

# Methods for Reconstructing Paleo Food Webs

Tanya Strydom <sup>1</sup>; Baran Karapunar <sup>2</sup>; Andrew P. Beckerman <sup>1</sup>; Alexander Dunhill <sup>2</sup>

**Abstract:** Food webs provide a powerful lens for understanding ecosystem structure and function, but reconstructing them in paleoecological contexts remains challenging because direct evidence of feeding interactions is rarely preserved. A wide range of models now exist for predicting interactions and inferring network structure, yet these models differ substantially in their assumptions, mechanisms, and data requirements. Here, we evaluate which network construction approaches are most suitable for paleo food webs given the constraints of the fossil record, and we assess how model choice influences the networks we infer. Using the Toarcian Oceanic Anoxic Event as a case study, we compare six modelling approaches encompassing mechanistic, structural, and theory-based methods. We show that the models produce strikingly different network structures and pairwise interactions, and that these differences propagate into ecological inference—including conclusions about extinction dynamics. Our results highlight the importance of aligning model choice with research questions and underscore the interpretative risks of treating all food-web reconstruction methods as interchangeable.

**Keywords:** food web, network construction

1 There has been growing interest in using deep-time fossil data to understand how ecological communities  
2 respond to environmental change (Dillon et al., 2022; Kiessling et al., 2019). In modern systems, species  
3 interactions and the networks that they form have become central to studying biodiversity, energy flow, and  
4 community stability (*e.g.*, Thuiller et al. (2024)). Consequently, paleoecologists have increasingly sought to  
5 reconstruct ancient food webs to gain similar insights across major environmental transitions (*e.g.*, Dunhill et  
6 al. (2024); Hao et al. (2025); Yeakel et al. (2014)). However, we are faced by the limitation that interactions  
7 cannot be directly observed in the fossil record (with the exception of rare instances) and as a result, the  
8 construction of paleo food webs depends on models that allow us to infer feeding relationships from preserved  
9 traits, analogies to modern taxa, or ecological theory. While numerous models exist for predicting interactions  
10 (see Morales-Castilla et al. (2015); Pichler & Hartig (2023); Strydom et al. (2021); Allesina et al. (2008) for  
11 broader reviews), only a subset can reliably be applied in paleo contexts, where data on traits, abundances,  
12 and community composition are inherently incomplete and biased.

13 The growing interest in paleo food webs has outpaced a clear discussion of *which* construction methods are  
14 suitable for which purposes. Different models generate different kinds of networks (feasible, realised, or purely  
15 structural) and these differences can fundamentally alter ecological interpretations. In this study, we evaluate  
16 a suite of methods that can be feasibly applied to paleo communities and explore how their underlying  
17 assumptions shape both network structure and ecological inference. Specifically we focus on identifying a suite  
18 of models that are appropriate for use with paleo data that can feasibly be constructed within the limitations  
19 that are imposed by fossil data while still spanning the larger network space. Here we use the data from  
20 Dunhill et al. (2024) as a case study to understand how different models recover different networks, both in  
21 terms of structure as well as pairwise interactions and establish if there are consequences for using networks  
22 that are constructed using different models in terms of making inferences about the behaviour of the system  
23 by looking at how the model type influences what we infer to be the dominant driver of extinctions across a  
24 mass extinction event.

## 25 1 Constructing paleo webs

### 26 1.1 Challenges specific to building paleo networks

27 Reconstructing paleo food webs presents challenges that differ from those encountered in modern systems.  
28 First, the fossil record provides an incomplete and selectively preserved subset of the original community.  
29 Preservation biases (driven by habitat, skeletal composition, and sedimentary environment) mean that  
30 some trophic groups are over-represented (*e.g.*, hard-shelled organisms), while others (*e.g.*, soft-bodied taxa,

31 plankton) are systematically under-sampled. This directly constrains the kinds of models that can be applied,  
32 because models requiring complete assemblages or accurate guild representation will perform poorly when  
33 preservation is uneven. Second, there is inherent uncertainty about the true community boundaries. Fossil  
34 assemblages may represent time-averaged accumulations, transported material, or mixed habitats. As a result,  
35 any species list is best interpreted as a set of taxa that could have interacted, rather than a snapshot of  
36 a specific subset of interaction species. Third, many extinct taxa have ambiguous trait states, especially  
37 regarding diet and behaviour. Even when functional morphology is preserved, ecological behaviour is seldom  
38 directly evident. Such uncertainty propagates differently across model families: mechanistic models tend to  
39 accommodate broad trait assignments, whereas theory-driven models are more sensitive to uncertainty in  
40 body size, foraging mode, or feeding constraints. These limitations do not render reconstruction impossible  
41 but highlight the importance of choosing a model whose assumptions match the type of ecological inference  
42 being attempted.

## 43 1.2 Understanding the approaches to network construction

44 Network construction approaches can be broadly grouped into three methodological and conceptual approaches  
45 Figure 1. The first are mechanistic models evaluate whether an interaction is *feasibly possible*. These models  
46 typically use trait-based rules (*e.g.*, feeding mode, body size, or functional morphology) or evolutionary  
47 relationships to determine whether a species *could* consume another. Mechanistic approaches tend to produce  
48 metawebs - the full set of all plausible interactions given biological constraints. Theory-driven models, embed  
49 assumptions from ecological theory (such as niche theory or foraging ecology) to generate realised interactions  
50 and networks. Structural models are similar to theory-driven models, with the exception that these models are  
51 species agnostic and as such can only be used to make inferences about network structure. Both theory-driven  
52 and structural models aim to reproduce characteristic patterns observed in modern food webs, such as  
53 intervality, trophic hierarchies, or body-size-scaled feeding ranges. They do not necessarily require detailed  
54 trait information and instead rely on ecological rules or statistical distributions consistent with empirical food  
55 webs.

56 [Figure 1 about here.]

57 Most existing paleo-specific approaches fall within the mechanistic tradition (*e.g.*, Shaw et al. (2024);  
58 Roopnarine (2006); Fricke et al. (2022)). While these are well-suited for reconstructing feasible interactions,  
59 they represent only a subset of the broader space of network construction methods. Incorporating theory-driven  
60 models allows paleoecologists to explore realised interaction structures and address a wider suite of ecological  
61 questions—provided their assumptions are compatible with the limitations of fossil data. Here we present a

62 range of models Table 1 that carry specific assumptions and data requirements. For instance, allometric models  
 63 depend on quantitative body-size estimates, which must be inferred from size classes or functional morphology  
 64 in the fossil record. Structural models such as the niche model require only richness and connectance, but their  
 65 species-agnostic nature limits their usefulness for trait-based or diet-specific questions. Mechanistic models  
 66 rely on accurate assignment of feeding traits, which may be uncertain for extinct taxa but are often more  
 67 tractable than estimating abundances or interaction strengths. Understanding how these limitations intersect  
 68 with what fossil data can reliably provide is essential for selecting an appropriate modelling approach.

Table 1: Six different models that can be used to construct food webs for both this specific community but are also broadly suited to paleo network prediction. These models span all facets of the network representation space (metaweb, realised, and structural network) and are suitable for an array of different paleo communities as the data requirements fall within the limitations set by the fossil record.

