high-level mechanisms that are structuring networks (Box 1).

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

## 3.1 How do we predict food webs?

Selecting a model for the task of network prediction should come down to two things; what *aspect* of a food web one is interested in predicting, and what data are available, necessary, and sufficient. As shown in panel B of [Figure 1](#fig-anatomy) the interest in a network is (usually) at either the ‘structural’ or ‘interaction’ level and the development of models for the task of network prediction often focus on high fidelity (performance) at one of these scales. With this in mind it is beneficial to think of the different model families relative to these two different goals; here we refer to models that are used to predict the structure of a network as **topology generators** and models developed to infer the interactions for a given species pool as **interaction predictors**. It is meaningful to make this distinction because although it is possible to construct a food web given using an *interaction predictor* the models themselves lack any sort of parametrisation of the network structure and so the resulting network is a poor reflection of the actual network structure [31]. This is primarily because *interaction predictors* are models that evaluate the feasibility of an interaction between species pairs and not in the context of feasibility at the community level. Models themselves are a reflection of the different goals and intentions of the research program from which they are developed and are often ‘described’ by a specific mechanism that will determine the resulting structure or interactions (Box 1). Models such as the niche [8] or cascade [32] were developed with the intent of being used to understand the *structural* aspects of food webs, specifically how links are distributed amongst species in the community, whereas bayesian [33] or trait hierarchy [34] models have been developed on the basis that the traits of a species are the underlying mechanism in determining the feasibility of interactions (*i.e.,* species has the capacity to eat species ). Along with predicting different anatomical parts of a food web the different models have varying degrees of data that are needed to ‘parametrise’ the network. Once these two limitations are assessed and addressed it is then possible to select the model (or model family) that will best be able to capture food web feature that the researcher is most interested in (see Box 2 - Assessing model outputs). It is thus clear that (realistically) there will probably never be a ‘best fit’ tool that is able to construct a food web that will span the entire range of needs, and rather the responsibility lies with the researcher to be aware of not only the underlying philosophy of the specific toolset (as this could have knock-on effects when using those networks for downstream analyses/simulations; pers. comms. Beckerman, 2024), but also how well the tool can retrieve the specific network or interaction properties that is of interest.

In order for a model to formalise a ‘complete’ food web it is necessary to formalise two aspects of the network, ‘who eats whom’ (to determine the links between nodes) as well as the structure of the network (to limit the distribution of links), however most models are inclined to focus on one of the two aspects [panel B of 1](#fig-anatomy).

Crucially most topology generators lack some key data on the interaction between species (this can be because of how the model itself defines species or the way in which links are assigned in the network) and interaction predictors lack some sort of parametrisation of network structure [just because two species can interact it does not mean that they will, 35].

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

## 3.2 Categorizing Model Families

**NEED A TRANSITION PARAGRAPH**

As there are many food web models to choose from it is perhaps useful to think about the models in terms of model families, a summary of these families is presented in [Table 1](#tbl-families) highlights the differences and similarities of the philosophies and assumptions that determine a network. A more extensive overview of the different models that fall with in the different model families can be found in [SuppMat 1](https://beckslab.github.io/ms_t_is_for_topology/notebooks/model_descriptions-preview.html) and for a more detailed breakdown of the different ‘traits’ of the model families refer to [SuppMat 2](https://beckslab.github.io/ms_t_is_for_topology/notebooks/model_qualitative-preview.html).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: A summary of the different families of tools that can be used to generate food webs. Here ‘type’ leans on the ideas from [36] in terms of reconstructing networks from other known (observed) interactions and inferring interactions without relying on observed interactions.   | Model family | Assumption | Theory | Type | Key reference | | --- | --- | --- | --- | --- | | null | Links are randomly distributed within a network |  | network inference |  | | neutral | Network structure is random, but species abundance determines links between nodes | abundance | network inference | [37] | | resource | Networks are interval, species can be ordered on a ‘niche axis’ |  | network inference | [38] | | generative | Networks are determined by their structural features |  | network reconstruction |  | | energetic | Interactions are determined by energetic costs | abundance + energy | network reconstruction |  | | graph embedding | Interactions can be predicted from the latent traits of networks | feasibility | network reconstruction | [20] | | trait matching | Interactions can be inferred by a mechanistic framework/relationships | feasibility | network reconstruction | [18] | | binary classifiers | Interactions can be predicted by learning the relationship between interactions and ecologically relevant predictors | feasibility | network reconstruction | [19] | | expert knowledge | ‘Boots on the ground’ ecological knowledge and observations | feasibility | network reconstruction |  | | data scavenging | Webscraping to create networks from online databases |  | network reconstruction | [39] (f you squint?) | | co-occurrence | co-occurrence patterns arise from interactions so we can use these patterns to reverse engineer the interactions | co-occurrence | network inference |  | |

