

# **Reconstructing deep-time food webs: model assumptions drive paleoecological inference**

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 vary widely in their assumptions, mechanisms, and data requirements. Here, we evaluate which network construction approaches are most suitable for paleo-food-web reconstruction given the constraints of the fossil record, and we assess how model choice shapes the networks we infer. Using the Toarcian Oceanic Anoxic Event (Early Jurassic, ~183 Ma) as a case study, we compare six modelling approaches encompassing mechanistic, structural, and theory-based methods. These models yield strikingly different network structures and pairwise interactions, and that these discrepancies propagate into ecological inference, including conclusions about extinction dynamics. Our results highlight the importance of the need to align 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 species interacted with each  
2 other in past ecosystems (*e.g.*, Dunne et al. (2008); Dunne et al. (2014)) and how ecological communities  
3 responded to past environmental change (Dillon et al., 2022; Kiessling et al., 2019). In modern ecosystems,  
4 species interactions and the networks that they form have become central to studying biodiversity, energy  
5 flow, and community stability (*e.g.*, Thuiller et al. (2024)). Network prediction tools derived from modern  
6 ecology have increasingly been applied by paleoecologists to reconstruct ancient food webs to gain insight into  
7 how the biosphere responds to major environmental transitions (*e.g.*, Dunhill et al. (2024); Hao et al. (2025);  
8 Yeakel et al. (2014)). However, we are faced by the limitation that interactions cannot be directly observed in  
9 the fossil record (with the exception of rare instances *e.g.*, Jenny et al. (2019); Vullo (2011)) and as a result,  
10 the reconstruction of paleo food webs depends on models that allow us to infer feeding relationships from  
11 preserved traits, analogies to modern taxa, or ecological theory. While numerous models exist for inferring  
12 interactions (see Morales-Castilla et al. (2015); Pichler & Hartig (2023); Strydom et al. (2021); Allesina et  
13 al. (2008) for broader reviews), only a subset can reliably be applied in paleo contexts, where data on traits,  
14 abundances, and community composition are inherently incomplete and biased.

15 The growing interest in paleo food webs has outpaced a clear discussion of *which* construction methods are  
16 suitable for which purposes. Different models generate different kinds of networks (*e.g.*, feasible, realised, or  
17 purely structural) and these differences can fundamentally alter ecological interpretations. In this study, we  
18 evaluate a suite of methods that can be feasibly applied to paleo communities and explore how their underlying  
19 assumptions shape both network structure and ecological inference. Specifically we focus on identifying a suite  
20 of models that are appropriate for use with paleo data that can feasibly be constructed within the limitations  
21 that are imposed by fossil data while still intersecting across the different network types. Here we use the  
22 data from Dunhill et al. (2024) to reconstruct communities across the Toarcian extinction event using a suite  
23 of different modelling approaches. We ask (i) how do different models vary in the network structures and  
24 pairwise interactions they recover, and (ii) to what extent does model choice influence ecological inference,  
25 including which factors are interpreted as the dominant drivers of extinction across a major extinction event..

## 26 1 Constructing paleo webs

### 27 1.1 Challenges specific to building paleo networks

28 Reconstructing paleo food webs presents challenges that differ from those encountered in modern systems.  
29 First, the fossil record provides an incomplete and selectively preserved subset of the original community.  
30 Preservation biases (driven by habitat, skeletal composition, and sedimentary environment) mean that

some trophic groups are over-represented (e.g., hard-shelled organisms), while others (e.g., soft-bodied taxa, plankton) are systematically under-sampled. This directly constrains the kinds of models that can be applied, because models requiring complete assemblages or accurate guild representation will perform poorly when preservation is uneven (**REF**). Furthermore, there is inherent uncertainty about the true community compositions. Fossil assemblages may represent time-averaged “death assemblage” accumulations, transported material, or mixed habitats, rather than true representations of “life assemblages” of organisms that interacted with one another. As a result, any paleontological assemblage is best interpreted as a set of taxa that could have interacted, rather than a snapshot of a specific subset of interacting organisms. Moreover, certain traits cannot be reliably estimated, such as body size due to disarticulation (e.g., echinoderms) or partial preservation (e.g., sharks), or species abundances due to time-averaging, transportation, or preservation. These shortcomings limit the potential models that can be applied to reconstruct trophic networks from fossil assemblages. Additionally, many extinct taxa have ambiguous trait states, especially regarding diet and behaviour. Even when functional morphology is preserved, ecological behaviour is seldom directly evident. Such uncertainty propagates differently across model families: mechanistic models (e.g., **xxx**, **Figure 1**) tend to accommodate broad trait assignments, whereas theory-driven models (e.g., **xxx**) are more sensitive to uncertainty in body size, foraging mode, or feeding constraints. These limitations do not render reconstruction impossible but highlight the importance of choosing a model whose assumptions match the type of ecological inference being attempted.