Model family	Assumptions	Data needs	'Limitation'	Network type	Key reference
random	Links are randomly distributed within a network	richness, number of links	parameter assumptions, species agnostic	structural network	Erdős & Rényi (1959)
niche	Networks are interval, species can be ordered on a 'niche axis'	richness, connectance	parameter assumptions, species agnostic	structural network	Williams & Martinez (2008)
allometric diet	Interactions are determined by energetic costs (foraging ecology)	body mass, biomass (abundance)	does not account for forbidden links in terms of trait compatibility, assumptions on body size and biomass (abundance) from fossil data	theoretical network	Petchey et al. (2008)
breadth model (ADBM)					

Model family	Assumptions	Data needs	'Limitation'	Network type	Key reference
l-matrix	Interactions inferred using allometric rules (ratio of body sizes between predator and prey), with links being constrained by a Ricker function	body mass, number of producer species	does not account for forbidden links in terms of trait compatibility, assumptions on body size from fossil data, assumptions as to the number of producer species	theoretical network	Schneider et al. (2016)
paleo food web inference model (PFIM)	Interactions can be inferred by a mechanistic framework/relationships	feeding traits for taxa, mechanistic feeding rules	Assumption made as to the feeding mechanisms, need to elucidate traits (although this is a way smaller issue)	mechanistic web	Shaw et al. (2024)

Model family	Assumptions	Data needs	‘Limitation’	Network type	Key reference
body size ratio model	Interactions inferred using allometric rules (ratio of body sizes between predator and prey). Logit of the linking probability used to further constrain links to an ‘optimal size range’ for prey.	body mass	does not account for forbidden links in terms of evolutionary compatibility, assumptions on body size from fossil data	theoretical network	Rohr et al. (2010)

## 69 2 Case study: Toarcian mass extinction event

### 70 2.1 Dataset overview

#### 71 2.1.1 Species occurrence

72 We used fossil occurrence data spanning the upper Pliensbachian (~185 Ma) to the upper Toarcian (~175  
 73 Ma) of the Cleveland Basin, following Dunhill et al. (2024) . The dataset comprises four paleo-communities  
 74 representing the pre-extinction, post-extinction, early recovery, and late recovery intervals of the Toarcian  
 75 Oceanic Anoxic Event. Each assemblage was treated as a community of potentially interacting taxa. (*[Insert*  
 76 *here]: total number of taxa; numbers per time bin; brief note on sampling completeness?]*)  
 77 Although the Toarcian occurs within “deep time,” the depositional environment of the Cleveland Basin  
 78 provides an unusually well-resolved record of community change, making it a suitable system for comparing  
 79 network reconstruction methods.

80 **2.1.2 Defining modes of life (traits)**

81 We used the modes of life (traits) as identified in Dunhill et al. (2024), who defined four traits: motility (fast,  
82 slow, facultative, non-motile), tiering (pelagic, erect, surficial, semi-infaunal, shallow infaunal, deep infaunal),  
83 feeding (predator, suspension feeder, deposit feeder, mining, grazer), and size: gigantic ( $>500$  mm), very large  
84 ( $>300\text{--}500$  mm), large ( $>100\text{--}300$  mm), medium ( $>50\text{--}100$  mm), small ( $>10\text{--}50$  mm), tiny ( 10 mm), for each  
85 fossil species based on the ecological traits defined in the Bambach ecospace model (Bambach et al., 2007).

86 **2.1.3 Constructing networks**

87 For each paleo community, we constructed 100 networks using each of the models listed in Table 1 (6 models  $\times$   
88 4 time intervals  $\times$  100 replicates = 2,400 networks). Networks were then simplified by removing disconnected  
89 species, ensuring that all nodes participated in at least one interaction. Models requiring body-size inputs  
90 (ADBM, l-matrix, and body-size ratio models) were parameterised by drawing body masses from uniform  
91 distributions bounded by the size-class limits assigned in Dunhill et al. (2024). This approach propagates  
92 uncertainty inherent in fossil size estimates while preserving consistent relative sizes among species within a  
93 replicate. For each replicate, the same set of body masses was used across models that depend on size. For  
94 structural models (random and niche), connectance was drawn uniformly from 0.07–0.34 to ensure networks  
95 spanned a realistic range of empirical food-web connectances, while holding richness constant. The same  
96 connectance value was used for both models within a replicate to facilitate direct comparison. For each  
97 network, we calculated the metrics listed in Table 2, capturing macro-, meso-, and micro-scale structural  
98 properties.

99 **2.2 Models capture different network structure but in unexpected ways**

100 When quantifying network structure, we are essentially asking how interactions are distributed across species  
101 and how these patterns scale from individual nodes to the whole community. Structural metrics are informative  
102 because they reflect underlying ecological processes: how energy flows through trophic levels, how disturbances  
103 propagate, where redundancy or fragility exists, and how species specialise or generalise in their diets. To  
104 capture these different facets, we evaluated a suite of macro-, meso-, and micro-scale metrics Table 2, ranging  
105 from global properties like connectance and complexity to motifs and species-level generality and vulnerability.

Table 2: Network properties used for analysis.

Metric	Definition	Scale	Reference (for maths), can make footnotes probs
Richness	Number of nodes in the network	Macro	
Links	Normalized standard deviation of links (number of consumers plus resources per taxon)	Micro	
Connectance	$L/S^2$ , where $S$ is the number of species and $L$ the number of links	Macro	
Max trophic level	Prey-weighted trophic level averaged across taxa	Macro	Williams & Martinez (2004)
Diameter	Diameter can also be measured as the average of the distances between each pair of nodes in the network	Macro	Delmas et al. (2018)
Complexity	SVD complexity of a network, defined as the Pielou entropy of its singular values	Macro	Strydom et al. (2021)
Redundancy	$(L - (S - 1))/S$ , where $S$ is the number of species and $L$ the number of links. Indicates the number of edges beyond what is needed for a minimum-connected tree	Macro	

Metric	Definition	Scale	Reference (for maths), can make footnotes probs
S1	Number of linear chains, normalised	Meso	Milo et al. (2002); Stouffer et al. (2007)
S2	Number of omnivory motifs, normalised	Meso	Milo et al. (2002); Stouffer et al. (2007)
S4	Number of apparent competition motifs, normalised	Meso	Milo et al. (2002); Stouffer et al. (2007)
S5	Number of direct competition motifs, normalised	Meso	Milo et al. (2002); Stouffer et al. (2007)
Generality	Normalized standard deviation of generality of a species standardized by $L/S$	Micro	Williams & Martinez (2000)
Vulnerability	Normalized standard deviation of vulnerability of a species standardized by $L/S$	Micro	Williams & Martinez (2000)

- 106 Despite being supplied with the same species pools, the different models generated networks with systematically  
 107 different structural signatures. A MANOVA revealed a strong multivariate effect of model type on network  
 108 structure (Pillai's Trace = 3.89,  $p < 0.001$ ), indicating that each modelling approach produces a distinct  
 109 'structural fingerprint'. Follow-up ANOVAs confirmed that model choice had substantial effects on every  
 110 metric we examined, with effect sizes typically exceeding 0.80. The only exception was maximum trophic  
 111 level ( $\eta^2 = 0.19$ ), suggesting some convergence in vertical structure even when structure differs widely.
- 112 Post-hoc comparisons Figure 2 revealed three broad clusters. PFIM consistently yielded the most connected,  
 113 dense networks, reflecting its mechanistic emphasis on trait-based feasibility. Niche, random, and body-size  
 114 ratio models formed an intermediate group, producing networks with moderate connectance and motif  
 115 frequencies. ADBM and l-matrix formed a tight cluster characterised by constrained feeding ranges, reflecting

their shared basis in energetic and allometric theory. Although these groupings broadly align with *a priori* expectations about model families, several patterns emerged that were less intuitive. Most notably, the body-size ratio model which is theoretically grounded but aligned more closely with structural models than with the fully allometric ones. This suggests that even slight differences in how body-size constraints are implemented can shift a model's position within the network-structure landscape. A Linear Discriminant Analysis further illustrated the distinctiveness of the model families Figure 3. Classification accuracy of the LDA was 85%, demonstrating that the combination of dependent variables effectively discriminates among model types, indicating the strong imprint that model assumptions leave on inferred ecological patterns. The PFIM model is strongly separated, while ADBM and l-matrix networks cluster closely together. The niche and random models occupy intermediate positions. Importantly these results indicate that model choice is the dominant driver of inferred network structure, often overwhelming the ecological signal embedded within the species pool itself. Models designed to reproduce feasible interactions and those designed to generate realised niche structure occupy fundamentally different portions of 'network space', even when operating on identical taxa. Such structural divergences could have direct implications for ecological inference, particularly when comparing networks across time or when using network metrics to infer processes such as community stability, trophic organisation, or susceptibility to cascading extinctions.

[Figure 2 about here.]

[Figure 3 about here.]

