T is for Topology

Tanya Strydom   Jennifer A. Dunne   Timothée Poisot

Andrew P. Beckerman

2024-04-19

There are many reasons one might want to generate a network and there are many tools on the market that might make that possible. However not all tools are created equally and there is reason to assume that not all networks will suit most purposes. Here the aim is to compare and contrast the different topology generating tools that are on the market and see where they shine and where they fall flat. There probably isn't one model to rule them all but it doesn't mean that we shouldn't be critical when we think about the model we want to use.

1 Introduction

- In order to construct a 'perfect' network *i.e.,* one which *perfectly* captures the dynamics for a specific community at a specific time one needs to consider and account for many different moving parts (*e.g.,*). So when developing a model it makes sense that you prioritise the aspect of the prediction/construction task that has the most value for your research goal, acknowledging that a model might fall short in others. The thing is that with the growing suite of approaches to generating networks it is important that we don't lose sight of the core philosophy behind the model we use and to ensure that we are using the model best suited to what we want to be accomplishing.

    - 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 and Hutchinson 2020; Pringle 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.

1

– This is also going to be very important as network ecology as a field starts being incorporated into other fields *e.g.,* biogeography (Thuiller et al. 2024)

- What do we mean when we talk about an 'interaction network'? Is it the flow of biomass/energy? Is it an effect on fitness (feeding)? (can we really represent indirect interactions as actual links?)

  – "For example, nodes can be individuals and the links contacts between those individuals (Melian et al. 2011), nodes can be species and the links the interactions (e.g. feeding, mutualistic, parasitic) between them (de Ruiter et al. 1995), or nodes can be ecological communities or ecosystems and the links fluxes between these localities (Proulx et al. 2005; Chadès et al. 2011)" - Poisot, Stouffer, and Kéfi (2016)

- It is perhaps useful to start with asking why do we want/need models to generate networks. This can be broadly thought of to fall into two categories. Build networks because we want to build concepts vs build networks because we want specificity ('real world', on the ground applications). Broadly this means that we either want to construct/predict a collection of interactions (generate networks) or a network of interactions (predict interactions).

  – Also need to discuss the nuance between predicting something vs generating something, which also then bleeds into need to be aware when we use the term 'model' as well…
  – "For example, dynamic models focus on how structure relates to population dynamics and community stability (2, 7–11). Evolutionary models incorporate the processes that control the formation and expansion of food webs (12, 13). Static models include rules that determine structural attributes of food webs (14–19)." - Petchey et al. (2008)

Arguably the need for methods and tools for constructing interaction networks 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 running further simulations (*e.g.,* extinction simulations) or understanding some higher-level process/concept (*e.g.,* understanding energy flows), importantly these networks do not require any level of species specificity *per se* and it is more the arrangements of the nodes (species) 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 inherit challenges of inventorying interaction in the field (Morales-Castilla et al. 2015; Strydom et al. 2021). Of course these two categories are not distinct, mutually exclusive, groups but can rather be viewed as operating on a gradient ranging from a need for generality (*i.e.,* creating a network that, when taken in aggregate, the distribution of links (interactions) between species

are ecologically plausible) to a need for specificity (*i.e.,* local-level predictions between specific species).

- 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?

  - Core mechanistic differences that models will work at — some are really concerned about (and thus constrained by) structure, others are more mechanistic in nature *i.e.,* species *a* has the capacity to eat species *b* because traits

  - 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).

- A breakdown of wanting to generate a network; statement of need and core philosophies

- A breakdown of wanting to predict an interaction; statement of need (Jordano 2016b, 2016a; Poisot et al. 2021) and core philosophies (trait-matching, coexistence, evolutionary backbones)

- Stands to reason then that we have developed methods that specialise in one or the other. Which comes at a cost of 'performance' in other aspects. Knowing how the different model families stack up to each other is thus valuable.

Joel E. Cohen, Newman, and Steele (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 and 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'.

Other works to keep in mind: Pringle and Hutchinson (2020), Pringle (2020)

> "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 and Hutchinson (2020)

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

2 Model families

Given the large number of models that have been developed it is perhaps more meaningful to group models into families with the idea that models from the same family will yield similar results because they play by similar rules. These rules referring to the underlying philosophy as to what structures either networks or the interactions within them (see Figure 1 panel A). Although there have been efforts to compare and contrast different models (*e.g.,* Williams and 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.
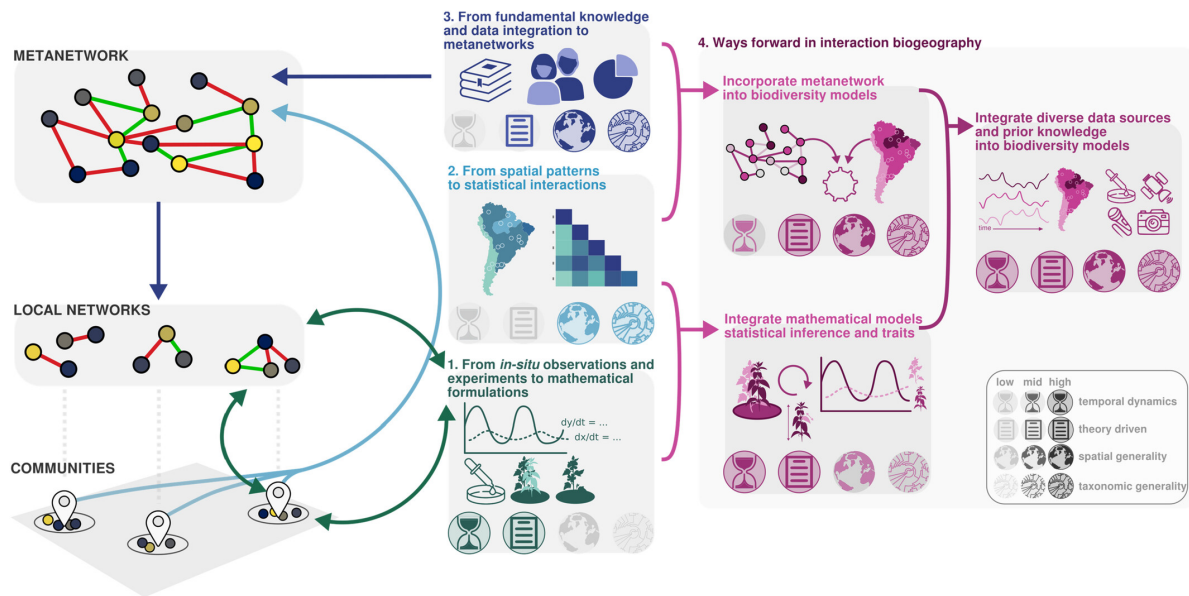