# 4. Making Progress with 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’etre*) 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)…

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

# 5. Concluding remarks

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

* In certain situations structure is ‘enough’ but there may be use cases where we are really interested in the node-level interactions *i.e.,* species identity is a thing we care about and need to be able to retrieve specific interactions at specific nodes correctly.
* Why do interaction models do so badly at predicting structure? Nuance of metaweb vs realisation but also time? At the core of it interaction models are trained on existing interaction data; this is data that are most likely closer to a metaweb than a local realisation even if they are being inventoried at a small scale…
  + We can briefly shoehorn downsampling here maybe??
* It will be interesting to bring up the idea that if a model is missing a specific pairwise link but doing well overall then when does it matter?
  + The fact that *some* people are concerned about the taxonomic resolution and cascading effects those might have on our understanding of network structure [10,40], but that puts us in a place where we are at risk of losing our ability to distinguish the wood from the tree - are we not (at least at times) concerned more with understanding ecosystem level processes than with needing to understand things *perfectly* at the species level.
  + I don’t think these ‘rare’/nuanced links (e.g. carnivorous hippos) are going to rock the boat when we think about networks at the structural level.

“The resolution of food-web data is demonic because it can radically change network topology and associated biological inferences in ways that are unknowable in the absence of better data.” - [40] The counter to this is that structural models are often not working at the species level and thus the structure remains ‘unchanged’ when you increase the resolution - I don’t think that people are that concerned with the structure of real world networks barring connectance and since that scales with species richness anyway your final proportion will probably still remain the same…

* I think a big take home will (hopefully) be how different approaches do better in different situations and so you as an end user need to take this into consideration and pick accordingly. I think [41] might have (and share) some thoughts on this. I feel like I need to look at [42] but maybe not exactly in this context but vaguely adjacent.
  + I think this is sort of the crux of the argument presented in [43] as well.

*“we highlight an interesting paradox: the models with the best performance measures are not necessarily the models with the closest reconstructed network structure.”* - [44]

* Do we need network models to predict interactions and interaction models to predict structure?
  + “Another argument for the joint prediction of networks and interactions is to reduce circularity and biases in the predictions. As an example, models like linear filtering generate probabilities of non-observed interactions existing, but do so based on measured network properties.” - [45]
  + Aligning (dove-tailing) with this the idea of ensemble modelling as presented by [46]
* Close out with a call to action that we have models that predict networks very well and models that predict interactions very well but nothing that is doing well at predicting both - this is where we should be focusing our attention when it comes to furthering model development…
* Do we expect there to be differences when thinking about unipartite vs bipartite networks? Is there underlying ecology/theory that would assume that different mechanisms (and thus models) are relevant in these two ‘systems’.
  + The [47] paper looks at some methods but is specifically looking at a bipartite world…

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

“That being said, there is a compelling argument for the need to ‘combine’ these smaller functional units with larger spatial networks [48] and that we should also start thinking about the interplay of time and space [49]. Although deciding exactly what measure might actually be driving differences between local networks and the regional metaweb might not be that simple [50].”

## 5.1 Time

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

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

# 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 *a priori* assumptions as that will constrain the network structure |
| 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 *all* 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 *per se* (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) |

# Outstanding questions

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

# References

1. Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, J. Murray

2. Proulx, S.R. *et al.* (2005) [Network thinking in ecology and evolution](https://doi.org/10.1016/j.tree.2005.04.004). *Trends in Ecology & Evolution* 20, 345–353

3. Brimacombe, C. *et al.* (2023) [Shortcomings of reusing species interaction networks created by different sets of researchers](https://doi.org/10.1371/journal.pbio.3002068). *PLOS Biology* 21, e3002068

4. Jordano, P. (2016) [Chasing Ecological Interactions](https://doi.org/10.1371/journal.pbio.1002559). *PLOS Biology* 14, e1002559

5. Jordano, P. (2016) Sampling networks of ecological interactions. *Functional Ecology* DOI: [10.1111/1365-2435.12763](https://doi.org/10.1111/1365-2435.12763)