### 2.3 Some networks don't share any interactions and some share a lot

Beyond differences in global structure, researchers are often interested in specific ecological relationships, *e.g.*, who eats who, which species share predators, and how trophic roles change across communities or time. For these types of questions, it is essential to understand how models differ at the level of pairwise interactions. To quantify this, we measured interaction turnover between networks to allow us to assess the degree to which two models predict the same or different links for the same set of species. This is analogous to -diversity but applied to links rather than species (Poisot et al., 2012). Specifically we only looked at the dissimilarity of interaction between shared species. Even when supplied with identical species pools, models varied dramatically in the interactions they inferred Figure 4. Some pairs of models showed substantial agreement, whereas others shared almost no interactions at all. Note here that we did not include the Random or Niche models as these networks are species agnostic and as such are not designed for inferring species-specific pairwise links.

The body-size ratio model had consistently high turnover relative to all others, indicating that it inferred

<sup>147</sup> diets that were largely distinct. This reflects the strong constraints imposed by its logit-based linking rule,  
<sup>148</sup> which sharply restricts prey to a narrow ‘optimal’ size range. Small differences in body-size estimates or  
<sup>149</sup> functional groupings therefore lead to disproportionately large changes in inferred interactions. In contrast, the  
<sup>150</sup> ADBM and L-matrix showed low turnover between each other, reflecting their shared theoretical foundations  
<sup>151</sup> (operationalising foraging decisions and energetic constraints using similar allometric principles). As a result,  
<sup>152</sup> they tend to produce similar pairwise interactions even when implemented independently. The PFIM exhibited  
<sup>153</sup> intermediate turnover, sharing more interactions with size-based theoretical models than with the body-size  
<sup>154</sup> ratio model. This makes sense: although PFIM uses categorical traits and hierarchical feeding rules rather  
<sup>155</sup> than quantitative foraging theory, these constraints will still produce broadly similar trophic groupings.

<sup>156</sup> Taken together, these patterns show that pairwise interactions differ far more across models than global  
<sup>157</sup> metrics alone might suggest. Two models with superficially similar connectance or trophic structure may  
<sup>158</sup> nonetheless infer completely different diets for individual taxa. This has large implications for any question  
<sup>159</sup> focused on species-level ecology, including predator-prey specialisation, trophic niche breadth, or the identity  
<sup>160</sup> of keystone consumers. These findings reinforce the importance of selecting a model whose assumptions  
<sup>161</sup> align with the intended inference. If the goal is to explore the full set of *possible* interactions a species could  
<sup>162</sup> have had, a mechanistic model such as PFIM is appropriate. If instead the goal is to infer the likely *realised*  
<sup>163</sup> interactions or energy pathways, models grounded in allometric foraging theory (ADBM, L-matrix) will  
<sup>164</sup> provide more ecologically coherent results. Conversely, models like the body-size ratio may be too restrictive or  
<sup>165</sup> idiosyncratic for diet-based questions because they force interactions into narrow, trait-determined windows.

<sup>166</sup> [Figure 4 about here.]

## <sup>167</sup> 2.4 Model choice changes the narrative

<sup>168</sup> The structural and interaction-level differences documented above raise a central question: do different models  
<sup>169</sup> also lead to different interpretations of ecological dynamics? In other words, does model choice merely affect the  
<sup>170</sup> architecture of the reconstructed networks, or does it shape the actual stories we tell about how communities  
<sup>171</sup> collapsed and recovered during the Toarcian extinction event? Using the pre-extinction networks as starting  
<sup>172</sup> points, we simulated species losses under a suite of ecologically plausible extinction scenarios, including  
<sup>173</sup> trait-based removals (e.g., body size, motility), network-position removals (e.g., vulnerability, generality), and  
<sup>174</sup> random extinctions. In each case, we allowed for cascading secondary extinctions. For each model and scenario,  
<sup>175</sup> we then measured how closely the simulated post-extinction network resembled the real fossil community.

176 **2.4.1 Inferred extinction drivers**

177 To assess how consistently different modelling approaches evaluate extinction scenarios, we quantified the  
178 agreement in scenario rankings produced by multiple models across several network metrics. For each model,  
179 each extinction scenario, and each network metric we calculated the mean absolute difference (MAD) between  
180 the observed metric value and the value predicted following the simulated extinction sequence. Lower MAD  
181 values indicate a closer match to the empirical network structure and therefore a better-performing extinction  
182 scenario for that model and metric. Additionally, we used a derivative of Gupta et al. (2022) true skill statistic  
183 (TSS, see Equation 1), where a score below zero indicates that the simulated extinction performs no better  
184 than random, and a score of one indicates a perfect match between real and simulated. Here we calculated  
185 both a node-level TSS as well as link-level TSS, by parsing out the TSS into two components we are able  
186 to assess if differences between real and simulated networks are due to node-level (the wrong species being  
187 removed) or link-level (the wrong links be recovered) mismatches. Because the extinction simulations do not  
188 allow for the origination of species, when calculating the TSS we only retained species that were present in  
189 both the pre and post extinction community and so any node-level mismatches between real and simulated  
190 networks was due to the wrong species being removed and not because new species were not.

$$TSS = \frac{TruePositive}{TruePositive + FalseNegative} + \frac{TrueNegative}{TrueNegative + FalsePositive} - 1 \quad (1)$$

191 For each network metric, we treated each model as an independent evaluator of scenario performance. MAD  
192 and TSS values were converted to within-model rankings, with rank 1 assigned to the scenario with the  
193 smallest MAD (*i.e.*, the closest match to the empirical value) or highest TSS score. Ranking was performed  
194 independently for each combination of model and network metric to avoid assumptions about comparability  
195 across metrics. To evaluate whether different models produced consistent rankings of extinction scenarios,  
196 we quantified rank correlation among models separately for each network metric. Agreement among model  
197 rankings was assessed using Kendall's rank correlation coefficient ( ), which measures the degree of agreement  
198 between two ordinal rankings. Kendall's was selected because it is robust for small sample sizes, handles  
199 tied ranks appropriately, and provides a direct measure of the probability that model pairs agree or disagree  
200 on the relative ordering of scenarios. Kendall's ranges from  $-1$  to  $+1$ , where  $+1$  indicates perfect agreement  
201 between rankings,  $0$  reflects no relationship, and  $-1$  represents complete disagreement such that one ranking is  
202 the exact reverse of the other.

203 When we look at Kendall's for the MAD across network structure and models Figure 5 we see that generally  
204 there is a positive correlation between the different different models. This implies that different models

are often recovering a similar ranking of extinction mechanisms (as in the ‘signal’ as to which extinction mechanisms may be the most plausible are the same). Although there is not a strong agreement between models as values tend to be low it is positive to observe that it is not often that we have a completely different ranking of extinction mechanisms, with the exception of complexity and the number of direct competition motifs. When looking at the macro-level network properties the random model often showcases a disagreement in terms of the MAD. This is unsurprising as we expect random networks to produce networks that are not ecologically sound and thus will not behave as one may expect (Ings et al., 2009). Interestingly we once again see the strong similarity between the L-matrix and the ADBM (have a high Kendall’s  $\tau$ ). meaning that they recover a similar ranking of extinction mechanisms, this is unsurprising given that we know these networks tend to recover a similar structure Figure 2. Broadly when we look at the behaviour of the different model families (with the exception of the Random model) we see that they recover similar structural signals with regards to the mechanisms potentially driving extinctions.

[Figure 5 about here.]

When looking at the node-level TSS scores (Figure 5, TSS, panel 2) we see that in general the signal of the extinction mechanism is maintained across the different models. However as many of the extinction mechanisms are determined by the *traits* of the node it is not surprising that we see a similar signal as the species are being removed in the exact same order. Link-level TSS scores (Figure 5, TSS, panel 1) do not show the same within extinction mechanism ranking/signal. We see that the Random and PFIM models have high TSS scores (*i.e.*, have a ‘good fit’), however in the case of the PFIM this is to be expected as the links are deterministic and so if you have the same two species pools you will recover the same links. The ‘stochastic’ element of the theoretical models (ADBM, l-matrix, and log-ratio) means that they create a degree of noise at the link-level and thus they are probably inappropriate to use for the type of extinction mechanism question we are asking here - specifically does the real and the simulated network look the same. Link-level TSS is perhaps also not an appropriate approach to determine the ‘best fit’ extinction mechanism if used in isolation and we advocate that the node-level TSS score (or alternatively some measure of  $\beta$  diversity is used). Finally, if we were to focus only on node level TSS we do not observe any strong differences between the models and it suggests that node-level driven (topological) extinction processes are insensitive to model type.

