

# T is for Topology

Tanya Strydom <sup>1</sup>; Jennifer A. Dunne <sup>2</sup>; Timothée Poisot <sup>3,4</sup>; Andrew P. Beckerman <sup>1</sup>

**Abstract:** Although it has been acknowledged that communities consist not only of co-occurring species but that they also interact being able to quantify those interactions and assemble them into interaction networks has been a limiting factor in the integration of network ecology into other fields of ecology. As the field of network ecology has matured there has been an accompanying expansion in the development of theory and tools that are centred around generating networks or predicting the interactions between species. Notably many of these tools have been developed with different underlying philosophies, ideas, and mechanisms as to what structures the interactions between species. It is thus critically important that those wanting to adopt these network generating tools be aware of how the the specific questions being asked maps to the underlying assumptions made when generating networks, as well as the limitations of how the networks/interactions are delimited. Here we provide an overview of the canonical network generating models, comparing and contrasting the underlying assumptions, data requirements, and resulting network predictions made by the different families in an attempt to provide guidance for those interested in adopting the generation of networks into their workflow. [R1. a discussion on the underlying assumptions we are making when we delimit a network]. [R2. an overview of how the different model families differ - ordination space/benchmarking]. [R3. identifying the relevant questions/bodies of theory that the networks generated by different families are suited to answer]. When choosing to construct an interaction network the researcher is faced with many assumptions and considerations that should be made and it is important to be aware of these limitations to avoid constructing (something poetic to capture the idea of falsity/false idols). Being aware of these choices is particularly important as the availability of these tools grows and network ecology starts to be adopted into other aspects of ecology and conservation biology.date: last-modified

**Keywords:** food web, network construction

1 It can be argued that the interaction between species (or individuals) is one of the main determinants of the  
2 emergent properties that are studied in other fields of ecology, *e.g.*, the range of plant will be determined by  
3 the range of its pollinator, and although the importance of species interactions and the resulting networks  
4 that they form has been an acknowledged part of the ecological canon since the penning of the ‘entangled  
5 bank’ (Darwin, 1859) (if not even earlier, stemming from Greek Antiquity (Thanos, 1994)), the adoption of  
6 network ecology into other disciplines of ecology has been limited. This has primarily been driven by two  
7 limitations; firstly, it is extremely challenging to actually record species interactions in the field (Jordano,  
8 2016a, 2016b), which has resulted in a limited coverage of ‘real world’ interaction data (Poisot et al., 2021),  
9 and secondly has been the need to develop terminology and tools that help us to construct, conceptualise,  
10 and analyse these networks. Although recording interactions in the field remains a challenge, although the  
11 development of both practical tools (*i.e.*, tools that help us to record or measure interactions, (e.g. ref maybe  
12 Pringle & Hutchinson, 2020)), as well as discussions around the development of tools to predict or infer  
13 them (Morales-Castilla et al., 2015; Strydom et al., 2021), has allowed us to begin filling in these ‘global  
14 gaps’, albeit in a more synthetic manner (Poisot, Gravel, et al., 2016). Additionally, there has been extensive  
15 development in in the ways in which we formalise networks (Dale & Fortin, 2010), and the tools and language  
16 that we use to quantify the structure and properties of networks (Delmas et al., 2019). All together these  
17 tools mean that, as a field, network ecology can (and should) be integrated into the broader fields of ecology  
18 (*e.g.*, Thuiller et al., 2024) and conservation biology. However (as with any new tool or model), it is important  
19 that one has a firm grasp of how networks (particularly synthetic ones) are generated and how the underlying  
20 philosophy thereof maps onto the questions being asked. Here we provide; a discussion of the underlying  
21 assumptions that are made when we attempt to delimit and describe a food webs, a synthesis of the different  
22 families of tools that are used to construct food webs, and a discussion linking network ecology to some of  
23 the outstanding questions in ecology.

24 Three themes that are aimed at: providing a standardisation of terms that are used when we are  
25 talking about both networks as well as what we mean when we are generating networks. The  
26 final theme aims to map network ecology to some of the outstanding questions in ecology

## 27 **0.1 The anatomy of a food web**

28 Although we specifically focus on food webs (interactions representing feeding links) it is beneficial to take  
29 a step back and acknowledge the diversity of form that an interaction network can encapsulate. The idea  
30 of an interaction network seems simple, it is the representation of the interactions (edges) between species  
31 (nodes), the definition of an ‘edge’ and a ‘node’, 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 between individuals), community (the links between species), or metacommunity (fluxes between locations) level. Even if we are to limit our definition of a network to represent community-level processes there are still many ways to define what is captured by the edges and nodes [insert some e.g.]. It is thus clear that the way that a network is coded (constructed) can influence the resulting observations and conclusions that are made (Brimacombe et al., 2023; Proulx et al., 2005), and it is important to have a strong grasp of what information a network is attempting to convey.

Even if one were to limit their scope to thinking of interaction networks only in terms of food webs there are still many ways to define the various components of the network, 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?, what is the spatial extent?). All these decisions will have an impact on the resultant structure and potential use-cases of the network.

### 0.1.1 How do we define a node?

Although this may seem an elementary question in the context of food webs — a node should represent a species, the reality is that nodes can often represent an aggregate of different (taxonomic) species - so called ‘trophic species’, and it is not uncommon that networks can have nodes that represent both taxonomic and trophic species (*e.g.*, there are many that do the basal ‘plant/phytoplankton’ node but include at least one REF). 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 [mutualism ref, at least there should be one of them], 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.

### 0.1.2 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) feeding links, potential (Dunne, 2006) feeding links, or energy flows[??] (Petchey et al., 2008). How we quantify 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 (*i.e.*, a ‘present’ interaction is one implies that species  $a$  has the ability to consume species  $b$  but it does not mean that this interaction is realised in the field) will be meaningless if you are interested in understanding the flow of energy through the system as the links within the network are over connected. 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 effect of one species on another, Berlow et al., 2004). Although there is a clear argument for moving away from a purely binary way of representing interactions [probabilities preprint] this of course also means that there is an additional layer to the interpretation these links.

### 0.1.3 Putting the parts together; what does it mean?

It is clear that there are many ways to define, code, and construct food webs, however what may be less clear is understanding *why* there is such a diversity of thought. Here it may be meaningful to contextualise the different ‘types’ of food webs within the larger questions (or needs) that have been driving them. Some of the earliest work on food webs was linked to the idea of niche space, and more specifically, the idea of trophic niches and how this would influence the dimensionality of a networks (Cohen, 1977). This introduced the idea that a single dimension (the “niche axis,” Allesina et al., 2008) constrains the interactions between species; in this instance it makes sense to think of species in terms of what they consume and what they are consumed by, as they are occupying the same space in the niche axis. Networks that are defined in this way may be useful for understanding how the flow of energy (resources) are constrained between ‘species’, particularly how it moves through the trophic levels. This ‘niche-based’ way of thinking might be beneficial when thinking about networks at the structural level, and when trying to map large-scale processes [ref?] however there was also a need to develop ways of thinking that were more geared to thinking about why does species  $a$  predate species  $b$ , broadly this is the result of two things; a predator needs to have the correct traits to be able to capture, kill, and consume, its prey (a mismatch between predator and prey is termed a forbidden link, Jordano (2016b)) and it needs to be energetically feasible [feeding ecology ref]. When we think of interactions in these terms it makes sense that nodes are defined at the species level (or at least as species that have the same traits and/or energy content), however the links between them can be quantified in different ways... [this is lazy writing]