Figure 2: I like the use of the different source indicator items (not to dissimilar from Tall Tom's nature paper but also different). This is from Thuiller et al. (2024)

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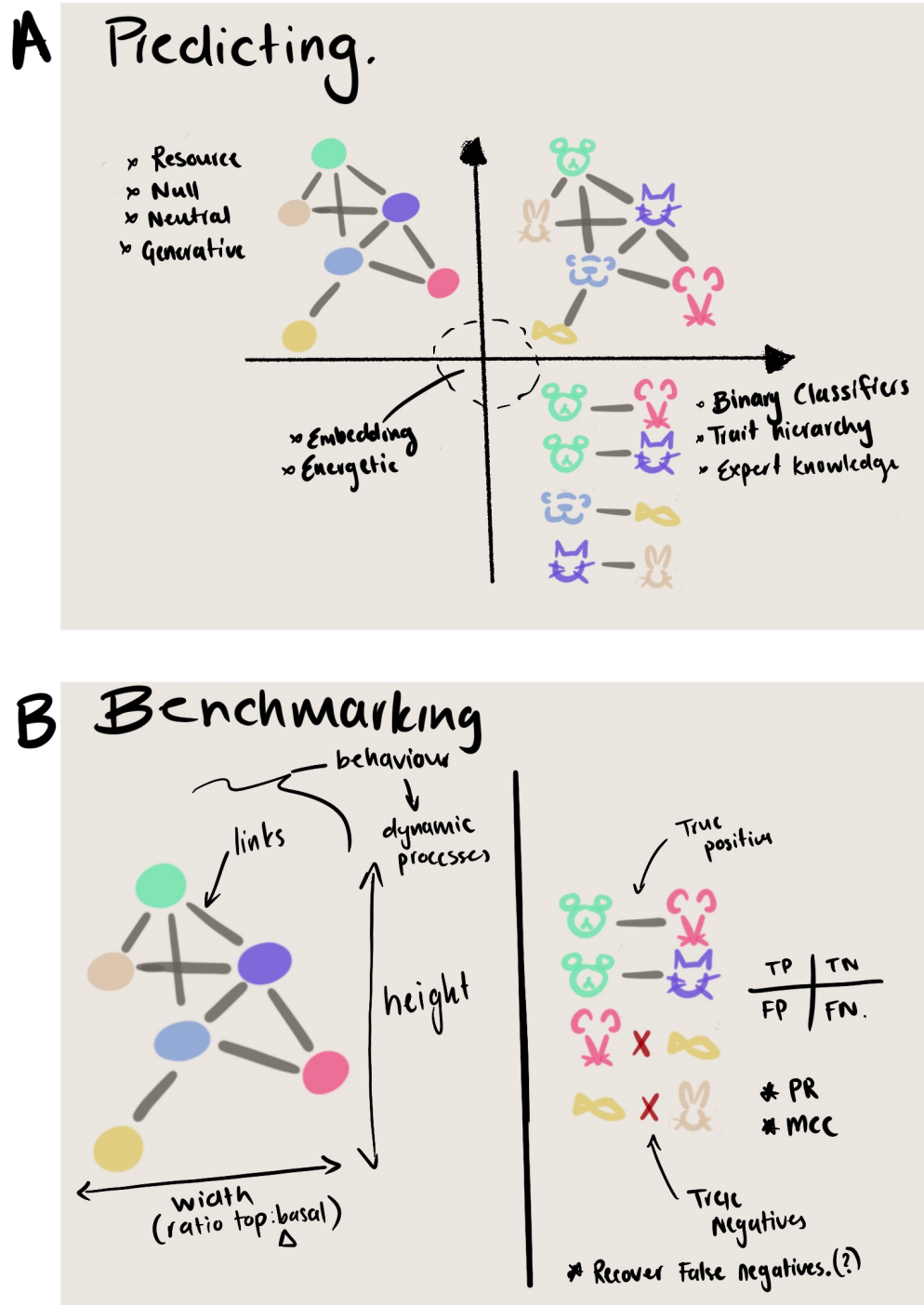
Figure 1: Conceptual figure of the 'network prediction'. Panel A shows where the model families fall in the the context of being models that predict networks or models that predict interactions space. Panel B serves to highlight the characteristics one might like to 'test'/benchmark for a model based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

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, Alonso, and Pascual (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 (Joel E. Cohen, Briand, and Newman 1990), which rests on the idea that species feed on one another in a hierarchical manner; the niche model (Williams and 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 and 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, Alonso, and Pascual (2008)).