Broad inferred extinction mechanisms were relatively robust across models. This is probably in part because species are removed in the same order, node-level outcomes (which species survive) tended to agree. However, at the link level, where secondary cascades depend sensitively on inferred interactions, models often showed limited agreement. PFIM produced consistent link-level outcomes due to its deterministic rules, whereas theory-driven models (ADBM, L-matrix, body-size ratio) generated more variable trajectories due to stochastic

237 link assignment. As a result, different models sometimes reconstructed different pathways of collapse, inferred  
238 different trophic groups as being the most affected, or ranked extinction scenarios differently. In other words,  
239 while the high-level narrative (e.g., traits matter) is stable, the fine-grained story of ecosystem disruption  
240 (who lost interactions first, how cascades unfolded, which species acted as bottlenecks) changes depending on  
241 the chosen model. Thus, the Toarcian mass extinction looks subtly but meaningfully different through the  
242 lens of each modelling framework. Model choice therefore must be treated as a core component of ecological  
243 inference, not a neutral preprocessing step. Some broad signals are robust—especially those driven by species  
244 traits—but many of the finer details that paleoecologists care about, such as trophic cascading pathways,  
245 keystone taxa, or the ordering of collapse, depend strongly on the chosen model. Researchers must therefore  
246 treat model choice not as a technical detail but as a central component of ecological inference.

### 247 3 Model Choice as an Ecological Inference Decision

248 Reconstructing food webs from fossil data is an exercise in inference under uncertainty, and our results  
249 demonstrate that the choice of network construction model is itself a major ecological decision. Despite using  
250 the same species pools, different models produced networks with profoundly different structural properties,  
251 interaction patterns, and inferred extinction dynamics. These differences emerge not from the fossil data  
252 themselves but from the assumptions embedded within each modelling approach. As a consequence, network  
253 reconstruction cannot be treated as a neutral methodological step: model choice fundamentally shapes the  
254 ecological narratives we extract from the fossil record.

#### 255 3.1 What our results demonstrate about model families

256 Across every structural metric we measured, model identity explained the majority of variation. PFIM-produced  
257 networks were consistently the most connected, while the ADBM and l-matrix produced sparser networks  
258 with tighter energetic constraints. Structural models (niche and random) fell between these extremes. These  
259 clusters reflect the conceptual differences between feasible, realised, and structural network representations.  
260 Importantly, these differences were not superficial. Pairwise interaction turnover revealed that models often  
261 disagreed not just on the number of links but on who interacts with whom. Such discrepancies directly affect  
262 diet-based ecological questions and interpretations of trophic roles.

#### 263 3.2 Consequences for interpreting paleoecological dynamics

264 The downstream effects of model choice were most evident when evaluating extinction dynamics. Models often  
265 recovered comparable high-level patterns (e.g., trait-driven extinction signals remained robust across models),

266 but the specific ranking of extinction mechanisms varied. Node-level patterns were relatively stable because  
267 extinction scenarios based on intrinsic traits remove species in the same order regardless of network structure.  
268 However, link-level outcomes—particularly those concerning secondary extinctions, network collapse, and  
269 recovery trajectories—were highly model-dependent.

270 This means that conclusions about *how* an extinction unfolded, *which* interactions disappeared first,  
271 or *what* drove cascading losses are not simply properties of the fossil data but emerge from the assumptions  
272 embedded in the chosen model. Researchers therefore need to be explicit about the type of ecological signal  
273 they are trying to recover and select models accordingly.

### 274 3.3 Aligning ecological questions with model choice

275 Our findings support a simple but crucial principle: different ecological questions require different types of  
276 network representations. The assumption that all food-web models are interchangeable is not supported by  
277 our results.

278 Below is a conceptual summary (corresponding to a possible Figure 6):

Ecological question	Recommended model type	Rationale
What <i>could</i> have eaten what? (feasible interactions; trophic potential)	Mechanistic (PFIM, trait-based, evolutionary-signal models)	Captures all trait-feasible links; appropriate for reconstructing dietary breadth or assessing potential niches.
What interactions were <i>likely realised</i> ?	Theory-driven (ADBM, l-matrix)	Encodes foraging ecology and energetic constraints; provides ecologically plausible realised networks.
How do global structural properties vary?	Structural models (niche, random as baseline)	Useful for understanding generic patterns (connectance, trophic depth) without trait reliance.
How do traits influence dynamics (e.g., cascade sensitivity)?	Theory-driven models	Capture body-size scaling, energetic constraints, and trophic dependencies.

<b>Ecological question</b>	<b>Recommended model type</b>	<b>Rationale</b>
How robust are inferences to uncertainty?	Ensembles across multiple model families	Allows quantifying how conclusions depend on assumptions.

<sup>279</sup> **3.4 Implications for the use of food webs in paleoecology**

<sup>280</sup> Our analysis underscores three important implications for the broader field:

- <sup>281</sup> 1. **Interpretations of ancient food webs must explicitly acknowledge model assumptions:** Without this, differences in reconstructed networks may be mistakenly attributed to biological change rather than methodological artifacts.
- <sup>284</sup> 2. **Comparisons across studies must be standardised by model family:** Comparing a metaweb from one study to a realised network from another risks generating misleading conclusions about ecological differences across space or time.
- <sup>287</sup> 3. **Mechanistic paleo-specific models alone cannot capture the full spectrum of possible network archetypes:** Incorporating theory-driven models allows paleoecology to leverage decades of research on modern food-web structure and dynamics, broadening the range of ecological questions that can be asked.

<sup>291</sup> **3.5 Recommendations and best practices**

<sup>292</sup> To support consistent, transparent use of network reconstruction methods in paleoecology, we propose the <sup>293</sup> following guidelines (which could be highlighted as a boxed summary):

- <sup>294</sup> 1. **Start with the ecological question:**, then identify the network representation (feasible, realised, structural) best aligned with it.
- <sup>296</sup> 2. **Use multiple models whenever possible:** to evaluate whether key conclusions are model-dependent.
- <sup>297</sup> 3. **Report all assumptions explicitly:**, including body-size estimation methods, trait inference procedures, and parameter choices.
- <sup>299</sup> 4. **Perform sensitivity analyses:**, particularly for size-based models and connectance assumptions.
- <sup>300</sup> 5. **Avoid comparing networks built with different model families** unless explicitly testing model effects.

### 304 3.6 Concluding remarks

305 Our results reveal that network construction models differ far more than is commonly appreciated in  
306 paleoecology. Because food webs are increasingly used to infer ecological resilience, collapse dynamics, and  
307 responses to environmental change, it is crucial to recognise that these inferences are shaped as much by  
308 modelling decisions as by fossil data. By mapping model families to ecological questions and demonstrating  
309 how model assumptions influence inferred patterns, our study provides a framework for more transparent,  
310 robust, and question-driven reconstruction of paleo food webs.

[Figure 6 about here.]