*something, something, introducing that the same problem (different philosophies) is also a thing that you need to think about when aggregating interactions/generating networks.*

## 0.2 Constructing ecological networks

maybe a more direct link here to the fact that when working with networks its often synthetic ones *i.e.*, the product of some sort of modelling exercise; alternatively there has also been a push to develop predictive tools to create hypothetical (but plausible) networks for real world situations. Also talk about even deciding to create a network from field observations is in and of itself still a ‘model’ that has assumptions... for example decisions are made about delimiting, aggregation, and observation, the idea of aggregating over time or aggregating over space. Same can e said for different food web generating tools , they have their own underlying rules and assumptions that are made when constructing a food web, which will determine and influence the resulting structure or inferred interactions (Petchey et al., 2008)

Arguably the need for methods and tools that can be used to construct synthetic food webs arises from two different (but still aligned) places of interest within the field of network ecology. On the one side sits the researcher who is interested in generating a set of ecologically plausible networks for the purpose of understanding some higher-level process/concept (*e.g.*, understanding energy flows) in a more synthetic setting, whereby these networks do not require any level of species specificity *per se* and it is more the arrangement of the nodes and links within the context of network structure that is of value. This researcher is contrasted by one that is interested in constructing real-world, location specific, interaction data for a specific collection of species (community). This is driven by the need for researchers to find alternative ways to infer the interactions between species as a way to overcome the inherent challenges of inventorying interactions in the field (see Morales-Castilla et al. (2015) for a more mechanistic, and Strydom et al. (2021) for a more statistical overview of ways to approach this specific issue). 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). 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 on the researcher to be aware of not only the underlying philosophy of the specific toolset they wish to use used also how well the tool is able to retrieve the network properties that they desire.

### 0.2.1 Model families

As there are many food web generating tools to choose from it is perhaps useful to think about these tools in terms of families, where families represent tools that have a similar methodology and (more importantly)

have the same underlying philosophies and assumptions that determine the links between nodes as well as how these may be encoded, a summary of these model families are presented in Table 1. Although there have been efforts to compare and contrast different models (*e.g.*, Williams & Martinez, 2008 looked at ‘structural models’; and Pichler et al., 2020 looked at machine learning algorithms) there still lacks an overall synthesis as to how the different model families differ from each other - both in terms of what they are actually predicting as well as how well they are performing in the different facets of constructing a food web.

[Figure 1 about here.]

[Figure 2 about here.]

**Null models:** The interactions between species occurs regardless of the identity of the species (*i.e.*, species have no agency) and links are randomly distributed throughout the network. There is however the assumption that a network will be constrained by the number of links. Type I (Fortuna & Bascompte, 2006), where interactions happen proportionally to connectance and Type II (Bascompte et al., 2003), where interactions happen proportionally to the joint degree of the two species involved. These two models are equivalent to the Erdos-Renyi and Configuration models (Newman, 2010) respectively (check that though).

**Neutral models:** Based on the theory that interactions occur as the result of the abundance of species (*i.e.*, the species still has no agency but its abundance does?). See Pomeranz et al. (2019)

**Resource models:** Based on the idea that networks follow a trophic hierarchy and that species interactions can be determined using a single dimension [the “niche axis”; Allesina et al. (2008)]. Essentially these models can be viewed as being based on the idea of resource partitioning (niches) along a one-dimensional resource and that the number of links scale with species richness (linear link scaling). That is, there is some sort of hierarchical feeding based on how a ‘resource’ is partitioned. Broadly this family consists of three core models; the cascade model (Cohen et al., 1990), which rests on the idea that species feed on one another in a hierarchical manner; the niche model (Williams & Martinez, 2000), broadly all species are randomly assigned a ‘feeding niche’ and all species that fall in this niche can be consumed by that species; and the nested hierarchy model (Cattin et al., 2004), which adds some component of phylogenetic clustering/signal... so not a single dimension? **TODO.** Williams & Martinez (2008) provides a broader overview of some of the variations in these models as well as comparison between them regarding their ability to retrieve elements of networks structure (see also Allesina et al. (2008)).

**Generative models:** (this is maybe a bit of a bold term to use). MaxEnt (Banville et al., 2023), (maybe) stochastic block (Xie et al., 2017).

**Feeding models:** Broadly this family of models is rooted in feeding theory and allocates the links between

species based on energetics, which predicts the diet of a consumer based on energy intake. This means that the model is focused on predicting not only the number of links in a network but also the arrangement of these links based on the diet breadth of a species. The diet breadth model (Beckerman et al., 2006) as well as its allometrically scaled cousin the allometric diet breadth model (ADBM) (Petchey et al., 2008) determine links between species based on the energetic content, handling time, and density of species. See also DeAngelis et al. (1975)

Gravel et al. (2013) also poses an interesting cross-over between the adbm and niche model.

**Binary classifiers:** The task of predicting if an interaction will occur between a species pair is treated as a statistical binary classification task, where the task is to correlate ‘real world’ interaction data with a suitable ecological proxy for which data is more widely available (*e.g.*, traits). Model families often used include generalised linear models (*e.g.*, Caron et al., 2022), random forest (*e.g.*, Llewelyn et al., 2023), trait-based k-NN (*e.g.*, Desjardins-Proulx et al., 2017), and Bayesian models (Cirtwill et al., 2019; *e.g.*, Eklöf et al., 2013). See Pichler et al. (2020) for a more detailed overview on the performance of machine learning and statistical approaches for inferring trait-trait relationships.

**Graph embedding:** This family of approaches has been extensively discussed in Strydom et al. (2023) but can be broadly explained as an approach that estimates latent features from observed networks that can be used to predict interactions. Strydom et al. (2022) uses a transfer learning framework (specifically using a random dot product graph for embedding) based around the idea that interactions are evolutionarily conserved and that we can use known networks, and phylogenetic relationships, to predict interactions for a given species pool. **TODO** Log-ratio (Rohr et al., 2010)

**Trait matching:** Interactions are determined by a series of ‘feeding rules’, whereby the interaction between a species pair will only occur if all feeding rules are met. These rules are determined on an *a priori* basis using expert/ecological knowledge to determine the underlying feeding hierarchy using ecological proxies (Morales-Castilla et al., 2015). For example the Paleo Foodweb Inference Model (PFIM, Shaw et al., 2024) uses a series of rules for a set of trait categories (such as habitat and body size) to determine if an interaction can occur between a species pair. What sets this family of models apart from **expert knowledge** ones is that there is a formalisation of the feeding rules and thus there is some ability to transfer these rules to different communities.

**Expert knowledge:** Not so much about empirical observations but more the value of ‘local’ knowledge and having specific individuals sitting around a table and assigning a value of how confident they are that a specific species pair are likely to interact (*e.g.*, Dunne et al., 2008), this has the added advantage that

interactions can be scored in a more categorical as opposed to binary fashion, *e.g.*, Maiorano et al. (2020) score interactions as either obligate (typical food resources) or occasional (opportunistic feeding) interactions. I feel like its worth also mentioning downfalls *a la* Brimacombe et al. (2023)...