Generative models: (this is maybe a bit of a bold term to use). MaxEnt (Banville, Gravel, and Poisot 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, Petchey, and Warren 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, Goldstein, and O'Neill (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 (*e.g.,* Eklöf, Tang, and Allesina 2013; Cirtwill et al. 2019). 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.,* Jennifer A. 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, Simons, and Mungall 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.

3 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)

3.1 Benchmarking network models

- Maybe look at some of the historic papers that compare some of the 'resource models'

- See also Allesina, Alonso, and Pascual (2008) and the likelihood function that they use for model selection

- Look at Vermaat, Dunne, and Gilbert (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, Alonso, and Pascual (2008)

3.2 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 inherit 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)

8

– 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* Jennifer A. Dunne (2006)) and a network realisation. Maybe also Poisot, Stouffer, and Gravel (2015) that discuss how the local factors are going to play a role.

4 Data & Methods

4.1 Selecting models

This section depends on if we go the family route and where we introduce them. But a more extended description of each model can be found in the `Extended Model Description` notebook (I'm trying to work out how to embed this…)

I know tables are awful but in this case they may make more sense. Also I don't think I'm at the point where I can say that the table is complete/comprehensive but it getting there Not sure about putting in some papers that have used the model - totes happy to drop those I think…

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

Model family	Theory	Example used	Predicts	Needs (minimum)	Assembly	Constraints	Interaction
null	Network structure is random		network	network (species agnostic)	random	link	binary

Model family	Theory	Example used	Predicts	Needs (minimum)	Assembly	Constraints	Interaction
neutral	Network structure is random, but species abundance plays a role	Pomeranz et al. (2019)	network	abundance, number of links	mass effect	link	binary
resource	Networks are interval, species can be ordered on a 'niche axis'		network	richness, connectance	'random'	link	binary
generative	Networks are determined by their structural features	Maximum Entropy (Banville, Gravel, and Poisot 2023)	network	network (species agnostic)	'random'		binary
energetic	Interactions are determined by foraging theory (feeding links)	ADBM (Petchey et al. 2008)	interactions	body size	deterministic	energy	

Model family	Theory	Example used	Predicts	Needs (minimum)	Assembly	Constraints	Interaction
graph embedding	Interactions can be predicted from the latent traits of networks	Transfer Learning (Strydom et al. 2022)	interactions	interactions, phylogenetic tree, list of target species (species pool)			probabilistic
trait matching	Interactions can be inferred by a mechanistic framework/relationships	PFIM (Shaw et al. 2024)	interactions	prior (expert) knowledge of trait hierarchy/relationships, traits, list of target species (species pool)	mechanistic	trait matching (*sensu* forbidden links in a way)	
binary classifiers	Interactions can be predicted by learning the relationship between interactions and ecologically relevant predictors		interactions	interactions, traits, list of target species (species pool)	statistical		

Model family	Theory	Example used	Predicts	Needs (minimum)	Assembly	Constraints	Interaction
expert knowledge	'Boots on the ground' ecological knowledge and observations		interactions	list of target species (species pool)	mechanistic		
data scavenging	Webscraping to create networks from online databases		interactions	list of target species (species pool)			

4.2 Datasets used

4.2.1 Mangal networks

We queried the Mangal (Poisot et al. 2016) database and extracted a total of **TODO** networks. *Some sort of summary as to the geographic/taxonomic range??* Although these networks represent a high volume of interaction data they do not have accompanying 'metadata' that we would need for some of the more data-hungry model families (*e.g.,* local abundance), the Mangal networks were used to provide the 'starting values' for the random, resource, and generative families. This allows us to generate a large number of different networks that we can use to compare and contrast the performance of the various model families (see Section 4.3 for a more detailed breakdown). For each network from Mangal we generated **TODO** versions of that network using each model family.

> "These complex food webs differ in their level of resolution and sampling effort, which may introduce noise in the estimation of their properties, especially given their large number of interacting elements. However, because our MaxEnt models are applied on imperfect data, they aim at reproducing the sampled structure of food webs, not their actual structure." - Banville, Gravel, and Poisot (2023) (something to think about…)

4.2.2 Empirical networks

'Elite' number of datasets for interaction models

Although the availability of empirical interaction data is growing as techniques begin to improve and grow (Pringle and Hutchinson 2020), we still lack a way to define what is the 'ideal' interaction dataset.

New Zealand dataset(s): Pomeranz et al. (2019)

> Here I think we need to span a variety of domains, at minimum aquatic and terrestrial but maybe there should be a 'scale' element as well *i.e.,* a regional and local network. I think there is going to be a 'turning point' where structural will take over from mechanistic in terms of performance. More specifically at local scales bioenergetic constraints (and co-occurrence) may play a bigger role in structuring a network whereas at the metaweb level then mechanistic may make more (since by default its about who can potentially interact and obviously not constrained by real-world scenarios) *sensu* Caron et al. (2024). Although having said that I feel that contradicts the idea of backbones (*sensu* Bramon Mora (sp?) et al & Stouffer et al) But that might be where we get the idea of core *structure* vs something like linkage density. So core things like trophic level/chain length will be conserved but connectance might not (I think I understand what I'm trying to say here)

I think we should also use the Jennifer A. Dunne et al. (2008) work. Because 1) it gives the paleo-centric methods their moment in the sun and 2) I think it also brings up the interesting question of can we use modern structure to predict past ones?

4.3 Model benchmarking

For now the (still essentially pending) workflow/associated code can be found at the following repository BecksLab/topology_generators. This will reflect that which is shown in panel *B* of Figure 1.