## **312 References**

- 313 Allesina, S., Alonso, D., & Pascual, M. (2008). A general model for food web structure. *Science*, 320(5876),  
314 658–661. <https://doi.org/10.1126/science.1156269>

315 Bambach, R. K., Bush, A. M., & Erwin, D. H. (2007). Autecology and the Filling of Ecospace: Key Metazoan  
316 Radiations. *Palaeontology*, 50(1), 1–22. <https://doi.org/10.1111/j.1475-4983.2006.00611.x>

317 Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Dalla Riva, G. V., Fortin, M.-J., Gravel, D., Guimarães, P.  
318 R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2018). Analysing  
319 ecological networks of species interactions. *Biological Reviews*, 112540. <https://doi.org/10.1111/brv.12433>

320 Dillon, E. M., Pier, J. Q., Smith, J. A., Raja, N. B., Dimitrijević, D., Austin, E. L., Cybulski, J. D., De  
321 Entrambasaguas, J., Durham, S. R., Grether, C. M., Haldar, H. S., Kocáková, K., Lin, C.-H., Mazzini, I.,  
322 Mychajliw, A. M., Ollendorf, A. L., Pimiento, C., Regalado Fernández, O. R., Smith, I. E., & Dietl, G. P.  
323 (2022). What is conservation paleobiology? Tracking 20 years of research and development. *Frontiers in  
324 Ecology and Evolution*, 10. <https://doi.org/10.3389/fevo.2022.1031483>

325 Dunhill, A. M., Zarzyczny, K., Shaw, J. O., Atkinson, J. W., Little, C. T. S., & Beckerman, A. P. (2024).  
326 Extinction cascades, community collapse, and recovery across a Mesozoic hyperthermal event. *Nature  
327 Communications*, 15(1), 8599. <https://doi.org/10.1038/s41467-024-53000-2>

328 Erdős, P., & Rényi, A. (1959). On random graphs. i. *Publicationes Mathematicae Debrecen*, 6(3-4), 290–297.  
329 <https://doi.org/10.5486/pmd.1959.6.3-4.12>

330 Fricke, E. C., Hsieh, C., Middleton, O., Gorczynski, D., Cappello, C. D., Sanisidro, O., Rowan, J., Svenning,

- 331 J.-C., & Beaudrot, L. (2022). Collapse of terrestrial mammal food webs since the Late Pleistocene. *Science*,  
332 377(6609), 1008–1011. <https://doi.org/10.1126/science.abn4012>
- 333 Gupta, A., Furrer, R., & Petchey, O. L. (2022). Simultaneously estimating food web connectance and structure  
334 with uncertainty. *Ecology and Evolution*, 12(3), e8643. <https://doi.org/10.1002/ece3.8643>
- 335 Hao, X., Holyoak, M., Zhang, Z., & Yan, C. (2025). Global Projection of Terrestrial Vertebrate Food  
336 Webs Under Future Climate and Land-Use Changes. *Global Change Biology*, 31(2), e70061. <https://doi.org/10.1111/gcb.70061>
- 337 Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F., Edwards, F., Figueroa,  
338 D., Jacob, U., Jones, J. I., Lauridsen, R. B., Ledger, M. E., Lewis, H. M., Olesen, J. M., Veen, F. J. F.  
339 van, Warren, P. H., & Woodward, G. (2009). Ecological networks—beyond food webs. *The Journal of  
340 Animal Ecology*, 78(1), 253–269. <https://doi.org/10.1111/j.1365-2656.2008.01460.x>
- 341 Kiessling, W., Raja, N. B., Roden, V. J., Turvey, S. T., & Saupe, E. E. (2019). Addressing priority questions  
342 of conservation science with palaeontological data. *Philosophical Transactions of the Royal Society B:  
343 Biological Sciences*, 374(1788), 20190222. <https://doi.org/10.1098/rstb.2019.0222>
- 344 Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple  
345 building blocks of complex networks. *Science*, 298(5594), 824–827. [https://doi.org/10.1126/science.298.5594.824](https://doi.org/10.1126/science.298.<br/>346 5594.824)
- 347 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from  
348 proxies. *Trends in Ecology & Evolution*, 30(6), 347–356. <https://doi.org/10.1016/j.tree.2015.03.014>
- 349 Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2008). Size, foraging, and food web structure.  
350 *Proceedings of the National Academy of Sciences*, 105(11), 4191–4196. [https://doi.org/10.1073/pnas.0710672105](https://doi.org/10.1073/pnas.<br/>351 0710672105)
- 352 Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in  
353 Ecology and Evolution*, 14(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- 354 Poisot, T., Canard, E., Mouquet, N., & Hochberg, M. E. (2012). A comparative study of ecological  
355 specialization estimators. *Methods in Ecology and Evolution*, 3(3), 537–544. [https://doi.org/10.1111/j.2041-210x.2011.00174.x](https://doi.org/10.1111/j.2041-<br/>356 210x.2011.00174.x)
- 357 Rohr, R., Scherer, H., Kehrli, P., Mazza, C., & Bersier, L.-F. (2010). Modeling food webs: Exploring  
358 unexplained structure using latent traits. *The American Naturalist*, 176(2), 170–177. [https://doi.org/10.1086/653667](https://doi.org/10.<br/>359 1086/653667)
- 360 Roopnarine, P. D. (2006). Extinction cascades and catastrophe in ancient food webs. *Paleobiology*, 32(1),  
361 1–19. <https://www.jstor.org/stable/4096814>
- 362 Schneider, F. D., Brose, U., Rall, B. C., & Guill, C. (2016). Animal diversity and ecosystem functioning in

- 364 dynamic food webs. *Nature Communications*, 7(1), 12718. <https://doi.org/10.1038/ncomms12718>
- 365 Shaw, J. O., Dunhill, A. M., Beckerman, A. P., Dunne, J. A., & Hull, P. M. (2024). *A framework for*  
366 *reconstructing ancient food webs using functional trait data* (p. 2024.01.30.578036). bioRxiv. <https://doi.org/10.1101/2024.01.30.578036>
- 367
- 368 Stouffer, D. B., Camacho, J., Jiang, W., & Nunes Amaral, L. A. (2007). Evidence for the existence of a robust  
369 pattern of prey selection in food webs. *Proceedings of the Royal Society B: Biological Sciences*, 274(1621),  
370 1931–1940. <https://doi.org/10.1098/rspb.2007.0571>
- 371 Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz,  
372 N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap  
373 towards predicting species interaction networks (across space and time). *Philosophical Transactions of the*  
374 *Royal Society B: Biological Sciences*, 376(1837), 20210063. <https://doi.org/10.1098/rstb.2021.0063>
- 375 Thuiller, W., Calderón-Sanou, I., Chalmandrier, L., Gaüzère, P., O'Connor, L. M. J., Ohlmann, M., Poggiato,  
376 G., & Münkemüller, T. (2024). Navigating the integration of biotic interactions in biogeography. *Journal*  
377 *of Biogeography*, 51(4), 550–559. <https://doi.org/10.1111/jbi.14734>
- 378 Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774), 180–183.  
379 <https://doi.org/10.1038/35004572>
- 380 Williams, R. J., & Martinez, N. D. (2004). Stabilization of chaotic and non-permanent food-web dynamics.  
381 *The European Physical Journal B - Condensed Matter*, 38(2), 297–303. <https://doi.org/10.1140/epjb/e2004-00122-1>
- 382
- 383 Williams, R. J., & Martinez, N. D. (2008). Success and its limits among structural models of complex food  
384 webs. *The Journal of Animal Ecology*, 77(3), 512–519. <https://doi.org/10.1111/j.1365-2656.2008.01362.x>
- 385 Yeakel, J. D., Pires, M. M., Rudolf, L., Dominy, N. J., Koch, P. L., Guimarães, P. R., & Gross, T.  
386 (2014). Collapse of an ecological network in ancient egypt. *PNAS*, 111(40), 14472–14477. <https://doi.org/10.1073/pnas.1408471111>
- 387