## 1.2 Understanding the approaches to network construction

Network construction approaches can be broadly grouped into three methodological and conceptual approaches **Figure 1**. The first are mechanistic models which evaluate whether an interaction is *feasibly possible*. These models typically use trait-based rules (e.g., feeding mode, body size, or functional morphology) or evolutionary relationships to determine whether a species *could* consume another, by extension this also allows us to identify forbidden links (*i.e.*, interactions that are mechanistically incompatible Jordano (2016)). Mechanistic approaches tend to produce metawebs - the full set of all plausible interactions given biological constraints. Theory-driven models embed assumptions from ecological theory (such as niche theory or foraging ecology) to generate realised interactions and networks. Structural models are similar to theory-driven models, with the exception that these models are species agnostic and as such can only be used to make inferences about network structure (**i.e.**, **xxx**) rather than inferring realistic pairwise predator-prey links. Both theory-driven and structural models aim to reproduce characteristic patterns observed in modern food webs, such as intervality, trophic hierarchies, or body-size-scaled feeding ranges. They do not necessarily require detailed

62 trait information and instead rely on ecological rules or statistical distributions consistent with empirical food  
63 webs.

64 [Figure 1 about here.]

65 Most existing paleo-specific approaches fall within the mechanistic space (e.g., Shaw et al. (2024); Roopnarine  
66 (2006); Fricke et al. (2022)). While these are well-suited for reconstructing feasible interactions, they represent  
67 only a subset of the broader space of network construction methods. Incorporating theory-driven models allows  
68 paleoecologists to explore realised interaction structures and address a wider suite of ecological questions—  
69 provided their assumptions are compatible with the limitations of fossil data. Here we present a range of  
70 models Table 1 that carry specific assumptions and data requirements. For instance, allometric models depend  
71 on quantitative body-size estimates, which must be inferred from size classes or functional morphology in  
72 the fossil record. Structural models such as the niche model require only richness and connectance, but their  
73 species-agnostic nature limits their usefulness for trait-based or diet-specific questions. Mechanistic models  
74 rely on accurate assignment of feeding traits, which may be uncertain for extinct taxa but are often more  
75 tractable than estimating abundances or interaction strengths. Understanding how these limitations intersect  
76 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)

Model family	Assumptions	Data needs	'Limitation'	Network type	Key reference
allometric diet breadth model (ADBM)	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)
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)

Model family	Assumptions	Data needs	‘Limitation’	Network type	Key reference
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 from models (although this is a way smaller issue)	mechanistic web	Shaw et al. (2024)
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)

77 **2 Case study: Toarcian mass extinction event**

78 **2.1 Dataset overview**

79 **2.1.1 Species occurrence**

80 We used fossil occurrence data spanning the upper Pliensbachian (~185 Ma) to the upper Toarcian (~175  
81 Ma) of the Cleveland Basin, following Dunhill et al. (2024). The dataset comprises four paleo-communities  
82 representing the pre-extinction, post-extinction, early recovery, and late recovery intervals of the Toarcian  
83 Oceanic Anoxic Event. Each assemblage was treated as a community of potentially interacting taxa.

84 **2.1.2 Defining modes of life (traits)**

85 We used the modes of life (traits) as identified in Dunhill et al. (2024), who defined four traits: motility (fast,  
86 slow, facultative, non-motile), tiering (pelagic, erect, surficial, semi-infaunal, shallow infaunal, deep infaunal),  
87 feeding (predator, suspension feeder, deposit feeder, mining, grazer), and size: gigantic (>500 mm), very large  
88 (>300–500 mm), large (>100–300 mm), medium (>50–100 mm), small (>10–50 mm), tiny (10 mm), for each  
89 fossil taxon based on the ecological traits defined in the Bambach ecospace model (Bambach et al., 2007).

90 **2.1.3 Constructing networks**

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

103 **2.2 Models capture different network structure but in unexpected ways**

104 When quantifying network structure, we are essentially asking how interactions are distributed across taxa and  
105 how these patterns scale from individual nodes to the whole community. Structural metrics are informative  
106 because they reflect underlying ecological processes: how energy flows through trophic levels, how disturbances  
107 propagate, where redundancy or fragility exists, and how species specialise or generalise in their diets. To  
108 capture these different facets, we evaluated a suite of macro-, meso-, and micro-scale metrics Table 2, ranging  
109 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)

Metric	Definition	Scale	Reference (for maths), can make footnotes probs
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	
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)

<sup>110</sup> Despite being supplied with the same taxon pools, the different models generated networks with systematically different structural signatures. A MANOVA (using the metrics listed in Table 2 as the dependent variables and model as the grouping variable) revealed a strong multivariate effect of model type on network structure

<sup>113</sup> (Pillai's Trace = 3.89,  $p < 0.001$ ), indicating that each modelling approach produces a distinct 'structural  
<sup>114</sup> fingerprint'. Follow-up ANOVAs confirmed that model choice had substantial effects on every metric we  
<sup>115</sup> examined, with effect sizes typically exceeding 0.80. The only exception was maximum trophic level ( $\eta^2 =$   
<sup>116</sup> 0.19), suggesting some convergence in vertical structure even when structure differs widely.