**Data scavenging:** There are also a lot of published *interaction* data that are publicly available *e.g.*, the Global Biotic Interactions (GloBI) database (Poelen et al., 2014) and these can also be used to construct an interaction network by mining these sources to look for interactions for specific species pairs. This is done by matching species pairs against those within a dataset of trophic interactions to determine if an interaction is present or absent between the two species (*e.g.*, the WebBuilder tool developed by Gray et al., 2015). It is important to note that this methodology is only going to be able to infer observations that have been recorded in the field, and given the relative scarcity [*I say Poisot et al. (2021) but that's more an overview of complete networks but one can also get pairwise interactions from these types of data so I feel like its okay?*] and localised sampling of these types of datasets it is very likely that there will be many false negatives (missing pairwise interactions) using this approach.

**Co-occurrence:** Trying to infer interactions from the co-occurrence patterns of species pairs within the community *e.g.*, the geographical lasso (Ohlmann et al., 2018). This (for me) seems fundamentally flawed and Blanchet et al. (2020) seems to agree with me at least a little bit.

Table 1: Lets make a table that gives an overview of the different model families and some of their features.

Model		Network	Nodes	Links	
family	Theory	predicted	represent	represent	Interaction
null	Network structure is random	structure	agnostic	feeding links	binary
neutral	Network structure is random, but species abundance plays a role	structure	species	feeding links	binary
resource	Networks are interval, species can be ordered on a 'niche axis'	structure	trophic species	subdivision of resource	binary
generative	Networks are determined by their structural features	structure	agnostic	links	binary



Model family	Theory	Network predicted	Nodes represent	Links represent	Interaction
energetic	Interactions are determined by foraging theory (feeding links)	interactions	species	feeding links	quantitative
graph embedding	Interactions can be predicted from the latent traits of networks	interactions	species	potential feeding links	probabilistic
trait matching	Interactions can be inferred by a mechanistic framework/relationships	interactions	species	feeding links	binary
binary classifiers	Interactions can be predicted by learning the relationship between interactions and ecologically relevant predictors	interactions	species	feeding links	binary
expert knowledge	‘Boots on the ground’ ecological knowledge and observations	interactions	species	feeding links	binary
data scavenging	Webscraping to create networks from online databases	interactions	species	feeding links	binary
co-occurrence	co-occurrence patterns arise from interactions so we can use these patterns to reverse engineer the interactions	co-occurrence patterns	species	association links	binary

#### 17 0.2.1.1 When to use what?

1 [Figure 3 about here.]

2 [Figure 4 about here.]

[Figure 5 about here.]

Source: [Article Notebook](#)

[Figure 6 about here.]

Source: [Article Notebook](#)

#### 0.2.1.2 Model benchmarking

- ‘Testing’ the performance of a model is going to depend on some of the core limitations of the model itself thus it makes sense to think of two sets benchmarking rules for network and interaction prediction models respectively (see Figure 1 panel B).

- When it comes to network models we are concerned with the ‘preservation’ of structure and distribution of links across the network. For interaction models we want to ensure that we are able to retrieve interactions that really exist but also those that cannot exist (*sensu* forbidden links Jordano (2016b))

“As long as these predictions are not perfect, some interactions will be predicted at the ‘wrong’ position in the network; these measures cannot describe the structural effect of these mistakes. On the other hand, measures of network structure can have the same value with interactions that fall at drastically different positions; this is in part because a lot of these measures covary with connectance, and in part because as long as these values are not 0 or their respective maximum, there is a large number of network configurations that can have the same value.” - Poisot (2023)

##### 0.2.1.2.1 Benchmarking network models

- Maybe look at some of the historic papers that compare some of the ‘resource models’
- See also Allesina et al. (2008) and the likelihood function that they use for model selection
- Look at Vermaat et al. (2009)

“Possibly, the most striking caveat of the use of summary statistics is that it cannot tell us whether or not a model is able to fully replicate empirical networks.” - Allesina et al. (2008)

##### 0.2.1.2.2 Benchmarking interaction models

- Main concern with predicting interactions is that we want to test the ‘quality’ of the links we are predicting (both true positives and true negatives), but the inherent sparsity (meaning high class imbalance) means that we also need to look at the balance of these predictions.

- “Both precision and recall may be useful in cases where there is imbalanced data. However, it may be valuable to prioritize one over the other in cases where the outcome of a false positive or false negative is costly.”
- Caveat regarding the use of real world interaction data both for training and validating predictions? *e.g.*, Poisot, Ouellet, et al. et al 2021 and Catchen et al 2023
- See Poisot (2023)
  - skill (ability to make the right prediction; evaluate whether low prevalence can lull us into a false sense of predictive accuracy)
  - bias (trends towards systematically over-predicting one class)
  - class imbalance (the relative number of cases representing interactions)
- “These results suggest that learning from a dataset with very low connectance can be a different task than for more connected networks: it becomes increasingly important to capture the mechanisms that make an interaction exist, and therefore having a slightly more biased training dataset might be beneficial. As connectance increases, the need for biased training sets is less prominent, as learning the rules for which interactions do not exist starts gaining importance”
- Maybe also looking at how well a model can recover ‘missing links’ *i.e.*, false negatives *sensu* what we did in Strydom et al. (2022)
- Need to discuss the key differences and implications between predicting a metaweb (*sensu* Dunne (2006)) and a network realisation. Maybe also Poisot et al. (2015) that discuss how the local factors are going to play a role.

[Figure 7 about here.]

Source: [Article Notebook](#)

I really like this way of plotting results from Pichler et al. (2020)

[Figure 8 about here.]

### 0.3 Link network ecology to the outstanding questions in ecology

- Bring up the fact that delimiting a network is in and of itself fuzzy - we tend to think of them in terms of snapshots but in reality the final (empirical) network is often the result of aggregation over multiple timescales.

- Also the fact that *some* people are concerned about the taxonomic resolution and cascading effects those might have on our understanding of network structure (Pringle, 2020; Pringle & Hutchinson, 2020), 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. To say this in a different way maybe it comes down to thinking about the scale of organisation within a network... The classical levels of organisation within ecology (population, community, ...) are also relevant when we think about a networks.

- Brief history of the development of tools within the context of the two different fields? Sort of where the theory/body of work was based and how that has changed?
- 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.
- 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).

Cohen et al. (1985) states that “[*Their*] approach is more like gross anatomy than like physiology... that is, the gross anatomy is frozen, rather than in motion.”.

Interestingly Williams & Martinez (2008) also explicitly talk about *structural* food-web models in their introduction... so how I see it that means that there has always been this inherent acknowledgement that models are functioning at a specific 'network 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.” - Pringle & Hutchinson (2020) 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...

