

1

T is for Topology

2

Tanya Strydom¹, Andrew P. Beckerman¹

3

¹University of Sheffield,

Corresponding author: Tanya Strydom, t.strydom@sheffield.ac.uk

Abstract

Pending...

Plain Language Summary

We want to know a bit more about the different network topology generators (predict tools) and how they differ - *i.e.*, their strengths and weaknesses

1 Introduction

The standard run of the mill that we cannot always feasibly construct networks because 1. hard, 2. time (yay dinosaurs, but also the future and impending doom I guess), and 3. probably something else meaningful that's just slipping my mind at the moment. Some of the usual culprits will come in here like: Jordano (2016b); (Jordano, 2016a); Poisot et al. (2021); Strydom et al. (2021) (shameless self plug)

Maybe a brief history of the development of predictive tools/topo generators? Sort of where the theory/body of work was based and how that has changed? IS there a difference between toppo generator and predictive tool - I'm inclined to think that it aligns with the whole debate of high level structure vs node-level perfection

Maybe start here with discussing the 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 (read gob size), and then you get Rohr et al. (2010) and Strydom et al. (2022) that sit in the weird liminal latent space...

Here I will probably get on my (newly discovered) soapbox and wax lyrical about how in certain situations structure is enough (and that will probably be for some high-level things like thinking about energy flows etc., I can also see a world in which maybe you want to do some sort of robustness/extinction work - since then you're usually doing 'random' (within limits) extinctions) but there may be use cases where we are really interested in the node-level interactions *i.e.*, species identity is like a thing we need to care about and also 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).

At some point we are going to need to discuss the key differences and implications between predicting a metaweb (*sensu* Jennifer A. Dunne (2006)) and a network realisation. And here I can't help but think about Poisot et al. (2015) (and probably other papers) that discuss how the local factors are going to play a role and even the same pair of species may interact differently in different points in the landscape.

Do we need to delve into individual-based networks? (*sensu* Tinker 2012, Araújo 2008) I think its probably a step too far and one starts creeping into apples and pears type of comparisons. Especially since these work off of already existing networks (I seem to recall) and its more about about 'tweaking' those - so not so much *de novo* predictions. Although this might be useful to keep in mind when it comes to re-wiring... Also on that note do we open the re-wiring door here in this ms or wait it out a bit.

2 Data & Methods**2.1 Overview of topology generators**

I know table 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

52 it getting there Not sure about putting in some papers that have used the model -
53 totes happy to drop those I think...

Table 1: Lets make a table that gives an overview of the different topology generators that we will look at

Approach	Reference	Core Mechanism	<i>e.g.</i> , uses	Applications (potential)	End product	Theory basis	Assumptions
Cascade model	Joel E. Cohen et al. (1990); Joel E. Cohen et al. (1997)	structural					
Niche model	Williams & Martinez (2000); Williams & Martinez (2008)	structural			networks have a finite number of links and species are partitioned along a resource	A network configuration for a given number of species & connectance	connectance, and others
PFIM	Shaw et al. (2024)	mechanistic	Dunhill (in review)	prehistoric networks (but can be modified)	Metaweb (I'd argue)	traits determine interactions	feeding mechanisms are trait correlated
Log-ratio	Rohr et al. (2010)	latent trait space	Yeakel et al. (2014), Pires et al. (2020) (?)				
Nested hierarchy	Cattin et al. (2004)						
ADBM	Petchey et al. (2008)	mechanistic	probably multiple		working on it...	diet and body size determine interactions	size and resources (many) limitations

Approach	Reference	Core Mechanism	<i>e.g.</i> , uses	Applications (potential)	End product	Theory basis	Assumptions
Stochastic	Rossberg et al. (2006)						
Graph Embedding	Strydom et al. (2023)	latent trait space	Strydom et al. (2022)	Generate metaweb for specific species pool	Probabilistic metaweb	interactions are ‘en-coded’ and can thus be extracted and transferred to similar species	evolutionary backbones/signals
Trait-based	Caron et al. (2022)	mechanistic	Caron et al. (2023)		metaweb	traits determine interactions	feeding mechanisms are trait correlated

Might be nice to have a little appendix/supp mat that breaks down the models in detail so that they are all in one place so that someone (grad student being told they need to build networks) some day can go and educate themselves with slightly lower effort. This will also be useful for me should I end up having to do some actual coding - think of this as step one in the pseudo code process.