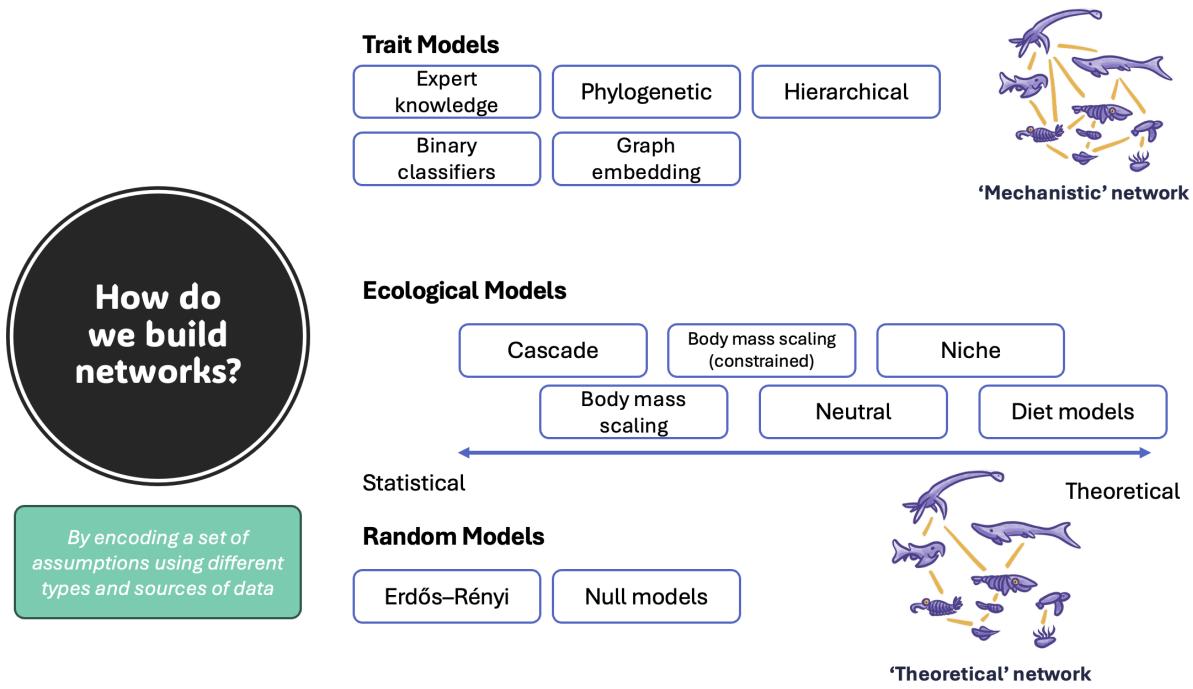


Figure 1: This obviously needs work but a variation on this to try and articulate the different approaches and broadly how they may differ.

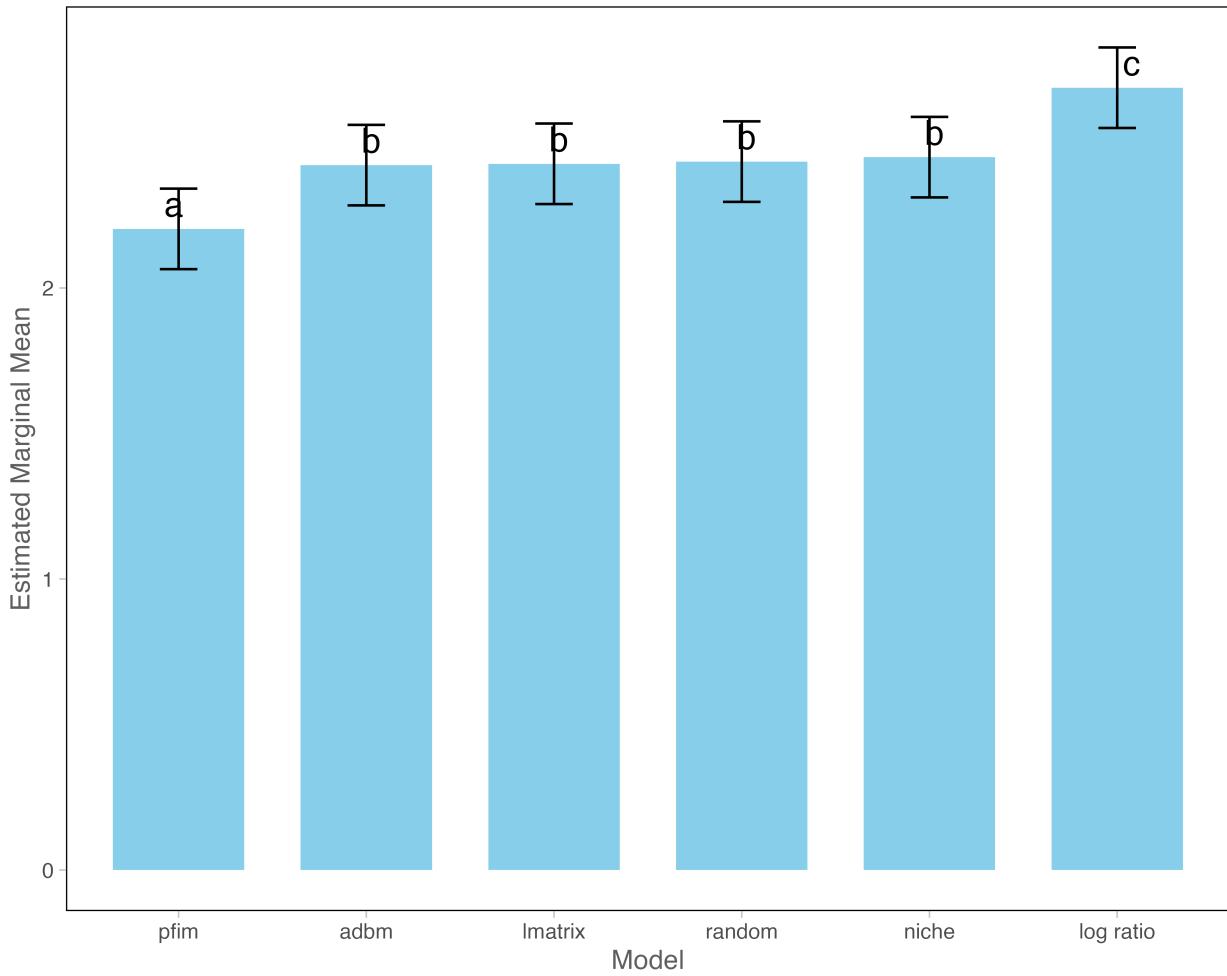


Figure 2: Estimated marginal means (EMMs) of ecological network metrics across six model types with 95% confidence intervals. Bars represent the predicted values for each model, and error bars indicate the 95% confidence limits. Letters above each bar denote Tukey-adjusted pairwise significance: models sharing the same letter are not significantly different, while models with different letters are significantly different ( $p < 0.05$ ). The plot reveals three tiers of model performance, with pfim consistently higher, log ratio, niche, and random at intermediate levels, and adbM and lmatrix lower, consistent with the MANOVA and post-hoc analyses.