“It makes no sense to describe the interaction structure of nodes which in themselves are poorly defined.” — Roslin et al. (2013, p. 2)

## 0.4 Discussion

- 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 Petchey et al. (2011) might have (and share) some thoughts on this (thanks Andrew). I feel like I need to look at Berlow et al. (2008) but maybe not exactly in this context but vaguely adjacent.
  - An interesting thing to also think about (and arguably it will be addressed based on some of the other thoughts and ideas) is data dependant and data independent ‘parametrisation’ of the models...
  - 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.
    - I think this is sort of the crux of the argument presented in Brimacombe et al. (2024)
- “we highlight an interesting paradox: the models with the best performance measures are not necessarily the models with the closest reconstructed network structure.” - Poisot (2023)*
- *Do we need network models to predict interactions and interaction models to predict structure?* (lets not think about that too hard or I might just have to sit in silence for a while...)
    - “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.” - Strydom et al. (2021)
    - Aligning (dove-tailing) with this the idea of ensemble modelling as presented by Becker et al. (2022)
  - It will be interesting to bring up the idea that if a model is missing a specific pairwise link but doing well at the structural level then when does it matter?
  - 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. (we need models that will fill the space in the top right quadrant of panel A in Figure 1)

## References

Source: [Article Notebook](#)

- Allesina, S., Alonso, D., & Pascual, M. (2008). A General Model for Food Web Structure. *Science*, 320(5876), 658–661. <https://doi.org/10.1126/science.1156269>
- Banville, F., Gravel, D., & Poisot, T. (2023). What constrains food webs? A maximum entropy framework for predicting their structure with minimal biases. *PLOS Computational Biology*, 19(9), e1011458. <https://doi.org/10.1371/journal.pcbi.1011458>
- Bascompte, J., Jordano, P., Melian, C. J., & Olesen, J. M. (2003). The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences*, 100(16), 9383–9387. <https://doi.org/10.1073/pnas.1633576100>
- 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)
- 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>
- 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>
- 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>
- 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>
- 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>
- 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>

Caron, D., Maiorano, L., Thuiller, W., & Pollock, L. J. (2022). Addressing the Eltonian shortfall with trait-based interaction models. *Ecology Letters*, 25(4), 889–899. <https://doi.org/10.1111/ele.13966>

Cattin, M.-F., Bersier, L.-F., Banašek-Richter, C., Baltensperger, R., & Gabriel, J.-P. (2004). Phylogenetic constraints and adaptation explain food-web structure. *Nature*, 427(6977), 835–839. <https://doi.org/10.1038/nature02327>

Cirtwill, A. R., Eklf, A., Roslin, T., Wootton, K., & Gravel, D. (2019). A quantitative framework for investigating the reliability of empirical network construction. *Methods in Ecology and Evolution*, 0(ja). <https://doi.org/10.1111/2041-210X.13180>

Cohen, J. E. (1977). Food webs and the dimensionality of trophic niche space. *Proceedings of the National Academy of Sciences*, 74(10), 4533–4536. <https://doi.org/10.1073/pnas.74.10.4533>

Cohen, J. E., Briand, F., & Newman, C. (1990). *Community Food Webs: Data and Theory*. Springer-Verlag.

Cohen, J. E., Newman, C. M., & Steele, J. H. (1985). A stochastic theory of community food webs I. Models and aggregated data. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 224(1237), 421–448. <https://doi.org/10.1098/rspb.1985.0042>

Dale, M. R. T., & Fortin, M.-J. (2010). From Graphs to Spatial Graphs. *Annual Review of Ecology, Evolution, and Systematics*, 41, 21–38. <https://www.jstor.org/stable/27896212>

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.

DeAngelis, D. L., Goldstein, R. A., & O'Neill, R. V. (1975). A Model for Tropic Interaction. *Ecology*, 56(4), 881–892. <https://doi.org/10.2307/1936298>

Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Riva, G. V. D., Fortin, M.-J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94(1), 16–36. <https://doi.org/10.1111/brv.12433>

Desjardins-Proulx, P., Laigle, I., Poisot, T., & Gravel, D. (2017). Ecological interactions and the Netflix problem. *PeerJ*, 5, e3644. <https://doi.org/10.7717/peerj.3644>

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

Dunne, J. A., Williams, R. J., Martinez, N. D., Wood, R. A., & Erwin, D. H. (2008). Compilation and Network Analyses of Cambrian Food Webs. *PLOS Biology*, 6(4), e102. <https://doi.org/10.1371/journal.pbio.0060102>

Eklöf, A., Tang, S., & Allesina, S. (2013). Secondary extinctions in food webs: A Bayesian network approach.

*Methods in Ecology and Evolution*, 4(8), 760–770. <https://doi.org/10.1111/2041-210X.12062>

Fortuna, M. A., & Bascompte, J. (2006). Habitat loss and the structure of plant-animal mutualistic networks: Mutualistic networks and habitat loss. *Ecology Letters*, 9(3), 281–286. <https://doi.org/10.1111/j.1461-0248.2005.00868.x>

Gravel, D., Poisot, T., Albouy, C., Velez, L., & Mouillot, D. (2013). Inferring food web structure from predator–prey body size relationships. *Methods in Ecology and Evolution*, 4(11), 1083–1090. <https://doi.org/10.1111/2041-210X.12103>

Gray, C., Figueroa, D. H., Hudson, L. N., Ma, A., Perkins, D., & Woodward, G. (2015). Joining the dots: An automated method for constructing food webs from compendia of published interactions. *Food Webs*, 5, 11–20. <https://doi.org/10.1016/j.fooweb.2015.09.001>

Jordano, P. (2016a). Chasing Ecological Interactions. *PLOS Biology*, 14(9), e1002559. <https://doi.org/10.1371/journal.pbio.1002559>

Jordano, P. (2016b). Sampling networks of ecological interactions. *Functional Ecology*. <https://doi.org/10.1111/1365-2435.12763>

Llewelyn, J., Strona, G., Dickman, C. R., Greenville, A. C., Wardle, G. M., Lee, M. S. Y., Doherty, S., Shabani, F., Saltr , F., & Bradshaw, C. J. A. (2023). Predicting predator–prey interactions in terrestrial endotherms using random forest. *Ecography*, 2023(9), e06619. <https://doi.org/10.1111/ecog.06619>

Maiorano, L., Montemaggiore, A., Ficetola, G. F., O’Connor, L., & Thuiller, W. (2020). TETRA-EU 1.0: A species-level trophic metaweb of European tetrapods. *Global Ecology and Biogeography*, 29(9), 1452–1457. <https://doi.org/10.1111/geb.13138>

Morales-Castilla, I., Matias, M. G., Gravel, D., & Ara jo, M. B. (2015). Inferring biotic interactions from proxies. *Trends in Ecology & Evolution*, 30(6), 347–356. <https://doi.org/10.1016/j.tree.2015.03.014>

Newman, M. E. J. (2010). *Networks. An introduction*. Oxford University Press.