<sup>117</sup> Post-hoc comparisons revealed three broad clusters Figure 2. PFIM consistently yielded the most connected,  
<sup>118</sup> dense networks, reflecting its mechanistic emphasis on trait-based feasibility. Niche, random, and body-size  
<sup>119</sup> ratio models formed an intermediate group, producing networks with moderate connectance and motif  
<sup>120</sup> frequencies. ADBM and l-matrix formed a tight cluster characterised by constrained feeding ranges, reflecting  
<sup>121</sup> their shared basis in energetic and allometric theory. Although these groupings broadly align with *a priori*  
<sup>122</sup> expectations about model families, several patterns emerged that were less intuitive. Most notably, the  
<sup>123</sup> body-size ratio model, which is theoretically grounded but, aligned more closely with structural models than  
<sup>124</sup> with the fully theoretic ones. This suggests that even slight differences in how body-size constraints are  
<sup>125</sup> implemented can shift a model's position within the network-structure landscape. A Linear Discriminant  
<sup>126</sup> Analysis (LDA) further illustrated the distinctiveness of the model families Figure 3. Classification accuracy of  
<sup>127</sup> the LDA was 85%, demonstrating that the combination of dependent variables effectively discriminates among  
<sup>128</sup> model types, indicating the strong imprint that model assumptions leave on inferred ecological patterns. The  
<sup>129</sup> PFIM model is strongly separated, while ADBM and l-matrix networks cluster closely together. The niche  
<sup>130</sup> and random models occupy intermediate positions. Importantly these results indicate that model choice is the  
<sup>131</sup> dominant driver of inferred network structure, often overwhelming the ecological signal embedded within the  
<sup>132</sup> taxon pool itself. Models designed to reproduce feasible interactions and those designed to generate realised  
<sup>133</sup> niche structure occupy fundamentally different portions of 'network space', even when operating on identical  
<sup>134</sup> taxa. Such structural divergences could have direct implications for ecological inference, particularly when  
<sup>135</sup> comparing networks across time or when using network metrics to infer processes such as community stability,  
<sup>136</sup> trophic organisation, or susceptibility to cascading extinctions.

<sup>137</sup> [Figure 2 about here.]

<sup>138</sup> [Figure 3 about here.]

### <sup>139</sup> 2.3 Some networks don't share any interactions and some share a lot

<sup>140</sup> Beyond differences in global structure, researchers are often interested in specific ecological relationships, *e.g.*,  
<sup>141</sup> who eats who, which species share predators, and how trophic roles change across communities or time. For  
<sup>142</sup> these types of questions, it is essential to understand how models differ at the level of pairwise interactions.  
<sup>143</sup> To quantify this, we measured interaction turnover between networks to allow us to assess the degree to

<sup>144</sup> which two models predict the same or different links for the same set of species (see Poisot et al., 2012 for  
<sup>145</sup> methods). This is analogous to -diversity but applied to links rather than species. Specifically we only  
<sup>146</sup> looked at the dissimilarity of interaction between shared species. Even when supplied with identical species  
<sup>147</sup> pools, models varied dramatically in the interactions they inferred Figure 4. Some pairs of models showed  
<sup>148</sup> substantial agreement, whereas others shared almost no interactions at all. Note here that we did not include  
<sup>149</sup> the Random or Niche models as these networks are species agnostic and as such are not designed for inferring  
<sup>150</sup> species-specific pairwise links.

<sup>151</sup> The body-size ratio model had consistently high turnover relative to all others, indicating that it inferred  
<sup>152</sup> diets that were largely distinct. This reflects the strong constraints imposed by its logit-based linking rule,  
<sup>153</sup> which sharply restricts prey to a narrow ‘optimal’ size range. Small differences in body-size estimates or  
<sup>154</sup> functional groupings therefore lead to disproportionately large changes in inferred interactions. In contrast, the  
<sup>155</sup> ADBM and L-matrix showed low turnover between each other, reflecting their shared theoretical foundations  
<sup>156</sup> (operationalising foraging decisions and energetic constraints using similar allometric principles). As a result,  
<sup>157</sup> they tend to produce similar pairwise interactions even when implemented independently. The PFIM exhibited  
<sup>158</sup> intermediate turnover, sharing more interactions with size-based theoretical models than with the body-size  
<sup>159</sup> ratio model. This makes sense: although PFIM uses categorical traits and hierarchical feeding rules rather  
<sup>160</sup> than quantitative foraging theory, these constraints will still produce broadly similar trophic groupings.