6. Poisot, T. *et al.* (2021) [Global knowledge gaps in species interaction networks data](https://doi.org/10.1111/jbi.14127). *Journal of Biogeography* 48, 1552–1563

7. Poisot, T. *et al.* (2016) [Describe, understand and predict: Why do we need networks in ecology?](https://www.jstor.org/stable/48582345) *Functional Ecology* 30, 1878–1882

8. Williams, R.J. and Martinez, N.D. (2000) [Simple rules yield complex food webs](https://doi.org/10.1038/35004572). *Nature* 404, 180–183

9. Clegg, T. *et al.* (2018) [The impact of intraspecific variation on food web structure](https://doi.org/10.1002/ecy.2523). *Ecology* 99, 2712–2720

10. Pringle, R.M. (2020) [Untangling Food Webs](https://doi.org/10.1515/9780691195322-020). In *Unsolved Problems in Ecology*, pp. 225–238, Princeton University Press

11. Dunne, J.A. (2006) The Network Structure of Food Webs. In *Ecological networks: Linking structure and dynamics* (Dunne, J. A. and Pascual, M., eds), pp. 27–86, Oxford University Press

12. Lindeman, R.L. (1942) [The Trophic-Dynamic Aspect of Ecology](https://doi.org/10.2307/1930126). *Ecology* 23, 399–417

13. Poisot, T. *et al.* (2016) [The structure of probabilistic networks](https://doi.org/10). *Methods in Ecology and Evolution* 7, 303–312

14. Berlow, E.L. *et al.* (2004) [Interaction strengths in food webs: Issues and opportunities](https://doi.org/10.1111/j.0021-8790.2004.00833.x). *Journal of Animal Ecology* 73, 585–598

15. Banville, F. *et al.* (2024) [Deciphering probabilistic species interaction networks](https://doi.org/10.32942/X28G8Z)EcoEvoRxiv

16. Blanchet, F.G. *et al.* (2020) [Co-occurrence is not evidence of ecological interactions](https://doi.org/10.1111/ele.13525). *Ecology Letters* 23, 1050–1063

17. Dansereau, G. *et al.* (2023) Spatially explicit predictions of food web structure from regional level data

18. Morales-Castilla, I. *et al.* (2015) [Inferring biotic interactions from proxies](https://doi.org/10.1016/j.tree.2015.03.014). *Trends in Ecology & Evolution* 30, 347–356

19. Pichler, M. *et al.* (2020) [Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks](https://doi.org/10.1111/2041-210X.13329). *Methods in Ecology and Evolution* 11, 281–293

20. Strydom, T. *et al.* (2023) [Graph embedding and transfer learning can help predict potential species interaction networks despite data limitations](https://doi.org/10.1111/2041-210X.14228). *Methods in Ecology and Evolution* 14, 2917–2930

21. Wootton, J.T. and Emmerson, M. (2005) [Measurement of Interaction Strength in Nature](https://doi.org/10.1146/annurev.ecolsys.36.091704.175535). *Annual Review of Ecology, Evolution, and Systematics* 36, 419–444

22. Beckerman, A.P. *et al.* (2006) [Foraging biology predicts food web complexity](https://doi.org/10.1073/pnas.0603039103). *Proceedings of the National Academy of Sciences* 103, 13745–13749

23. Petchey, O.L. *et al.* (2008) [Size, foraging, and food web structure](https://doi.org/10.1073/pnas.0710672105). *Proceedings of the National Academy of Sciences* 105, 4191–4196

24. Wootton, K.L. *et al.* (2023) [Towards a modular theory of trophic interactions](https://doi.org/10.1111/1365-2435.13954). *Functional Ecology* 37, 26–43

25. Pawar, S. *et al.* (2012) [Dimensionality of consumer search space drives trophic interaction strengths](https://doi.org/10.1038/nature11131). *Nature* 486, 485–489

26. Cherif, M. *et al.* (2024) [The environment to the rescue: Can physics help predict predator–prey interactions?](https://doi.org/10.1111/brv.13105) *Biological Reviews* n/a

27. Dunne, J.A. *et al.* (2008) [Compilation and Network Analyses of Cambrian Food Webs](https://doi.org/10.1371/journal.pbio.0060102). *PLOS Biology* 6, e102

28. Yeakel, J.D. *et al.* (2014) [Collapse of an ecological network in Ancient Egypt](https://doi.org/10.1073/pnas.1408471111). *PNAS* 111, 14472–14477