Ohlmann, M., Mazel, F., Chalm ndrier, L., Bec, S., Coissac, E., Gielly, L., Pansu, J., Schilling, V., Taberlet, P., Zinger, L., Chave, J., & Thuiller, W. (2018). Mapping the imprint of biotic interactions on  $\beta$ -diversity. *Ecology Letters*, 21(11), 1660–1669. <https://doi.org/10.1111/ele.13143>

Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2008). Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences*, 105(11), 4191–4196. <https://doi.org/10.1073/pnas.0710672105>

Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2011). Fit, efficiency, and biology: Some thoughts on judging food web models. *Journal of Theoretical Biology*, 279(1), 169–171. <https://doi.org/10.1016/j.jtbi.2011.03.019>

Pichler, M., Boreux, V., Klein, A.-M., Schleuning, M., & Hartig, F. (2020). Machine learning algorithms



- to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and Evolution*, 11(2), 281–293. <https://doi.org/10.1111/2041-210X.13329>
- Poelen, J. H., Simons, J. D., & Mungall, C. J. (2014). Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics*, 24, 148–159. <https://doi.org/10.1016/j.ecoinf.2014.08.005>
- Poisot, T. (2023). Guidelines for the prediction of species interactions through binary classification. *Methods in Ecology and Evolution*, 14(5), 1333–1345. <https://doi.org/10.1111/2041-210X.14071>
- Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., Mercier, B., Violet, C., & Vissault, S. (2021). Global knowledge gaps in species interaction networks data. *Journal of Biogeography*, n/a(n/a). <https://doi.org/10.1111/jbi.14127>
- Poisot, T., Cirtwill, A., Cazelles, K., Gravel, D., Fortin, M.-J., & Stouffer, D. (2016). The structure of probabilistic networks. *Methods in Ecology and Evolution*, 7(3), 303–312. <https://doi.org/10.1111/2041-210X.12941>
- Poisot, T., Gravel, D., Leroux, S., Wood, S. A., Fortin, M.-J., Baiser, B., Cirtwill, A. R., Araújo, M. B., & Stouffer, D. B. (2016). Synthetic datasets and community tools for the rapid testing of ecological hypotheses. *Ecography*, 39(4), 402–408. <https://doi.org/10.1111/ecog.01941>
- Poisot, T., Stouffer, D. B., & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary through space and time. *Oikos*, 124(3), 243–251. <https://doi.org/10.1111/oik.01719>
- Poisot, T., Stouffer, D. B., & Kéfi, S. (2016). Describe, understand and predict: Why do we need networks in ecology? *Functional Ecology*, 30(12), 1878–1882. <https://www.jstor.org/stable/48582345>
- Pomeranz, J. P. F., Thompson, R. M., Poisot, T., & Harding, J. S. (2019). Inferring predator–prey interactions in food webs. *Methods in Ecology and Evolution*, 10(3), 356–367. <https://doi.org/10.1111/2041-210X.13125>
- Pringle, R. M. (2020). Untangling Food Webs. In *Untangling Food Webs* (pp. 225–238). Princeton University Press. <https://doi.org/10.1515/9780691195322-020>
- 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>
- 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>
- Rohr, R. P., Scherer, H., Kehrli, P., Mazza, C., & Bersier, L.-F. (2010). Modeling Food Webs: Exploring Unexplained Structure Using Latent Traits. *The American Naturalist*, 176(2), 170–177. <https://doi.org/10.1086/653667>
- 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>

Stouffer, D. B., Camacho, J., Jiang, W., & Nunes Amaral, L. A. (2007). Evidence for the existence of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society B: Biological Sciences*, 274(1621), 1931–1940. <https://doi.org/10.1098/rspb.2007.0571>

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>

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>

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>

Thanos, C. A. (1994). Aristotle and Theophrastus on plant-animal interactions. In M. Arianoutsou & R. H. Groves (Eds.), *Plant-animal interactions in Mediterranean-type ecosystems* (pp. 3–11). Springer Netherlands. [https://doi.org/10.1007/978-94-011-0908-6\\_1](https://doi.org/10.1007/978-94-011-0908-6_1)

Thuiller, W., Calderón-Sanou, I., Chalmandrier, L., Gaüzère, P., O'Connor, L. M. J., Ohlmann, M., Poggiato, G., & Münkemüller, T. (2024). Navigating the integration of biotic interactions in biogeography. *Journal of Biogeography*, 51(4), 550–559. <https://doi.org/10.1111/jbi.14734>

Vermaat, J. E., Dunne, J. A., & Gilbert, A. J. (2009). Major dimensions in food-web structure properties. *Ecology*, 90(1), 278–282. <https://www.ncbi.nlm.nih.gov/pubmed/19294932>

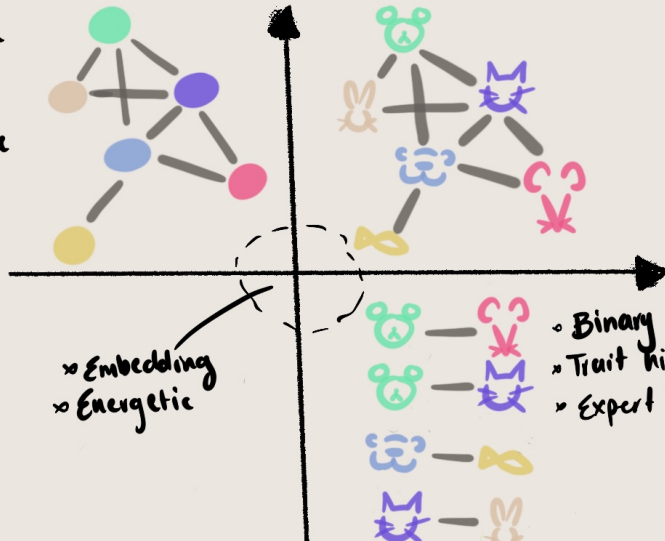
Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774), 180–183. <https://doi.org/10.1038/35004572>

Williams, R. J., & Martinez, N. D. (2008). Success and its limits among structural models of complex food webs. *Journal of Animal Ecology*, 77(3), 512–519. <https://doi.org/10.1111/j.1365-2656.2008.01362.x>

Xie, J.-R., Zhang, P., Zhang, H.-F., & Wang, B.-H. (2017). Completeness of Community Structure in Networks. *Scientific Reports*, 7(1), 5269. <https://doi.org/10.1038/s41598-017-05585-6>

# A Predicting.

- Resource
- Null
- Neutral
- Generative



- Embedding
- Energetic

- Binary Classifiers
- Trait hierarchy
- Expert knowledge

# B Benchmarking

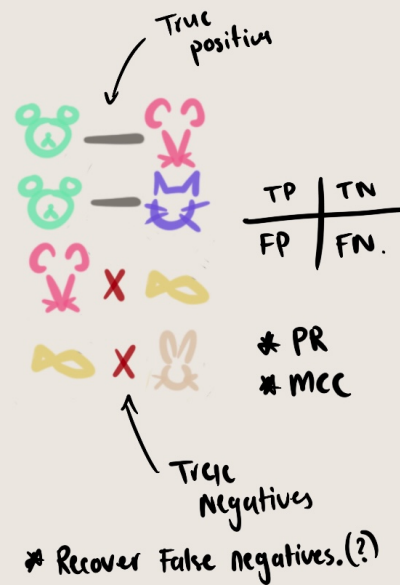
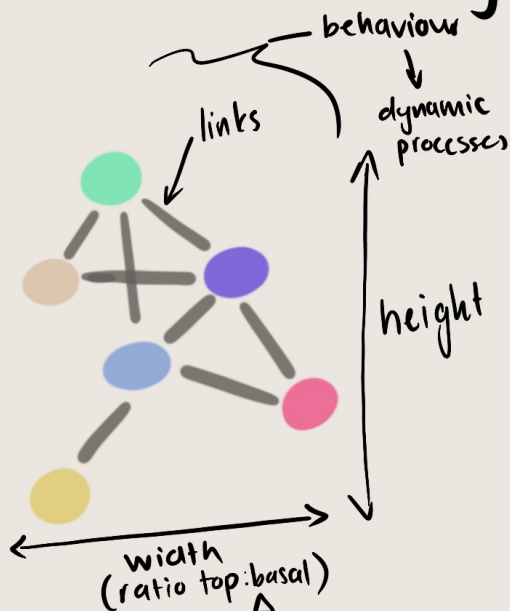