2.2 Datasets used

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. (2023). 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 Dunne (I think) Cambrian (also think) network (I was correct and its this one Jennifer A. Dunne et al. (2008)). 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? Here one might expect a more mechanistic approach to shine.

Draw the other datasets from **Mangal** because they will be nicely formatted and essentially at point and shoot level

2.3 Comparing different models

For now the (still essentially pending) workflow/associated code can be found at the following repository [BecksLab/topology_generators](#)

1. Shortlist/finalise the different topo generators
2. collate/translate into **Julia**
 - *e.g.*, some models will 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)
 - can also consider **R** but then it becomes a case of porting things left and right depending on how we decide to do the post analyses
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.
 - 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 et al. (2007)
 - Phylogeny stuff: Upham et al. (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
4. compare model performance based on the ideas currently listed in the results section.
5. Make a pretty picture that summarises things - maybe overlapping Venn circles that showcase which models do well in the different spheres/aspects of life

3 Results

How we want to compare and contrast. I think there won't be a 'winner' and thus we need to think of 'tests' that are going to measure performance in different situations/settings. With that in mind I think some valuable points to consider would be:

- Structural vs pairwise link predictions (graph vs node level)
 - % of links correctly retrieved
 - connectance
 - trophic level
 - generalism vs specialism
 - something related to false positives/negatives
 - intervality
- 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

maybe we can put these into broader categories - if we do start doing the venn overlap thing. *E.g.*, local scale predictions, regional scale predictions, pairwise interactions, structural (energetics), computationally cheap, low cost data

3.1 Qualitative stuff

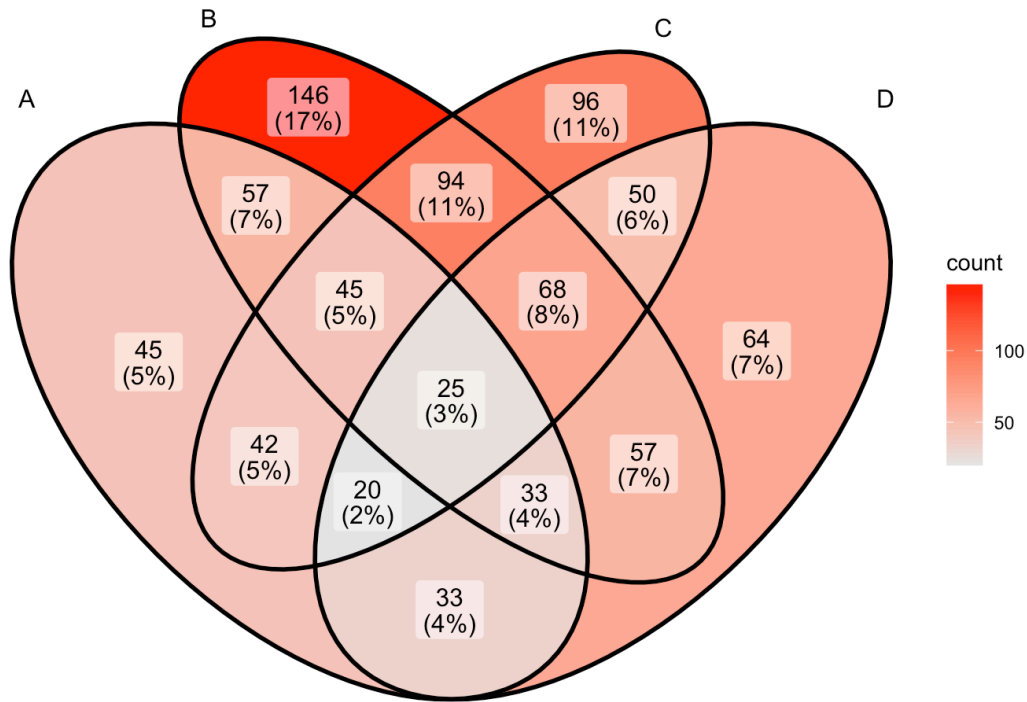


Figure 1: Venn diagram for qualitative analysis/overview of the fancy maths things

Source: [Article Notebook](#)

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...