- Data 'cost' (some methods might need a lot lot of supporting data vs something very light weight)
- I think it would be remiss to not also take into consideration computational cost
- Something about the network output - I'm acknowledging my biases and saying that probabilistic (or *maybe* weighted) links are the way

4.3.1 Network models

Want to compare real vs predicted and then get something that looks like Figure 4

- connectance, nestedness (Bastolla et al., 2009), modularity (Barber, 2007), asymmetry (Delmas et al., 2018), and Jaccard network dissimilarity (Canard et al., 2014)

- *Shape:* do the models construct tall 'pencil' vs flat 'pancake' networks (Beckerman 2024, pers comms), generality/vulnerability, chain length (?)

- *Structure:* Predicting 'structure' - SVD (Strydom, Dalla Riva, and Poisot 2021) but maybe something like nestedness as well (?)

- *Links:* are the number of links preserved (most network predicting models are to some extend link constrained but useful to see)

- *Motifs:* Staniczenko et al. (2010) uses S1, S2, S4, S5 from Stouffer et al. (2007)

  - S1: Number of linear chains

  - S2: Number of omnivory motifs

  - S4: Number of apparent competition motifs

  - S5: Number of direct competition motifs

4.3.2 Interaction models

- Based on Poisot (2023):

  - Precision-Recall (PR-AUC) - performance
  - Matthews correlation coefficient (MCC) - accuracy

- Maybe same measures we use for the network models

4.3.3 PVA (action plan)

1. Shortlist/finalise the different topo generators
2. collate/translate into `Julia`

   - *e.g.,* some models wil be in SpeciesInteractionNetworks.jl (new EcoNet); I know (parts of) the transfer learning stuff is and the niche model
   - others will need to be coded out (the more simpler models should be easier)

3. Curate networks for the different datasets/scenarios we select - I feel like there might be some scenarios that we can't do all models for all datasets but maybe I'm being a pessimist.

14

- Need to also think about where one might find the additional data for some of the models...
  - Body size: Herberstein et al. (2022) - Although maybe Andrew has strong thotsTM RE the one true body size database to rule them all...
  - Other trait sources: Wilman et al. (2014) and Jones et al. (2009)
  - This is where we'll get the paleo traits from if I'm correct Bambach, Bush, and Erwin (2007)
  - Phylogeny stuff: Upham, Esselstyn, and Jetz (2019) (what we used for TL but its only mammals...) but I'm sure there will be others
- Also limitation of scope... *e.g.,* do we even dare to think about including plants/basal producers (see *e.g.,* Valdovinos et al. (2023))
- Taxonomic harmonisation - something to think about and check

5 Results

Joel E. Cohen, Newman, and Steele (1985) actually tells us that the cascade model only really works for communities that range from 3-33 species... and Williams and Martinez (2008) also highlights how structural models really only work for small communities

5.1 Qualitative stuff

Maybe not the best term to use but thinking here about practical limitations of the different families. This can include thinking about:

- scale limitations (time or space); *e.g.,* a metaweb is going to encapsulate but not distinguish between different seasons or locations
- data needed. I think this can be in the form of real world datasets (*e.g.,* traits) but also *a priori* knowledge (*e.g.,* having to define the constraints of a niche model)
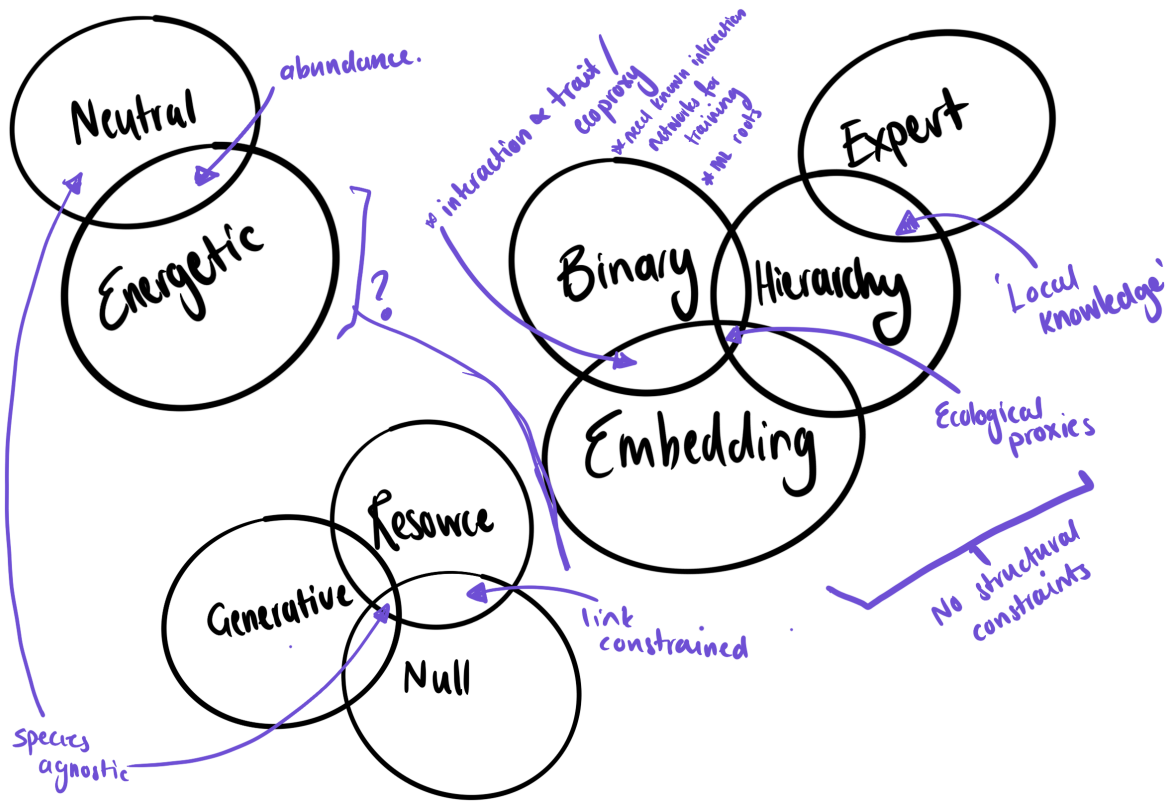- computational costs

Neutral

Energetic

abundance.