Figure 1: Conceptual figure of the ‘network prediction’. Panel A shows where the model families fall in the the context of being models that predict networks or models that predict interactions space. Panel B serves to highlight the characteristics one might like to ‘test’/benchmark for a model based on it being either a network or interaction predicting model

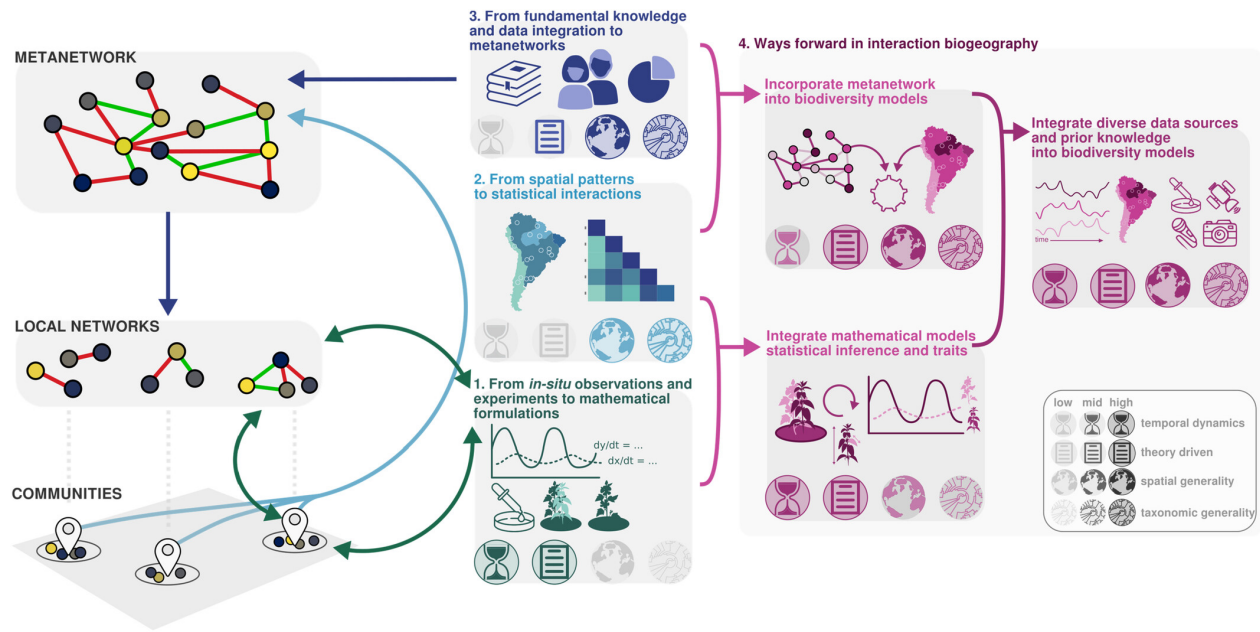


Figure 2: I like the use of the different source indicator items (not too dissimilar from Tall Tom's nature paper but also different). This is from Thuiller et al. (2024)

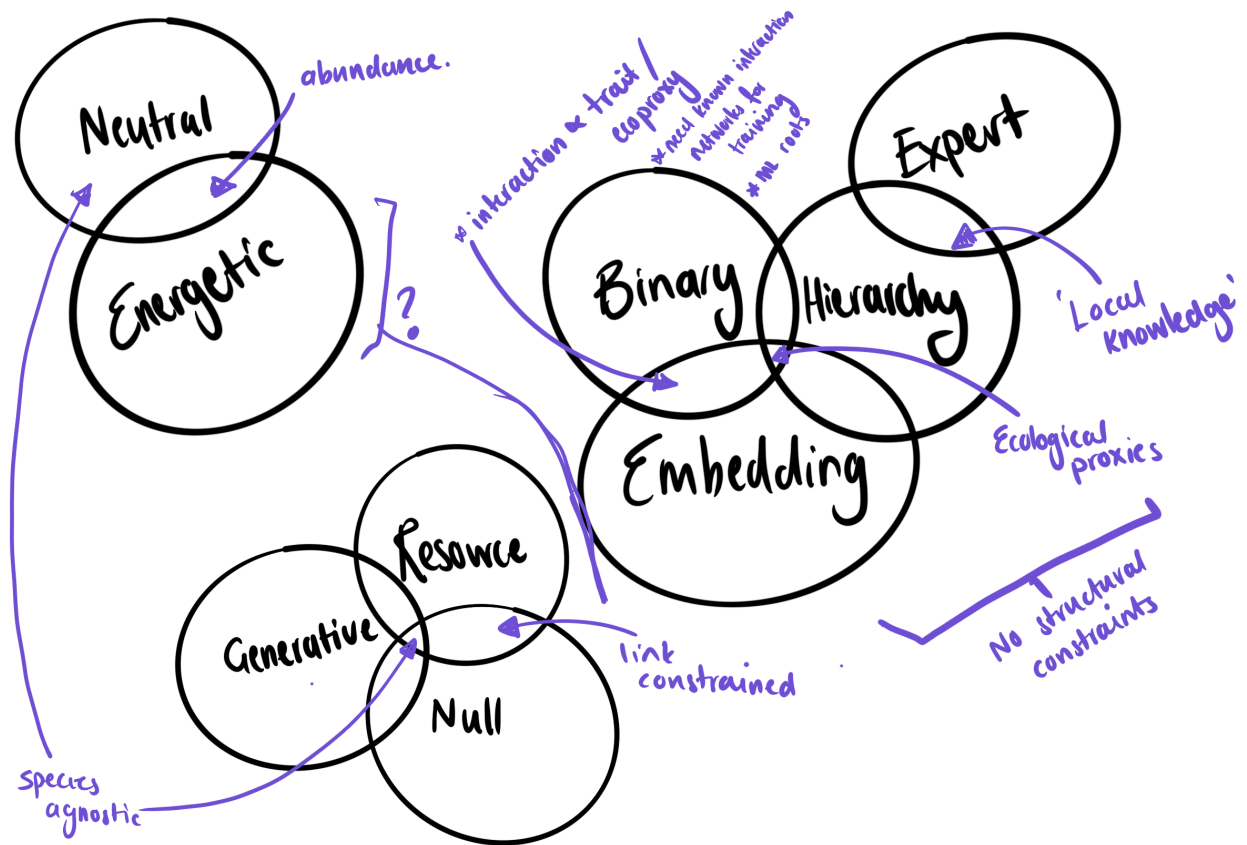


Figure 3: I still haven't given up on a sort of venn diagram idea but maybe it going to be more of a venn-flow chart hybrid...