<sup>161</sup> Taken together, these patterns show that pairwise interactions differ far more across models than global metrics  
<sup>162</sup> alone might suggest. Two models with superficially similar connectance or trophic structure may nonetheless  
<sup>163</sup> infer completely different diets for individual taxa. This has significant implications for any question focused on  
<sup>164</sup> species-level ecology, including predator-prey specialisation, trophic niche breadth, or the identity of keystone  
<sup>165</sup> consumers. These findings reinforce the importance of selecting a model whose assumptions align with the  
<sup>166</sup> intended inference. If the goal is to explore the full set of possible interactions a species could have had, a  
<sup>167</sup> mechanistic model such as PFIM is appropriate. If instead the goal is to infer the likely realised interactions  
<sup>168</sup> or energy pathways, models grounded in allometric foraging theory (ADBM, L-matrix) will provide more  
<sup>169</sup> ecologically coherent results. Conversely, models like the body-size ratio may be too restrictive or idiosyncratic  
<sup>170</sup> for diet-based questions because they force interactions into narrow, trait-determined windows.

<sup>171</sup> [Figure 4 about here.]

## <sup>172</sup> 2.4 Model choice changes the narrative

<sup>173</sup> The structural and interaction-level differences documented above raise a central question: do different  
<sup>174</sup> models also lead to different interpretations of extinction selectivity inference? In other words, does model

choice merely affect the architecture of the reconstructed networks, or does it shape the actual stories we tell about how communities collapsed and recovered during the Toarcian extinction event? Using the pre-extinction networks as starting points, we simulated species losses under a suite of ecologically plausible extinction scenarios, including trait-based removals (e.g., body size, motility), network-position removals (e.g., vulnerability, generality), and random extinctions. In each case, we allowed for cascading secondary extinctions. For each model and scenario, we then measured how closely the simulated post-extinction network resembled the real fossil community.

#### 2.4.1 Inferred extinction drivers

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

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

For each network metric, we treated each model as an independent evaluator of scenario performance. MAD and TSS values were converted to within-model rankings, with rank 1 assigned to the scenario with the smallest MAD (*i.e.*, the closest match to the empirical value) or highest TSS score. Ranking was performed independently for each combination of model and network metric to avoid assumptions about comparability across metrics. To evaluate whether different models produced consistent rankings of extinction scenarios, we quantified rank correlation among models separately for each network metric. Agreement among model rankings was assessed using Kendall's rank correlation coefficient (Kendall's  $\tau$ ), which measures the degree of

204 agreement between two ordinal rankings. Kendall's  $\tau$  was selected because it is robust for small sample sizes,  
205 handles tied ranks appropriately, and provides a direct measure of the probability that model pairs agree or  
206 disagree on the relative ordering of scenarios. Kendall's  $\tau$  ranges from  $-1$  to  $+1$ , where  $+1$  indicates perfect  
207 agreement between rankings,  $0$  reflects no relationship, and  $-1$  represents complete disagreement such that  
208 one ranking is the exact reverse of the other.

209 When we look at Kendall's  $\tau$  for the MAD across network structure and models Figure 5 we see that generally  
210 there is a positive correlation between the different models. This implies that different models are often  
211 recovering a similar ranking of extinction mechanisms (as in the ‘signal’ as to which extinction mechanisms  
212 may be the most plausible are the same). Although there is not a strong agreement between models, as values  
213 tend to be low, it is promising to observe that it is not often that we have a completely different ranking of  
214 extinction mechanisms, with the exception of complexity and the number of direct competition motifs. When  
215 looking at the macro-level network properties the random model often showcases a disagreement in terms of  
216 the MAD. This is unsurprising as we expect random networks to produce networks that are not ecologically  
217 sound and thus will not behave as one may expect (Ings et al., 2009). Interestingly we once again see the  
218 strong similarity between the L-matrix and the ADBM (have a high Kendall's  $\tau$ ), meaning that they recover a  
219 similar ranking of extinction mechanisms, this is unsurprising given that we know these networks tend to  
220 recover a similar structure Figure 2. Broadly when we look at the behaviour of the different model families  
221 (with the exception of the Random model) we see that they recover similar structural signals with regards to  
222 the mechanisms potentially driving extinctions.

223 [Figure 5 about here.]

224 When looking at the node-level TSS scores (Figure 5, TSS, panel 2) we see that, in general, the signal of  
225 the extinction mechanism is maintained across the different models. However as many of the extinction  
226 mechanisms are determined by the traits of the node it is not surprising that we see a similar signal as the  
227 taxa are being removed in the exact same order. Link-level TSS scores (Figure 5, TSS, panel 1) do not show  
228 the same within extinction mechanism ranking/signal. We see that the Random and PFIM models have high  
229 TSS scores (i.e., have a ‘good fit’), however in the case of the PFIM this is to be expected as the links are  
230 deterministic and so if you have the same two taxon pools you will recover the same links. The ‘stochastic’  
231 element of the theoretical models (ADBM, l-matrix, and bodymass-ratio) means that they create a degree of  
232 noise at the link-level and thus they are probably inappropriate to use for the type of extinction mechanism  
233 question we are asking here - specifically does the real and the simulated network look the same. Link-level  
234 TSS is perhaps also not an appropriate approach to determine the ‘best fit’ extinction mechanism if used in  
235 isolation and we advocate that the node-level TSS score (or alternatively some measure of  $\beta$  diversity) is used.