?

* interaction & trait / coproxy

* need known interaction networks for training

* m. roots

Binary

Hierarchy

Expert

'Local knowledge'

Embedding

Ecological proxies

No structural constraints

Resource

Generative

Null

link constrained

species agnostic

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...
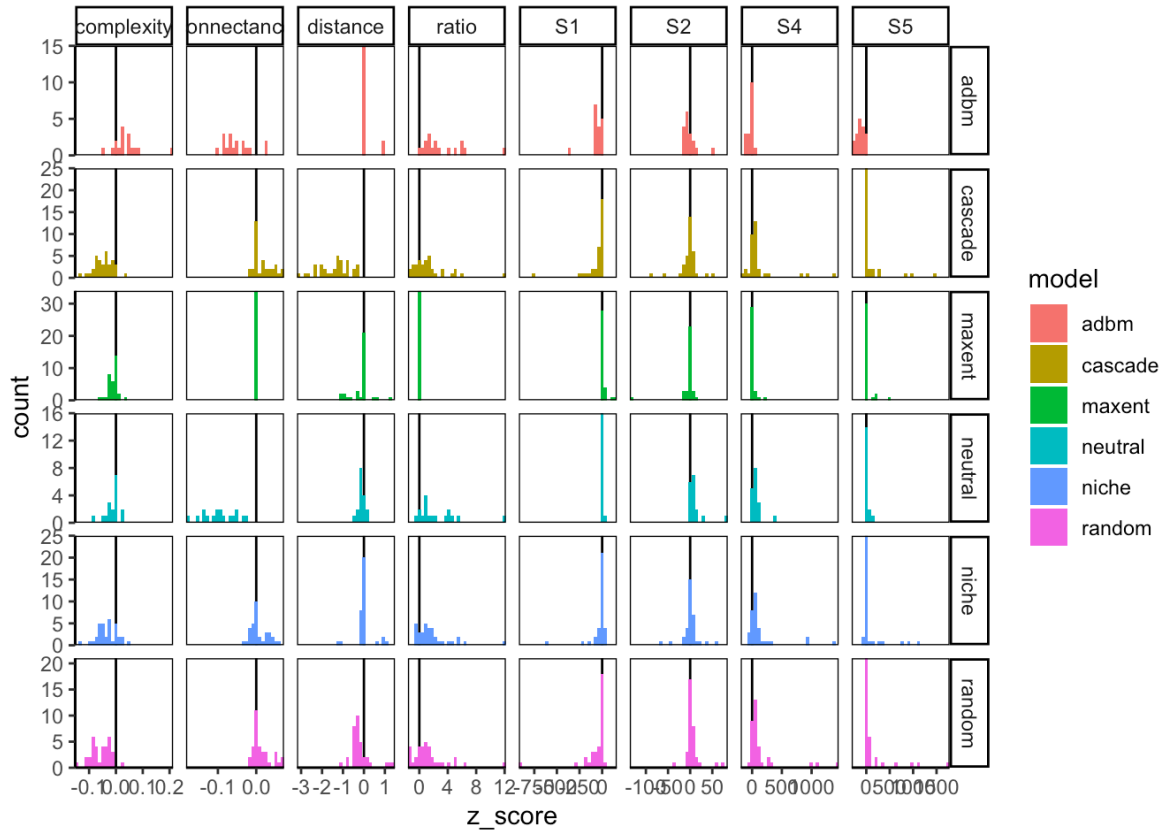
5.2 Quantitative stuff

Figure 4: DIfference between real and model network property. S1 - S5 represent the different motif structures identified in Stouffer et al. (2007).

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

6 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, Brose, and Martinez (2008) but maybe not exactly in this context but vaguely adjacent.
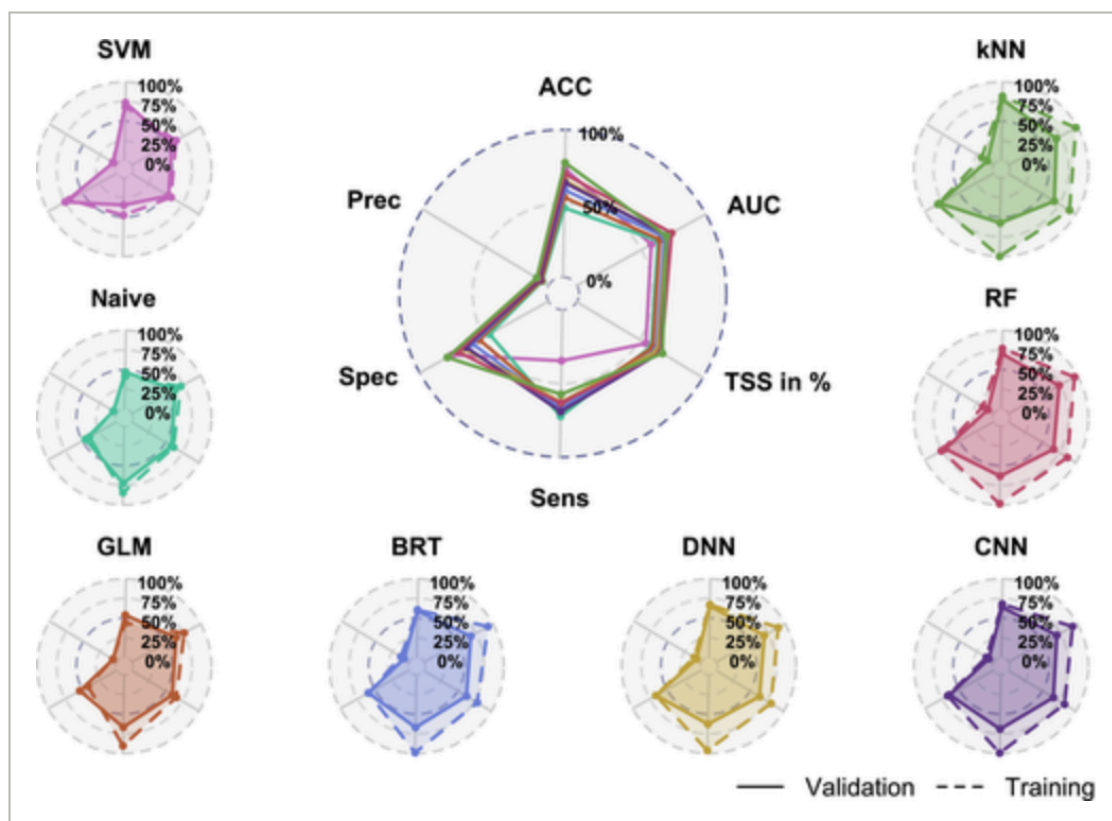
17