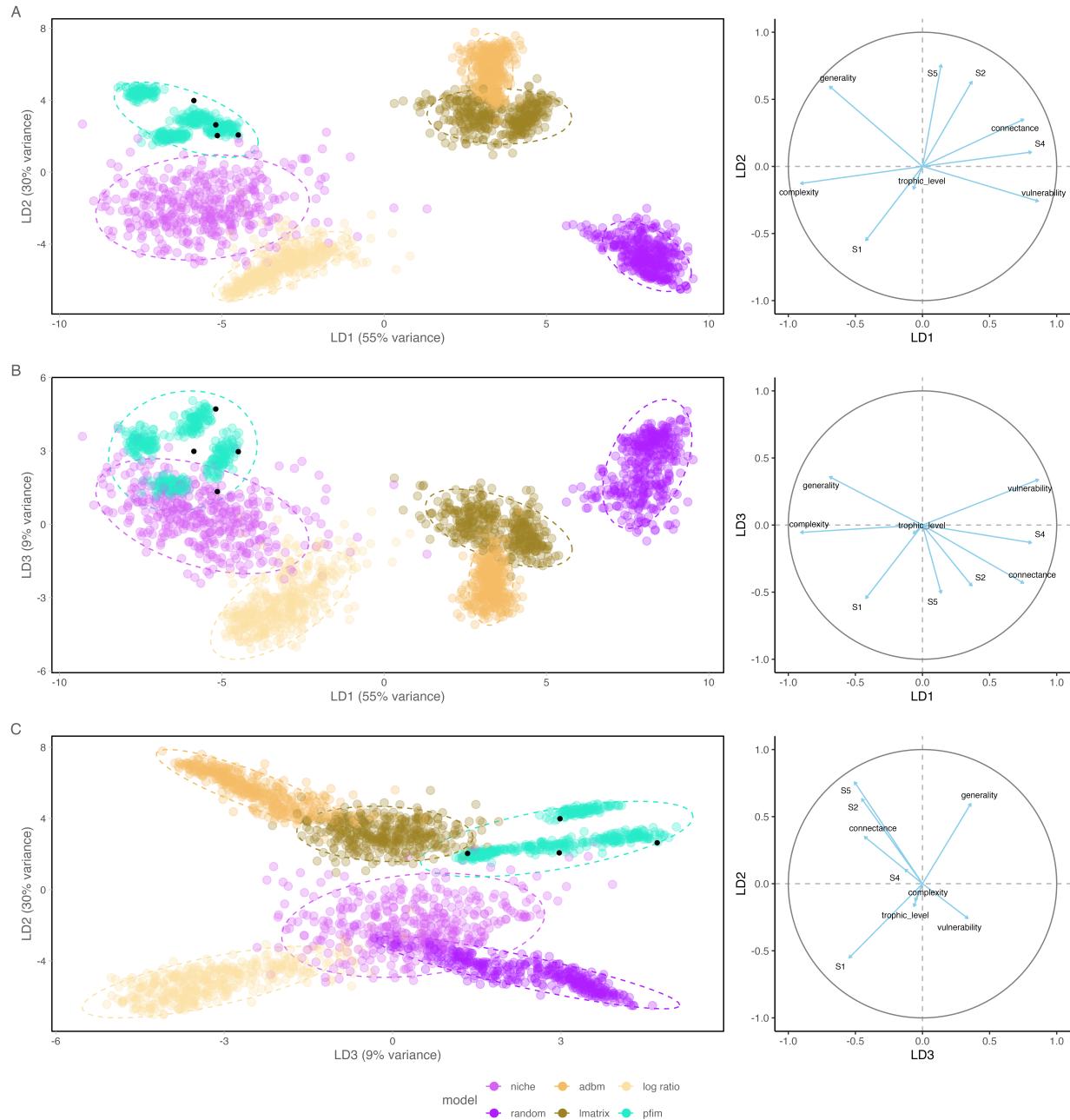


Figure 3: Linear discriminant analysis (LDA) of ecological network metrics for six model types. Each point represents a replicate, and ellipses indicate 95% confidence regions for each model. The second column represents the correlation of the various network metrics with the respective LDA axes.

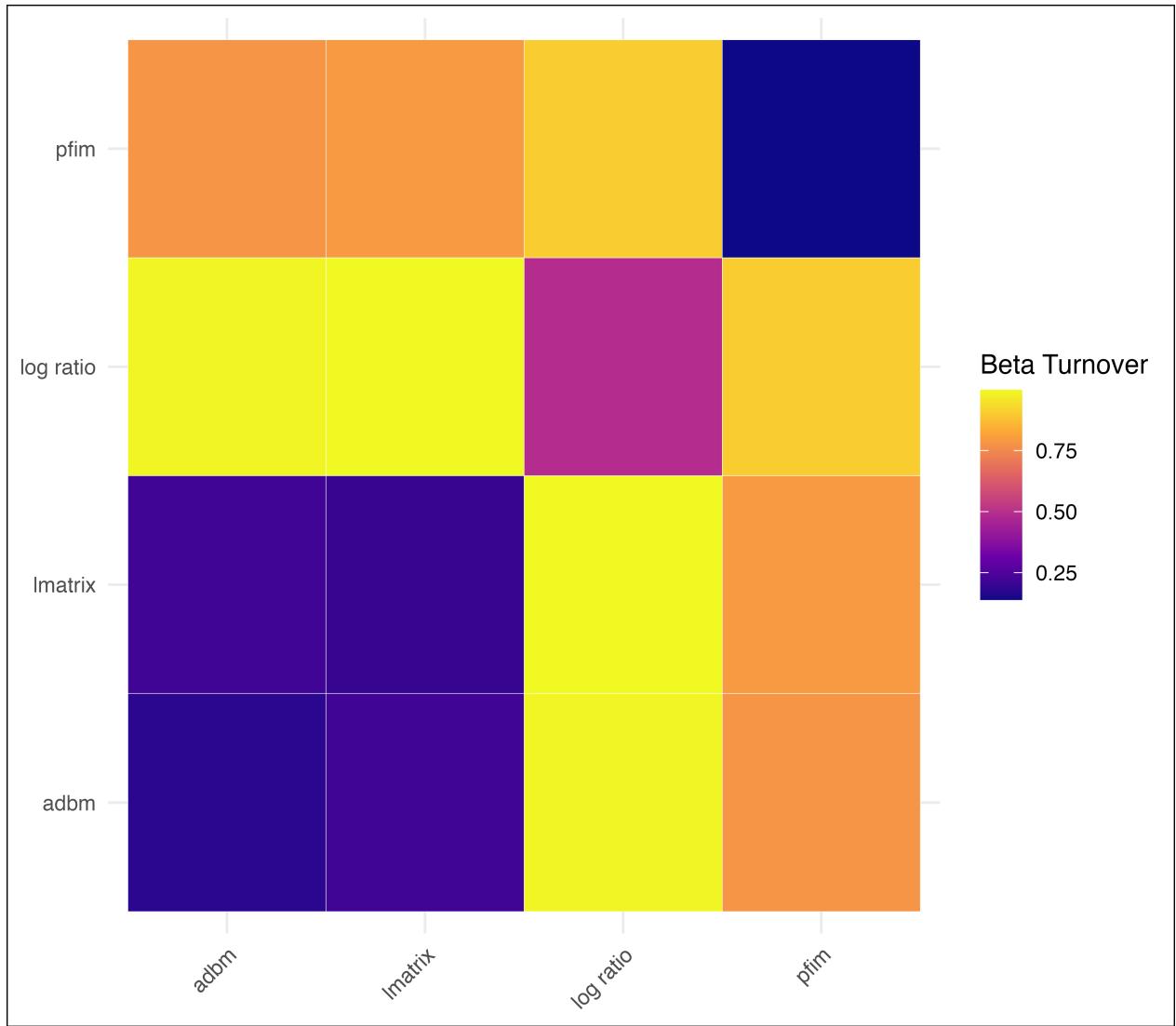
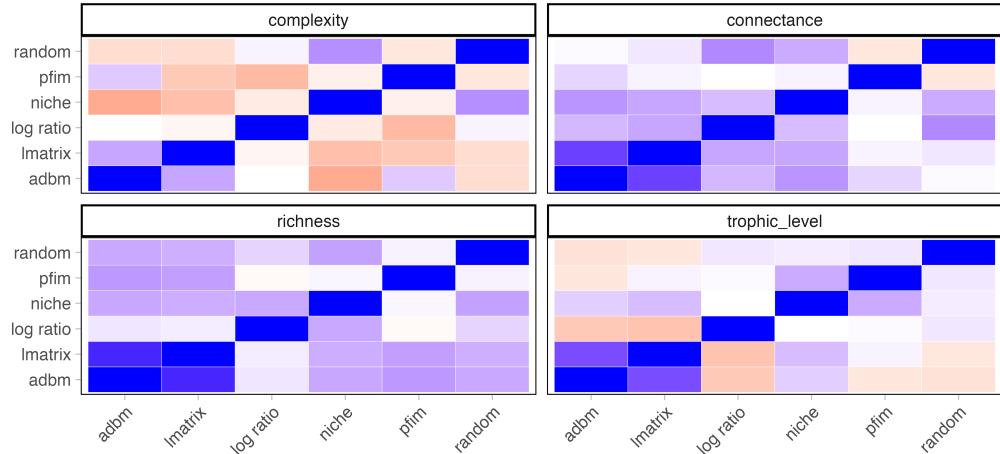
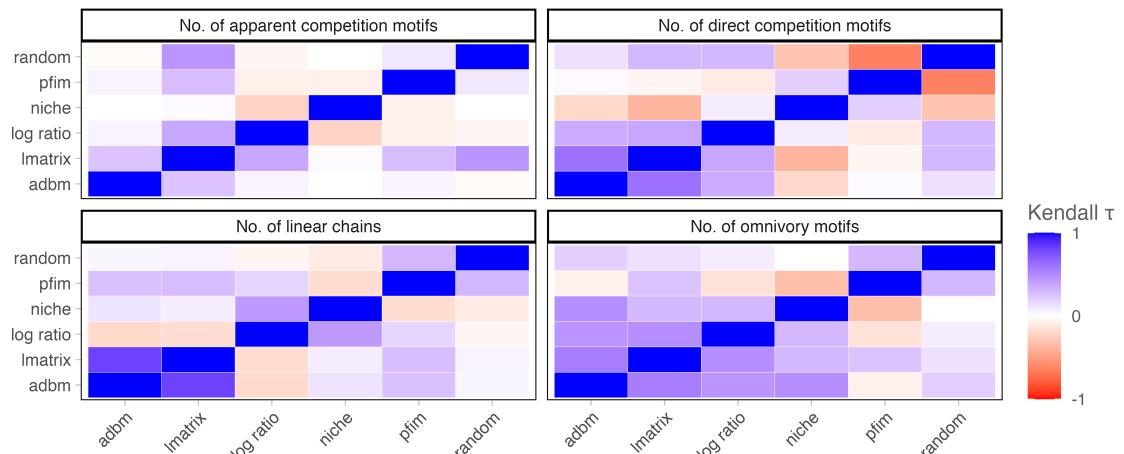