236 Finally, if we were to focus only on node level TSS, we do not observe any strong differences between the  
237 models, and it suggests that node-level driven (topological) extinction processes are insensitive to model type.  
238 Broadly inferred extinction mechanisms were relatively robust across models. This is probably in part because  
239 species are removed in the same order, node-level outcomes (which taxa survive) tend to agree. However,  
240 at the link level, where secondary cascades depend sensitively on inferred interactions, models often showed  
241 limited agreement. PFIM produced consistent link-level outcomes due to its deterministic rules, whereas  
242 theory-driven models (ADBM, L-matrix, body-size ratio) generated more variable trajectories due to stochastic  
243 link assignment. As a result, different models sometimes reconstructed different pathways of collapse, inferred  
244 different trophic groups as being the most affected, or ranked extinction scenarios differently. In other words,  
245 while the high-level narrative (e.g., traits matter) is stable, the fine-grained story of ecosystem disruption  
246 (who lost interactions first, how cascades unfolded, which species acted as bottlenecks) changes depending on  
247 the chosen model. Thus, the Toarcian extinction looks subtly but meaningfully different through the lens  
248 of each modelling framework. Model choice, therefore, must be treated as a core component of ecological  
249 inference, not a neutral preprocessing step. Some broad signals are robust—especially those driven by species  
250 traits—but many of the finer details that paleoecologists care about, such as trophic cascading pathways,  
251 keystone taxa, or the ordering of collapse, depend strongly on the chosen model. Researchers must therefore  
252 treat model choice not as a technical detail but as a central component of ecological inference.

### 253 3 Model Choice as an Ecological Inference Decision

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

#### 261 3.1 What our results demonstrate about model families

262 Across every structural metric we measured, model identity explained the majority of variation. PFIM-produced  
263 networks were consistently the most connected, while the ADBM and l-matrix produced sparser networks  
264 with tighter energetic constraints. Structural models (niche and random) fell between these extremes. These

265 clusters reflect the conceptual differences between feasible, realised, and structural network representations.  
266 Importantly, these differences were not superficial. Pairwise interaction turnover revealed that models often  
267 disagreed not just on the number of links but on who interacts with whom. Such discrepancies directly affect  
268 diet-based ecological questions and interpretations of trophic roles.

### 269 3.2 Consequences for interpreting paleoecological dynamics

270 The downstream effects of model choice were most evident when evaluating extinction dynamics. Models often  
271 recovered comparable high-level patterns (e.g., trait-driven extinction signals remained robust across models),  
272 but the specific ranking of extinction mechanisms varied. Node-level patterns were relatively stable because  
273 extinction scenarios based on intrinsic traits remove species in the same order regardless of network structure.  
274 However, link-level outcomes—particularly those concerning secondary extinctions, network collapse, and  
275 recovery trajectories—were highly model-dependent.

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

### 280 3.3 Aligning ecological questions with model choice

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

284 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</i> <i>realised</i> ?	Theory-driven (ADBM, l-matrix)	Encodes foraging ecology and energetic constraints; provides ecologically plausible ‘realised’ networks.

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

### **285 3.4 Implications for the use of food webs in paleoecology**

**286 Our analysis underscores three important implications for the broader field:**

- 287 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.
- 290 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.
- 293 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.

### **297 3.5 Recommendations and best practices**

**298 To support consistent, transparent use of network reconstruction methods in paleoecology, we propose the**  
**299 following guidelines (which could be highlighted as a boxed summary):**

- 300 1. Start with the ecological question:** then identify the network representation (feasible, realised, structural) best aligned with it.
- 302 2. Use multiple models whenever possible:** to evaluate whether key conclusions are model-dependent.

- 303 3. **Report all model assumptions explicitly and transparently:** including body-size estimation  
304 methods, trait inference procedures, and parameter choices.

305 4. **Perform sensitivity analyses:** particularly for size-based models and connectance assumptions.

306 5. **Avoid comparing networks built with different model families:** unless explicitly testing model  
307 effects.

308 6. **Distinguish between node-level and link-level inference:** as they respond differently to model  
309 choice.

### **3.6 Concluding remarks**

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

[Figure 6 about here.]