Figure 5: Cool way to conceptualise results from Pichler et al. (2020)

- 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, Bodner, and Fortin (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)

- 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, Stefano, David Alonso, and Mercedes Pascual. 2008. "A General Model for Food Web Structure." *Science* 320 (5876): 658–61. https://doi.org/10.1126/science.1156269.

Bambach, Richard K., Andrew M. Bush, and Douglas H. Erwin. 2007. "Autecology and the Filling of Ecospace: Key Metazoan Radiations." *Palaeontology* 50 (1): 1–22. https://doi.org/10.1111/j.1475-4983.2006.00611.x.

Banville, Francis, Dominique Gravel, and Timothée Poisot. 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., P. Jordano, C. J. Melian, and J. M. Olesen. 2003. "The Nested Assembly of Plant-Animal Mutualistic Networks." *Proceedings of the National Academy of Sciences* 100 (16): 9383–87. https://doi.org/10.1073/pnas.1633576100.

Beckerman, Andrew P., Owen L. Petchey, and Philip H. Warren. 2006. "Foraging Biology Predicts Food Web Complexity." *Proceedings of the National Academy of Sciences* 103 (37): 13745–49. https://doi.org/10.1073/pnas.0603039103.

Berlow, Eric L., Ulrich Brose, and Neo D. Martinez. 2008. "The 'Goldilocks Factor' in Food Webs." *Proceedings of the National Academy of Sciences* 105 (11): 4079–80. https://doi.org/10.1073/pnas.0800967105.

Brimacombe, Chris, Korryn Bodner, and Marie-Josee Fortin. 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, Chris, Korryn Bodner, Matthew Michalska-Smith, Timothée Poisot, and Marie-Josée Fortin. 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, Dominique, Ulrich Brose, Miguel Lurgi, F. Guillaume Blanchet, Dominique Gravel, and Laura J. Pollock. 2024. "Trait-Matching Models Predict Pairwise Interactions Across Regions, Not Food Web Properties." *Global Ecology and Biogeography* 33 (4): e13807. https://doi.org/10.1111/geb.13807.

Caron, Dominique, Luigi Maiorano, Wilfried Thuiller, and Laura J. Pollock. 2022. "Addressing the Eltonian Shortfall with Trait-Based Interaction Models." *Ecology Letters* 25 (4): 889–99. https://doi.org/10.1111/ele.13966.

Cattin, Marie-France, Louis-Félix Bersier, Carolin Banašek-Richter, Richard Baltensperger, and Jean-Pierre Gabriel. 2004. "Phylogenetic Constraints and Adaptation Explain Food-Web Structure." *Nature* 427 (6977): 835–39. https://doi.org/10.1038/nature02327.

Cirtwill, Alyssa R., Anna Eklf, Tomas Roslin, Kate Wootton, and Dominique Gravel. 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, Joel E, Frederic Briand, and Charles Newman. 1990. *Community Food Webs: Data and Theory.* Biomathematics. Berlin Heidelberg: Springer-Verlag.

Cohen, Joel E., C. M. Newman, and John Hyslop Steele. 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): 421–48. https://doi.org/10.1098/rspb.1985.0042.

DeAngelis, D. L., R. A. Goldstein, and R. V. O'Neill. 1975. "A Model for Tropic Interaction." *Ecology* 56 (4): 881–92. https://doi.org/10.2307/1936298.

Desjardins-Proulx, Philippe, Idaline Laigle, Timothée Poisot, and Dominique Gravel. 2017. "Ecological Interactions and the Netflix Problem." *PeerJ* 5: e3644. https://doi.org/10.

7717/peerj.3644.

Dunne, Jennifer A. 2006. "The Network Structure of Food Webs." In *Ecological Networks: Linking Structure and Dynamics*, edited by Jennifer A Dunne and Mercedes Pascual, 27–86. Oxford University Press.

Dunne, Jennifer A., Richard J. Williams, Neo D. Martinez, Rachel A. Wood, and Douglas H. Erwin. 2008. "Compilation and Network Analyses of Cambrian Food Webs." *PLOS Biology* 6 (4): e102. https://doi.org/10.1371/journal.pbio.0060102.

Eklöf, Anna, Si Tang, and Stefano Allesina. 2013. "Secondary Extinctions in Food Webs: A Bayesian Network Approach." Edited by Jessica Metcalf. *Methods in Ecology and Evolution* 4 (8): 760–70. https://doi.org/10.1111/2041-210X.12062.