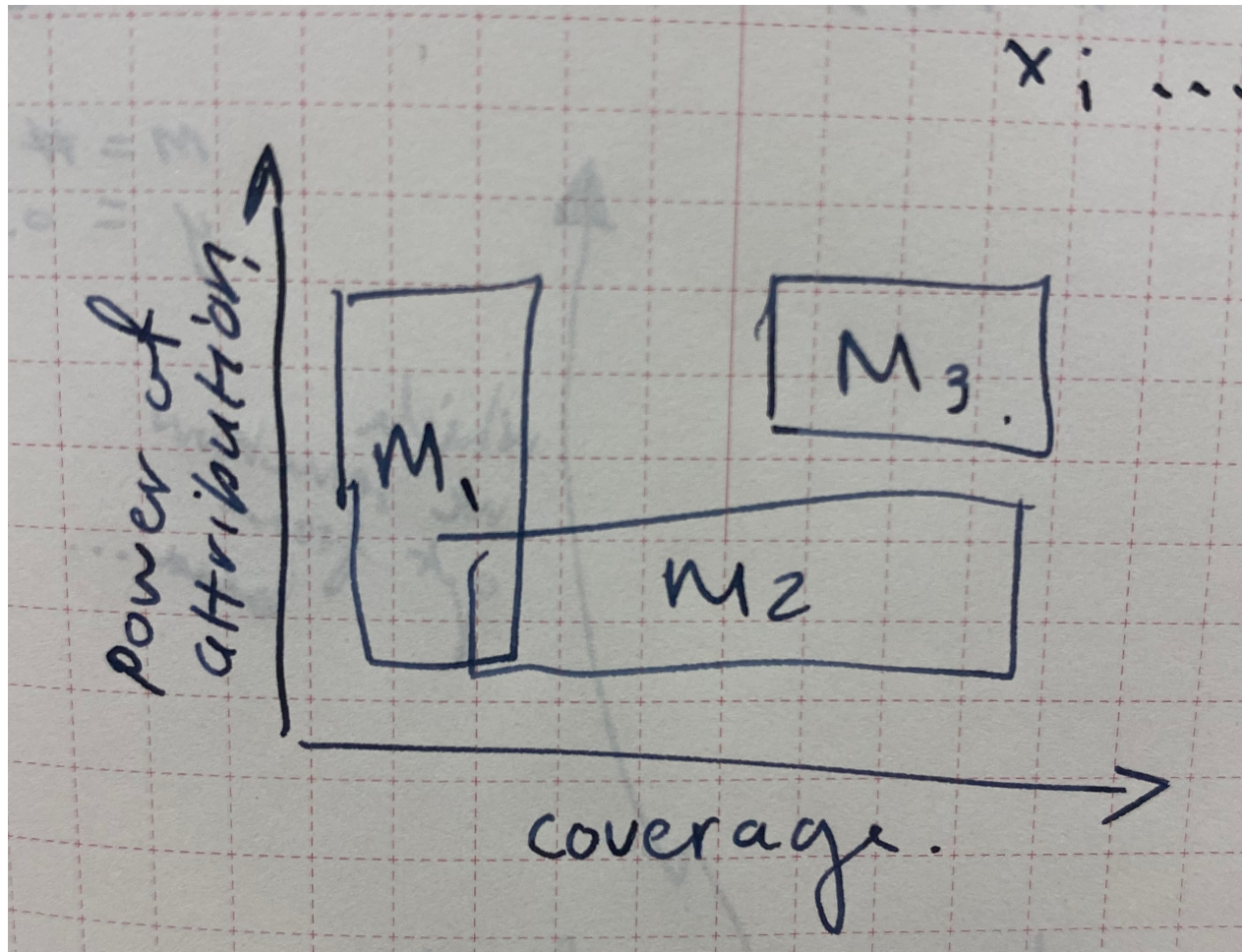


Figure 4: I like these schematics that Charlie Outhwaite presented at the EEB seminar (there was a series of them).