318 References

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

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

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

326 Dillon, E. M., Pier, J. Q., Smith, J. A., Raja, N. B., Dimitrijević, D., Austin, E. L., Cybulski, J. D., De  
327 Entrambasaguas, J., Durham, S. R., Grether, C. M., Haldar, H. S., Kocáková, K., Lin, C.-H., Mazzini, I.,  
328 Mychajliw, A. M., Ollendorf, A. L., Pimiento, C., Regalado Fernández, O. R., Smith, I. E., & Dietl, G. P.  
329 (2022). What is conservation paleobiology? Tracking 20 years of research and development. *Frontiers in  
330 Ecology and Evolution*, 10. <https://doi.org/10.3389/fevo.2022.1031483>

- 331 Dunhill, A. M., Zarzyczny, K., Shaw, J. O., Atkinson, J. W., Little, C. T. S., & Beckerman, A. P. (2024).  
332 Extinction cascades, community collapse, and recovery across a Mesozoic hyperthermal event. *Nature*  
333 *Communications*, 15(1), 8599. <https://doi.org/10.1038/s41467-024-53000-2>
- 334 Dunne, J. A., Labandeira, C. C., & Williams, R. J. (2014). Highly resolved early eocene food webs show  
335 development of modern trophic structure after the end-cretaceous extinction. *Proceedings of the Royal*  
336 *Society B: Biological Sciences*, 281(1782), 20133280. <https://doi.org/10.1098/rspb.2013.3280>
- 337 Dunne, J. A., Williams, R. J., Martinez, N. D., Wood, R. A., & Erwin, D. H. (2008). Compilation and Network  
338 Analyses of Cambrian Food Webs. *PLOS Biology*, 6(4), e102. <https://doi.org/10.1371/journal.pbio.0060102>
- 339 Erdős, P., & Rényi, A. (1959). On random graphs. i. *Publicationes Mathematicae Debrecen*, 6(3-4), 290–297.  
340 <https://doi.org/10.5486/pmd.1959.6.3-4.12>
- 341 Fricke, E. C., Hsieh, C., Middleton, O., Gorczynski, D., Cappello, C. D., Sanisidro, O., Rowan, J., Svenning,  
342 J.-C., & Beaudrot, L. (2022). Collapse of terrestrial mammal food webs since the Late Pleistocene. *Science*,  
343 377(6609), 1008–1011. <https://doi.org/10.1126/science.abn4012>
- 344 Gupta, A., Furrer, R., & Petchey, O. L. (2022). Simultaneously estimating food web connectance and structure  
345 with uncertainty. *Ecology and Evolution*, 12(3), e8643. <https://doi.org/10.1002/ece3.8643>
- 346 Hao, X., Holyoak, M., Zhang, Z., & Yan, C. (2025). Global Projection of Terrestrial Vertebrate Food  
347 Webs Under Future Climate and Land-Use Changes. *Global Change Biology*, 31(2), e70061. <https://doi.org/10.1111/gcb.70061>
- 348 Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F., Edwards, F., Figueroa,  
349 D., Jacob, U., Jones, J. I., Lauridsen, R. B., Ledger, M. E., Lewis, H. M., Olesen, J. M., Veen, F. J. F.  
350 van, Warren, P. H., & Woodward, G. (2009). Ecological networks—beyond food webs. *The Journal of*  
351 *Animal Ecology*, 78(1), 253–269. <https://doi.org/10.1111/j.1365-2656.2008.01460.x>
- 352 Jenny, D., Fuchs, D., Arkhipkin, A. I., Hauff, R. B., Fritschi, B., & Klug, C. (2019). Predatory behaviour and  
353 taphonomy of a Jurassic belemnoid coleoid (Diplobelida, Cephalopoda). *Scientific Reports*, 9(1), 7944.  
354 <https://doi.org/10.1038/s41598-019-44260-w>
- 355 Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30(12), 1883–1893.  
356 <https://doi.org/10.1111/1365-2435.12763>
- 357 Kiessling, W., Raja, N. B., Roden, V. J., Turvey, S. T., & Saupe, E. E. (2019). Addressing priority questions  
358 of conservation science with palaeontological data. *Philosophical Transactions of the Royal Society B:*  
359 *Biological Sciences*, 374(1788), 20190222. <https://doi.org/10.1098/rstb.2019.0222>
- 360 Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple  
361 building blocks of complex networks. *Science*, 298(5594), 824–827. <https://doi.org/10.1126/science.298.5594.824>
- 362
- 363