Fortuna, Miguel A., and Jordi Bascompte. 2006. "Habitat Loss and the Structure of Plant-Animal Mutualistic Networks: Mutualistic Networks and Habitat Loss." *Ecology Letters* 9 (3): 281–86. https://doi.org/10.1111/j.1461-0248.2005.00868.x.

Gravel, Dominique, Timothée Poisot, Camille Albouy, Laure Velez, and David Mouillot. 2013. "Inferring Food Web Structure from Predator–Prey Body Size Relationships." *Methods in Ecology and Evolution* 4 (11): 1083–90. https://doi.org/10.1111/2041-210X.12103.

Gray, Clare, David H. Figueroa, Lawrence N. Hudson, Athen Ma, Dan Perkins, and Guy Woodward. 2015. "Joining the Dots: An Automated Method for Constructing Food Webs from Compendia of Published Interactions." *Food Webs* 5 (December): 11–20. https://doi.org/10.1016/j.fooweb.2015.09.001.

Herberstein, Marie E., Donald James McLean, Elizabeth Lowe, Jonas O. Wolff, Md Kawsar Khan, Kaitlyn Smith, Andrew P. Allen, et al. 2022. "AnimalTraits - a Curated Animal Trait Database for Body Mass, Metabolic Rate and Brain Size." *Scientific Data* 9 (1): 265. https://doi.org/10.1038/s41597-022-01364-9.

Jones, Kate E., Jon Bielby, Marcel Cardillo, Susanne A. Fritz, Justin O'Dell, C. David L. Orme, Kamran Safi, et al. 2009. "PanTHERIA: A Species-Level Database of Life History, Ecology, and Geography of Extant and Recently Extinct Mammals." *Ecology* 90 (9): 2648–48. https://doi.org/10.1890/08-1494.1.

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

———. 2016b. "Sampling Networks of Ecological Interactions." *Functional Ecology*, September. https://doi.org/10.1111/1365-2435.12763.

Llewelyn, John, Giovanni Strona, Christopher R. Dickman, Aaron C. Greenville, Glenda M. Wardle, Michael S. Y. Lee, Seamus Doherty, Farzin Shabani, Frédérik Saltré, and Corey J. A. Bradshaw. 2023. "Predicting Predator–Prey Interactions in Terrestrial Endotherms Using Random Forest." *Ecography* 2023 (9): e06619. https://doi.org/10.1111/ecog.06619.

Maiorano, Luigi, Alessandro Montemaggiori, Gentile Francesco Ficetola, Louise O'Connor, and Wilfried Thuiller. 2020. "TETRA-EU 1.0: A Species-Level Trophic Metaweb of European Tetrapods." *Global Ecology and Biogeography* 29 (9): 1452–57. https://doi.org/10.1111/geb.13138.

Morales-Castilla, Ignacio, Miguel G. Matias, Dominique Gravel, and Miguel B. Araújo. 2015. "Inferring Biotic Interactions from Proxies." *Trends in Ecology & Evolution* 30 (6): 347–56. https://doi.org/10.1016/j.tree.2015.03.014.

Newman, Mark E. J. 2010. *Networks. An Introduction.* New York, NY: Oxford University Press.

Petchey, Owen L., Andrew P. Beckerman, Jens O. Riede, and Philip H. Warren. 2008. "Size, Foraging, and Food Web Structure." *Proceedings of the National Academy of Sciences* 105 (11): 4191–96. https://doi.org/10.1073/pnas.0710672105.

———. 2011. "Fit, Efficiency, and Biology: Some Thoughts on Judging Food Web Models." *Journal of Theoretical Biology* 279 (1): 169–71. https://doi.org/10.1016/j.jtbi.2011.03.019.

Pichler, Maximilian, Virginie Boreux, Alexandra-Maria Klein, Matthias Schleuning, and Florian Hartig. 2020. "Machine Learning Algorithms to Infer Trait-Matching and Predict Species Interactions in Ecological Networks." *Methods in Ecology and Evolution* 11 (2): 281–93. https://doi.org/10.1111/2041-210X.13329.

Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. "Global Biotic Interactions: An Open Infrastructure to Share and Analyze Species-Interaction Datasets." *Ecological Informatics* 24 (November): 148–59. https://doi.org/10.1016/j.ecoinf.2014.08.005.

Poisot, Timothée. 2023. "Guidelines for the Prediction of Species Interactions Through Binary Classification." *Methods in Ecology and Evolution* 14 (5): 1333–45. https://doi.org/10.1111/2041-210X.14071.

Poisot, Timothée, Benjamin Baiser, Jennifer Dunne, Sonia Kéfi, François Massol, Nicolas Mouquet, Tamara N. Romanuk, Daniel B. Stouffer, Spencer A. Wood, and Dominique Gravel. 2016. "Mangal – Making Ecological Network Analysis Simple." *Ecography* 39 (4): 384–90. https://doi.org/10.1111/ecog.00976.

Poisot, Timothée, Gabriel Bergeron, Kevin Cazelles, Tad Dallas, Dominique Gravel, Andrew MacDonald, Benjamin Mercier, Clément Violet, and Steve Vissault. 2021. "Global Knowledge Gaps in Species Interaction Networks Data." *Journal of Biogeography* n/a (n/a). https://doi.org/10.1111/jbi.14127.

Poisot, Timothée, Daniel B. Stouffer, and Dominique Gravel. 2015. "Beyond Species: Why Ecological Interaction Networks Vary Through Space and Time." *Oikos* 124 (3): 243–51. https://doi.org/10.1111/oik.01719.