Figure 4: Pairwise beta turnover in species interactions among four ecological network models (adbm, lmatrix, log-ratio, and pfim). Each cell represents the mean turnover value between a pair of models, with warmer colors indicating greater dissimilarity in inferred interactions. The diagonal is omitted. High turnover values (yellow) indicate strong disagreement in network structure between models, whereas lower values (blue–purple) indicate greater similarity.

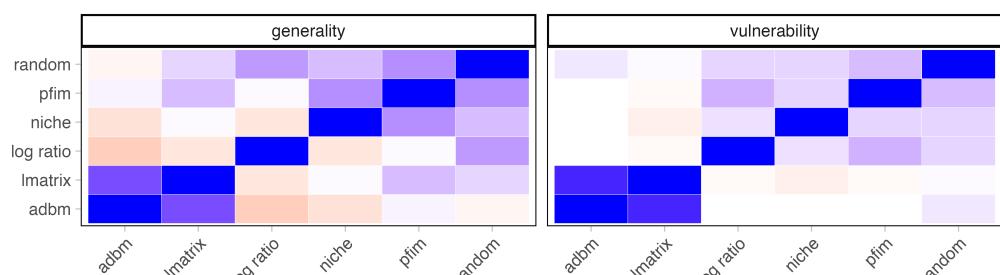
### Macro



### Meso



### Micro



### TSS

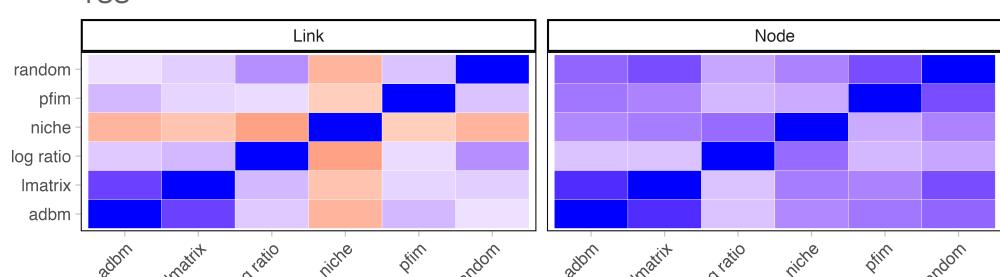


Figure 5: Heatmaps showing pairwise Kendall rank correlation coefficients ( $\tau$ ) between models for each network metric. Each panel corresponds to a different metric and displays the degree of agreement in extinction-scenario rankings across models based on mean absolute differences (MAD) between observed and predicted network values. Positive values (blue) indicate concordant rankings between models, whereas negative values (red) indicate opposing rankings. Warmer colours approaching zero represent little or no agreement. Panels

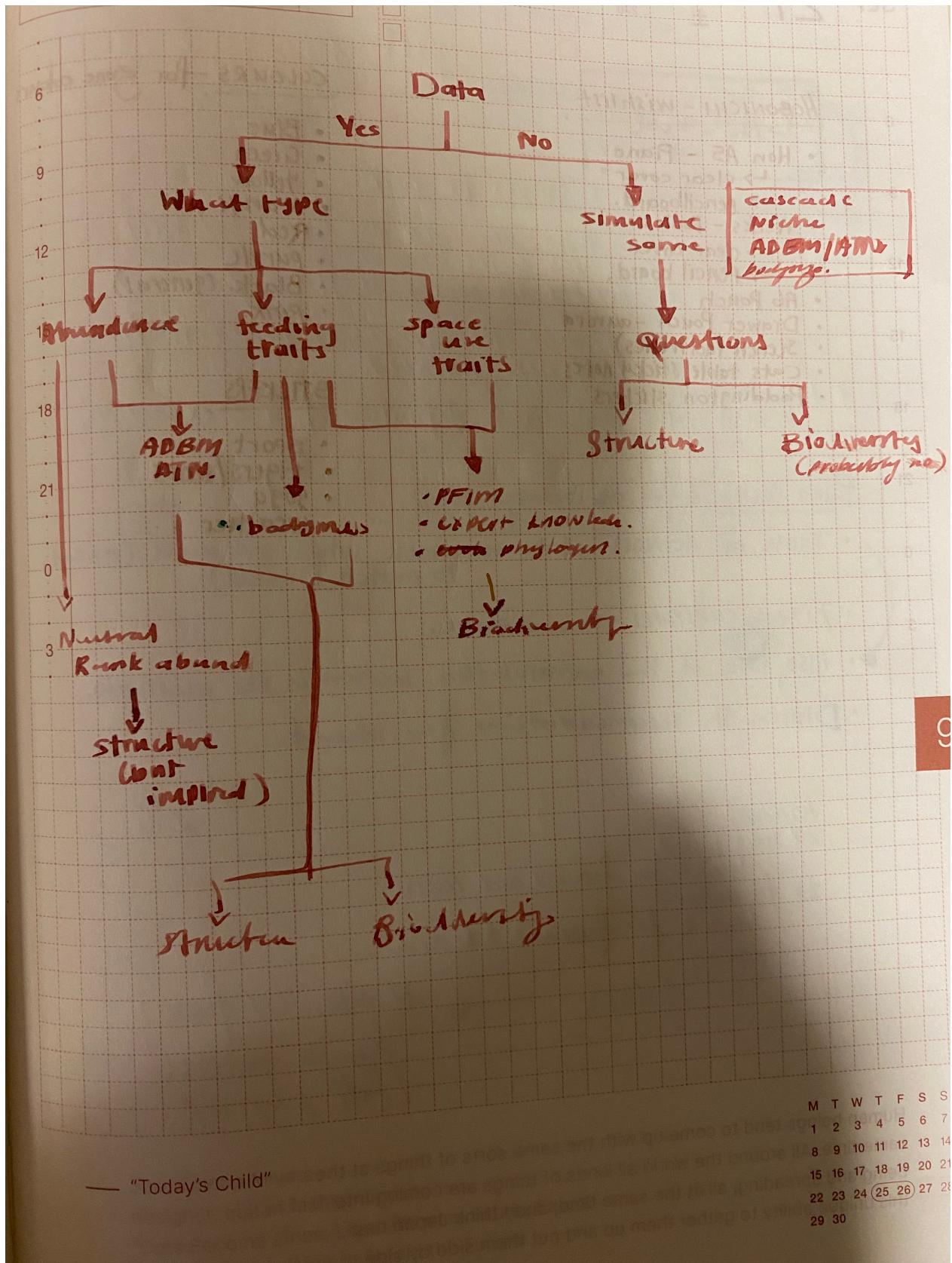


Figure 6: TODO.