- 364 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from  
365 proxies. *Trends in Ecology & Evolution*, 30(6), 347–356. <https://doi.org/10.1016/j.tree.2015.03.014>
- 366 Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2008). Size, foraging, and food web structure.  
367 *Proceedings of the National Academy of Sciences*, 105(11), 4191–4196. <https://doi.org/10.1073/pnas.0710672105>
- 368
- 369 Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in  
370 Ecology and Evolution*, 14(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- 371 Poisot, T., Canard, E., Mouquet, N., & Hochberg, M. E. (2012). A comparative study of ecological  
372 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/>373 210x.2011.00174.x)
- 374 Rohr, R., Scherer, H., Kehrli, P., Mazza, C., & Bersier, L.-F. (2010). Modeling food webs: Exploring  
375 unexplained structure using latent traits. *The American Naturalist*, 176(2), 170–177. <https://doi.org/10.1086/653667>
- 376
- 377 Roopnarine, P. D. (2006). Extinction cascades and catastrophe in ancient food webs. *Paleobiology*, 32(1),  
378 1–19. <https://www.jstor.org/stable/4096814>
- 379 Schneider, F. D., Brose, U., Rall, B. C., & Guill, C. (2016). Animal diversity and ecosystem functioning in  
380 dynamic food webs. *Nature Communications*, 7(1), 12718. <https://doi.org/10.1038/ncomms12718>
- 381 Shaw, J. O., Dunhill, A. M., Beckerman, A. P., Dunne, J. A., & Hull, P. M. (2024). *A framework for  
382 reconstructing ancient food webs using functional trait data* (p. 2024.01.30.578036). bioRxiv. <https://doi.org/10.1101/2024.01.30.578036>
- 383
- 384 Stouffer, D. B., Camacho, J., Jiang, W., & Nunes Amaral, L. A. (2007). Evidence for the existence of a robust  
385 pattern of prey selection in food webs. *Proceedings of the Royal Society B: Biological Sciences*, 274(1621),  
386 1931–1940. <https://doi.org/10.1098/rspb.2007.0571>
- 387 Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz,  
388 N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap  
389 towards predicting species interaction networks (across space and time). *Philosophical Transactions of the  
390 Royal Society B: Biological Sciences*, 376(1837), 20210063. <https://doi.org/10.1098/rstb.2021.0063>
- 391 Thuiller, W., Calderón-Sanou, I., Chalmandrier, L., Gaüzère, P., O'Connor, L. M. J., Ohlmann, M., Poggianto,  
392 G., & Münkemüller, T. (2024). Navigating the integration of biotic interactions in biogeography. *Journal  
393 of Biogeography*, 51(4), 550–559. <https://doi.org/10.1111/jbi.14734>
- 394 Vullo, R. (2011). Direct evidence of hyodont shark predation on Late Jurassic ammonites. *Naturwissen-  
395 senschaften*, 98(6), 545–549. <https://doi.org/10.1007/s00114-011-0789-9>
- 396 Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774), 180–183.

- 397 <https://doi.org/10.1038/35004572>
- 398 Williams, R. J., & Martinez, N. D. (2004). Stabilization of chaotic and non-permanent food-web dynamics.
- 399 *The European Physical Journal B - Condensed Matter*, 38(2), 297–303. <https://doi.org/10.1140/epjb/e2004-00122-1>
- 400
- 401 Williams, R. J., & Martinez, N. D. (2008). Success and its limits among structural models of complex food
- 402 webs. *The Journal of Animal Ecology*, 77(3), 512–519. <https://doi.org/10.1111/j.1365-2656.2008.01362.x>
- 403 Yeakel, J. D., Pires, M. M., Rudolf, L., Dominy, N. J., Koch, P. L., Guimarães, P. R., & Gross, T.
- 404 (2014). Collapse of an ecological network in ancient egypt. *PNAS*, 111(40), 14472–14477. <https://doi.org/10.1073/pnas.1408471111>
- 405

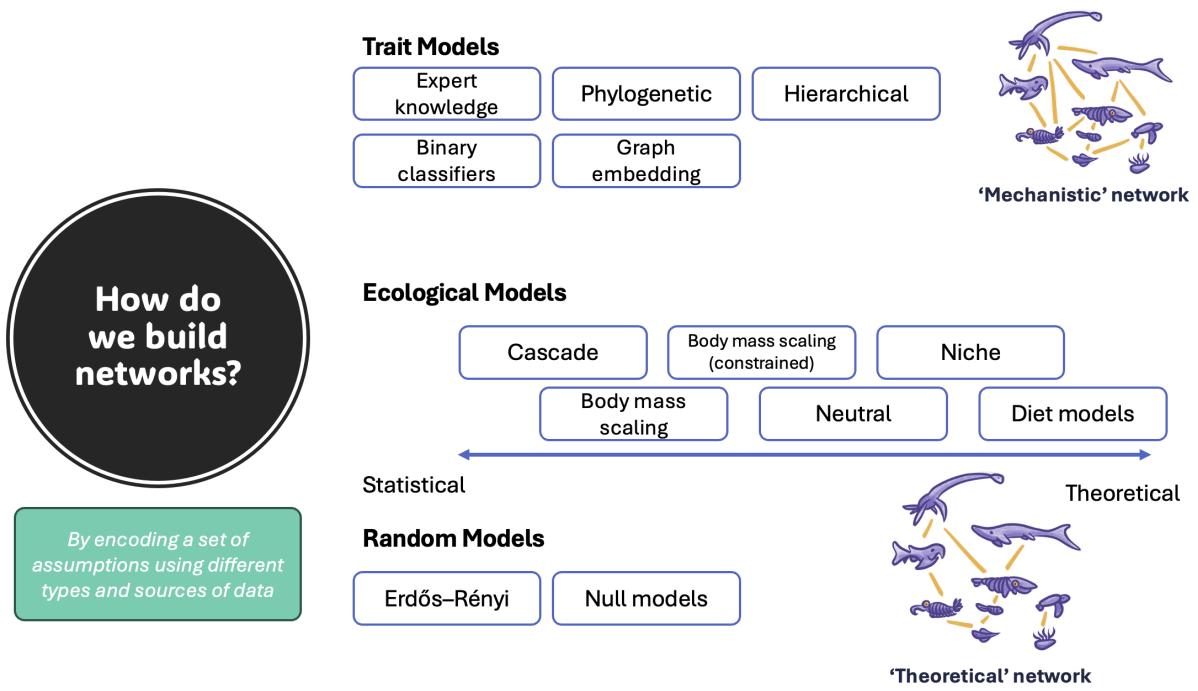


Figure 1: This obviously needs work but a variation on this to try and articulate the different approaches and broadly how they may differ.