Poisot, Timothée, Daniel B. Stouffer, and Sonia Kéfi. 2016. "Describe, Understand and Predict: Why Do We Need Networks in Ecology?" *Functional Ecology* 30 (12): 1878–82. https://www.jstor.org/stable/48582345.

Pomeranz, Justin P. F., Ross M. Thompson, Timothée Poisot, and Jon S. Harding. 2019. "Inferring Predator–Prey Interactions in Food Webs." *Methods in Ecology and Evolution* 10 (3): 356–67. https://doi.org/10.1111/2041-210X.13125.

Pringle, Robert M. 2020. "Untangling Food Webs." In *Untangling Food Webs*, 225–38. Princeton University Press. https://doi.org/10.1515/9780691195322-020.

Pringle, Robert M., and Matthew C. Hutchinson. 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.

Rohr, Rudolf Philippe, Heike Scherer, Patrik Kehrli, Christian Mazza, and Louis-Félix Bersier. 2010. "Modeling Food Webs: Exploring Unexplained Structure Using Latent Traits." *The American Naturalist* 176 (2): 170–77. https://doi.org/10.1086/653667.

Shaw, Jack O., Alexander M. Dunhill, Andrew P. Beckerman, Jennifer A. Dunne, and Pincelli

M. Hull. 2024. "A Framework for Reconstructing Ancient Food Webs Using Functional Trait Data." bioRxiv. https://doi.org/10.1101/2024.01.30.578036.

Staniczenko, Phillip P. A., Owen T. Lewis, Nick S. Jones, and Felix Reed-Tsochas. 2010. "Structural Dynamics and Robustness of Food Webs." *Ecology Letters* 13 (7): 891–99. https://doi.org/10.1111/j.1461-0248.2010.01485.x.

Stouffer, Daniel B, Juan Camacho, Wenxin Jiang, and Luís A Nunes Amaral. 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): 1931–40. https://doi.org/10.1098/rspb.2007.0571.

Strydom, Tanya, Salomé Bouskila, Francis Banville, Ceres Barros, Dominique Caron, Maxwell J. Farrell, Marie-Josée Fortin, et al. 2022. "Food Web Reconstruction Through Phylogenetic Transfer of Low-Rank Network Representation." *Methods in Ecology and Evolution* 13 (12): 2838–49. https://doi.org/10.1111/2041-210X.13835.

Strydom, Tanya, Salomé Bouskila, Francis Banville, Ceres Barros, Dominique Caron, Maxwell J. Farrell, Marie-Josée Fortin, et al. 2023. "Graph Embedding and Transfer Learning Can Help Predict Potential Species Interaction Networks Despite Data Limitations." *Methods in Ecology and Evolution* 14 (12): 2917–30. https://doi.org/10.1111/2041-210X.14228.

Strydom, Tanya, Michael D. Catchen, Francis Banville, Dominique Caron, Gabriel Dansereau, Philippe Desjardins-Proulx, Norma R. Forero-Muñoz, et al. 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.

Strydom, Tanya, Giulio V. Dalla Riva, and Timothée Poisot. 2021. "SVD Entropy Reveals the High Complexity of Ecological Networks." *Frontiers in Ecology and Evolution* 9. https://doi.org/10.3389/fevo.2021.623141.

Thuiller, Wilfried, Irene Calderón-Sanou, Loïc Chalmandrier, Pierre Gaüzère, Louise M. J. O'Connor, Marc Ohlmann, Giovanni Poggiato, and Tamara Münkemüller. 2024. "Navigating the Integration of Biotic Interactions in Biogeography." *Journal of Biogeography* 51 (4): 550–59. https://doi.org/10.1111/jbi.14734.

Upham, Nathan S., Jacob A. Esselstyn, and Walter Jetz. 2019. "Inferring the Mammal Tree: Species-level Sets of Phylogenies for Questions in Ecology, Evolution, and Conservation." *PLOS Biology* 17 (12): e3000494. https://doi.org/10.1371/journal.pbio.3000494.

Valdovinos, Fernanda S., Kayla R. S. Hale, Sabine Dritz, Paul R. Glaum, Kevin S. McCann, Sophia M. Simon, Elisa Thébault, William C. Wetzel, Kate L. Wootton, and Justin D. Yeakel. 2023. "A Bioenergetic Framework for Aboveground Terrestrial Food Webs." *Trends in Ecology & Evolution* 38 (3): 301–12. https://doi.org/10.1016/j.tree.2022.11.004.

Vermaat, Jan E, Jennifer A Dunne, and Alison J Gilbert. 2009. "Major Dimensions in Food-Web Structure Properties." *Ecology* 90 (1): 278–82. https://www.ncbi.nlm.nih.gov/pubmed/19294932.

Williams, Richard J., and Neo D. Martinez. 2000. "Simple Rules Yield Complex Food Webs." *Nature* 404 (6774): 180–83. https://doi.org/10.1038/35004572.

———. 2008. "Success and Its Limits Among Structural Models of Complex Food Webs." *Journal of Animal Ecology* 77 (3): 512–19. https://doi.org/10.1111/j.1365-2656.2008.01362.

x.

Wilman, Hamish, Jonathan Belmaker, Jennifer Simpson, Carolina de la Rosa, Marcelo M. Rivadeneira, and Walter Jetz. 2014. "EltonTraits 1.0: Species-level Foraging Attributes of the World's Birds and Mammals." *Ecology* 95 (7): 2027–27. https://doi.org/10.1890/13-1917.1.

Xie, Jia-Rong, Pan Zhang, Hai-Feng Zhang, and Bing-Hong Wang. 2017. "Completeness of Community Structure in Networks." *Scientific Reports* 7 (1): 5269. https://doi.org/10.1038/s41598-017-05585-6.