Figure 5: PCA of the trait table

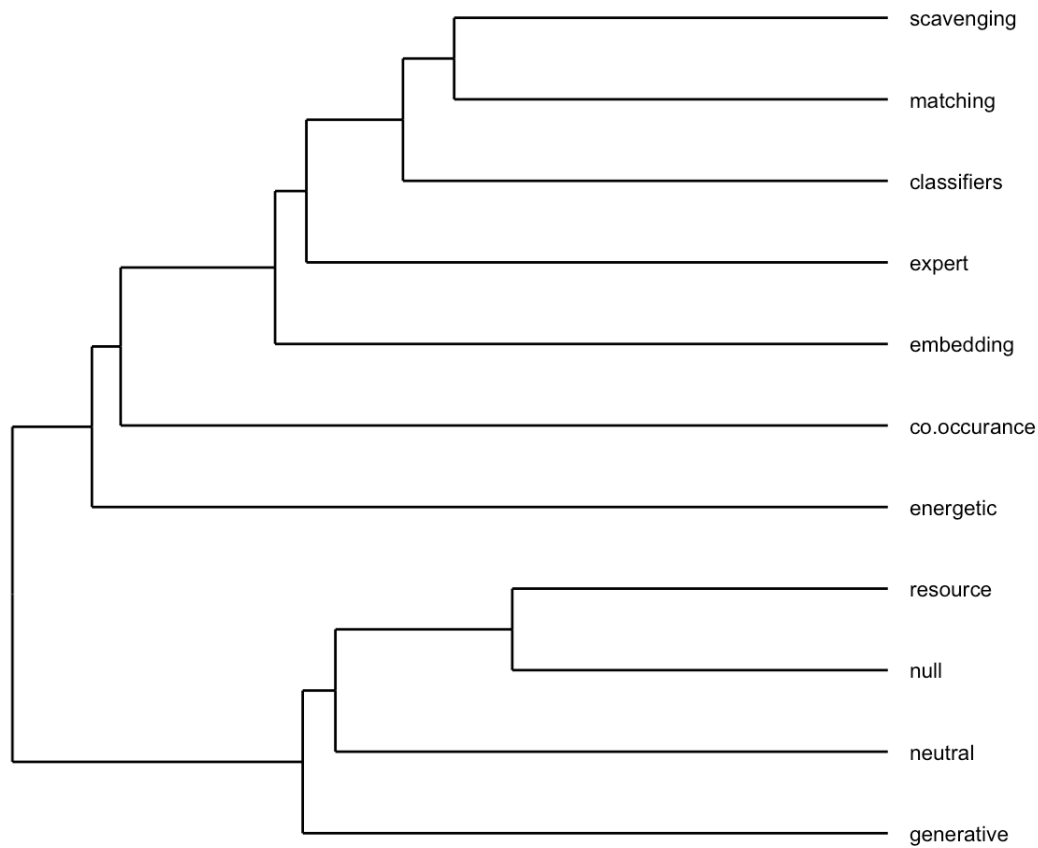


Figure 6: Dendrogram of the trait table



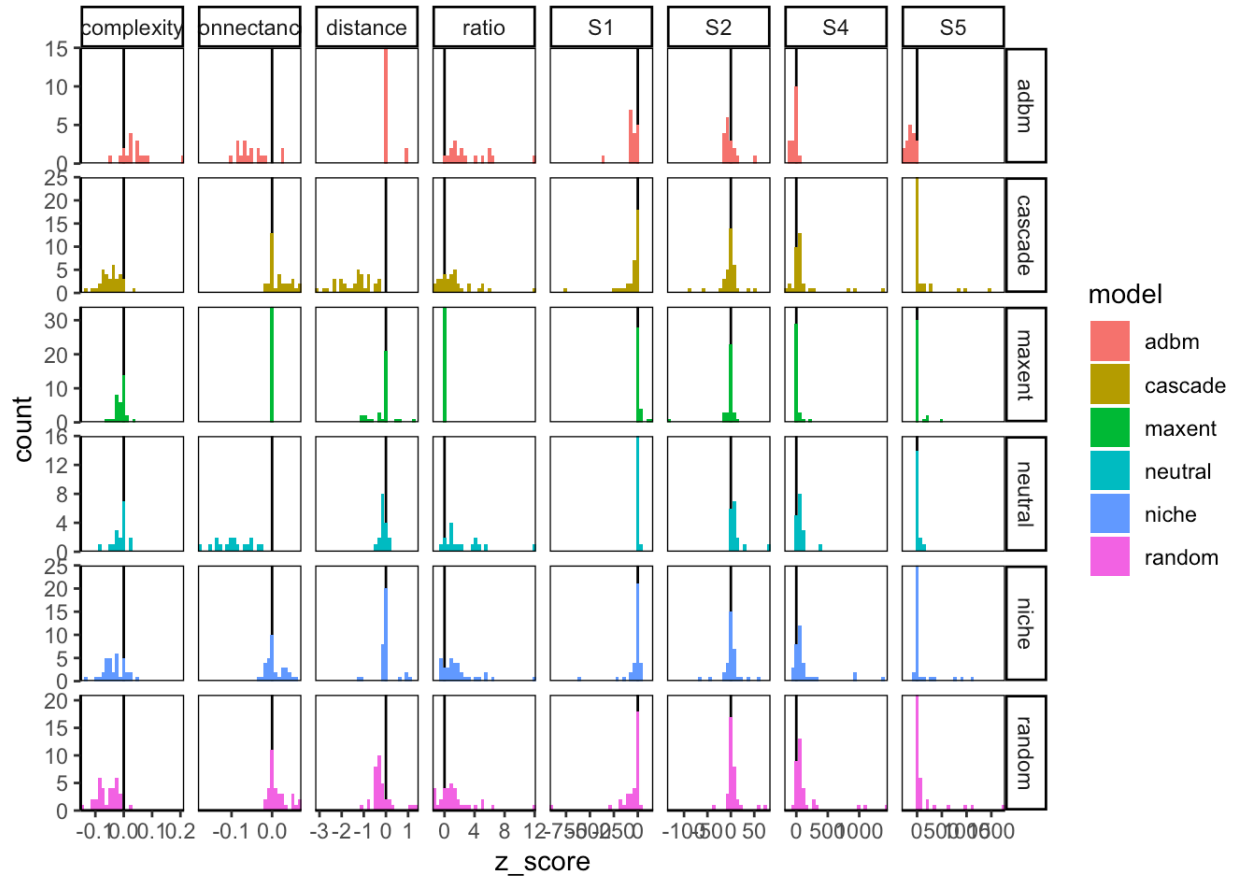


Figure 7: Difference between real and model network property. S1 - S5 represent the different motif structures identified in Stouffer et al. (2007).

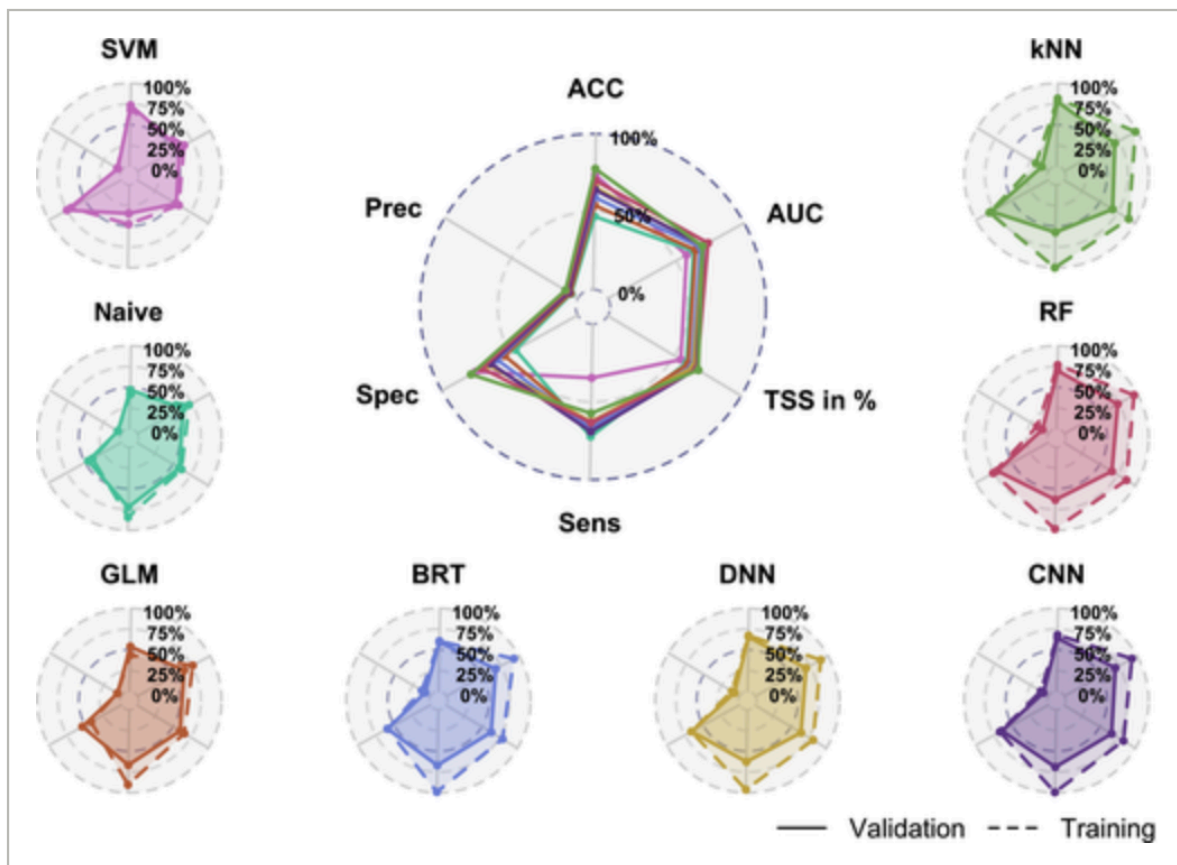


Figure 8: Cool way to conceptualise results from Pichler et al. (2020)