29. Pollock, L.J. *et al.* (2014) [Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM)](https://doi.org/10.1111/2041-210X.12180). *Methods in Ecology and Evolution* 5, 397–406

30. Dunn, R.R. *et al.* (2009) [The sixth mass coextinction: Are most endangered species parasites and mutualists?](https://doi.org/10.1098/rspb.2009.0413) *Proceedings. Biological Sciences* 276, 3037–3045

31. Caron, D. *et al.* (2024) [Trait-matching models predict pairwise interactions across regions, not food web properties](https://doi.org/10.1111/geb.13807). *Global Ecology and Biogeography* 33, e13807

32. Cohen, J.E. *et al.* (1990) *Community Food Webs: Data and Theory*, Springer-Verlag

33. Cirtwill, A.R. *et al.* (2019) [A quantitative framework for investigating the reliability of empirical network construction](https://doi.org/10.1111/2041-210X.13180). *Methods in Ecology and Evolution* 10, 902–911

34. Shaw, J.O. *et al.* (2024) [A framework for reconstructing ancient food webs using functional trait data](https://doi.org/10.1101/2024.01.30.578036)bioRxiv, 2024.01.30.578036

35. Poisot, T. *et al.* (2015) [Beyond species: Why ecological interaction networks vary through space and time](https://doi.org/10.1111/oik.01719). *Oikos* 124, 243–251

36. Momal, R. *et al.* (2020) [Tree-based inference of species interaction networks from abundance data](https://doi.org/10.1111/2041-210X.13380). *Methods in Ecology and Evolution* 11, 621–632

37. Canard, E. *et al.* (2012) [Emergence of Structural Patterns in Neutral Trophic Networks](https://doi.org/10.1371/journal.pone.0038295). *PLOS ONE* 7, e38295

38. Williams, R.J. and Martinez, N.D. (2008) [Success and its limits among structural models of complex food webs](https://doi.org/10.1111/j.1365-2656.2008.01362.x). *Journal of Animal Ecology* 77, 512–519

39. Poisot, T. *et al.* (2016) [Synthetic datasets and community tools for the rapid testing of ecological hypotheses](https://doi.org/10.1111/ecog.01941). *Ecography* 39, 402–408

40. Pringle, R.M. and Hutchinson, M.C. (2020) [Resolving Food-Web Structure](https://doi.org/10.1146/annurev-ecolsys-110218-024908). *Annual Review of Ecology, Evolution and Systematics* 51, 55–80

41. Petchey, O.L. *et al.* (2011) [Fit, efficiency, and biology: Some thoughts on judging food web models](https://doi.org/10.1016/j.jtbi.2011.03.019). *Journal of Theoretical Biology* 279, 169–171

42. Berlow, E.L. *et al.* (2008) [The “Goldilocks factor” in food webs](https://doi.org/10.1073/pnas.0800967105). *Proceedings of the National Academy of Sciences* 105, 4079–4080

43. Brimacombe, C. *et al.* (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)

44. Poisot, T. (2023) [Guidelines for the prediction of species interactions through binary classification](https://doi.org/10.1111/2041-210X.14071). *Methods in Ecology and Evolution* 14, 1333–1345

45. Strydom, T. *et al.* (2021) [A roadmap towards predicting species interaction networks (across space and time)](https://doi.org/10.1098/rstb.2021.0063). *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, 20210063

46. Becker, D.J. *et al.* (2022) [Optimising predictive models to prioritise viral discovery in zoonotic reservoirs](https://doi.org/10.1016/S2666-5247(21)00245-7). *The Lancet Microbe* 3, e625–e637

47. Terry, J.C.D. and Lewis, O.T. (2020) [Finding missing links in interaction networks](https://doi.org/10.1002/ecy.3047). *Ecology* 101, e03047

48. Fortin, M.-J. *et al.* (2021) [Network ecology in dynamic landscapes](https://doi.org/10.1098/rspb.2020.1889). *Proceedings of the Royal Society B: Biological Sciences* 288, rspb.2020.1889, 20201889

49. Estay, S.A. *et al.* (2023) Editorial: Patterns and processes in ecological networks over space. *Frontiers in Ecology and Evolution* 11

50. Saravia, L.A. *et al.* (2022) [Ecological network assembly: How the regional metaweb influences local food webs](https://doi.org/10.1111/1365-2656.13652). *Journal of Animal Ecology* 91, 630–642