I probably think about this point too much but a point of discussion that I think 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? I think this is covered with the whole node vs graph level performance but I kind of just want to bring it up here again because also one of those things that I think about a bit too much probably...

Thinking very long term here and maybe a bit beyond the scope but also thinking about a multi- model approach? So in other words using one model to build an initial network but maybe a second one to constrain it a bit better. I blame this thought on the over-connected PFIM food webs...

References

Source: [Article Notebook](#)

- Bambach, R. K., Bush, A. M., & Erwin, D. H. (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>
- 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>
- 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>
- Caron, D., Brose, U., Lurgi, M., Blanchet, G., Gravel, D., & Pollock, L. J. (2023, May). Trophic interaction models predict interactions across space, not food webs. *EcoEvoRxiv*. <https://doi.org/10.32942/X29K55>
- 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>
- Cohen, Joel E., Briand, F., & Newman, C. (1990). *Community Food Webs: Data and Theory*. Berlin Heidelberg: Springer-Verlag.
- Cohen, Joel E., Newman, C. M., & Steele, J. H. (1997). 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>
- Dunne, Jennifer A. (2006). The Network Structure of Food Webs. In Jennifer A. Dunne & M. Pascual (Eds.), *Ecological networks: Linking structure and dynamics* (pp. 27–86). Oxford University Press.
- Dunne, Jennifer 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>
- Herberstein, M. E., McLean, D. J., Lowe, E., Wolff, J. O., Khan, M. K., Smith, K., 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, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., et al. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, 90(9), 2648–2648. <https://doi.org/10.1890/08-1494.1>
- 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>
- 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>

- Pires, M. M., Rindel, D., Moscardi, B., Cruz, L. R., Guimarães, P. R., dos Reis, S. F., & Perez, S. I. (2020). Before, during and after megafaunal extinctions: Human impact on Pleistocene-Holocene trophic networks in South Patagonia. *Quaternary Science Reviews*, 250, 106696. <https://doi.org/10.1016/j.quascirev.2020.106696>
- 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., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., et al. (2021). Global knowledge gaps in species interaction networks data. *Journal of Biogeography*, n/a(n/a). <https://doi.org/10.1111/jbi.14127>
- 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>
- Rossberg, A. G., Matsuda, H., Amemiya, T., & Itoh, K. (2006). Food webs: Experts consuming families of experts. *Journal of Theoretical Biology*, 241(3), 552–563. <https://doi.org/10.1016/j.jtbi.2005.12.021>
- Shaw, J. O., Dunhill, A. M., Beckerman, A. P., Dunne, J. A., & Hull, P. M. (2024, January). A framework for reconstructing ancient food webs using functional trait data. *bioRxiv*. <https://doi.org/10.1101/2024.01.30.578036>
- Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., 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, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M. J., et al. (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., 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), 2917–2930. <https://doi.org/10.1111/2041-210X.14228>
- Upham, N. S., Esselstyn, J. A., & Jetz, W. (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, F. S., Hale, K. R. S., Dritz, S., Glaum, P. R., McCann, K. S., Simon, S. M., et al. (2023). A bioenergetic framework for aboveground terrestrial food webs. *Trends in Ecology & Evolution*, 38(3), 301–312. <https://doi.org/10.1016/j.tree.2022.11.004>
- 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. *The Journal of Animal Ecology*, 77(3), 512–519. <https://doi.org/10.1111/j.1365-2656.2008.01362.x>
- Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world’s birds and mammals. *Ecology*, 95(7), 2027–2027. <https://doi.org/10.1890/13-1917.1>
- Yeakel, J. D., Pires, M. M., Rudolf, L., Dominy, N. J., Koch, P. L., Guimarães, P. R., & Gross, T. (2014). Collapse of an ecological network in Ancient Egypt. *PNAS*, 111(40), 14472–14477. <https://doi.org/10.1073/pnas.1408471111>