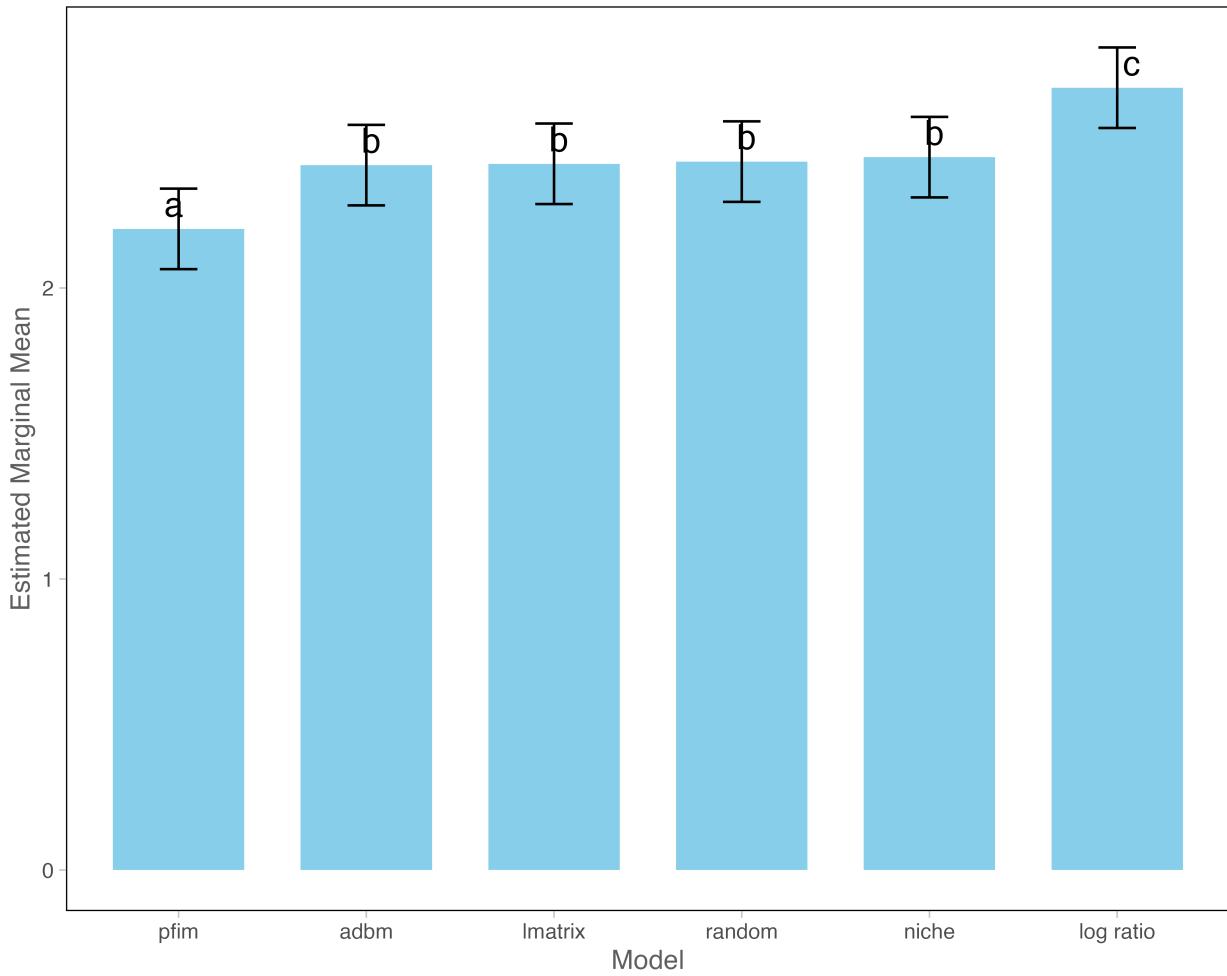


Figure 2: Estimated marginal means (EMMs) of ecological network metrics across six model types with 95% confidence intervals. Bars represent the predicted values for each model, and error bars indicate the 95% confidence limits. Letters above each bar denote Tukey-adjusted pairwise significance: models 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 adb and l-matrix lower, consistent with the MANOVA and post-hoc analyses.

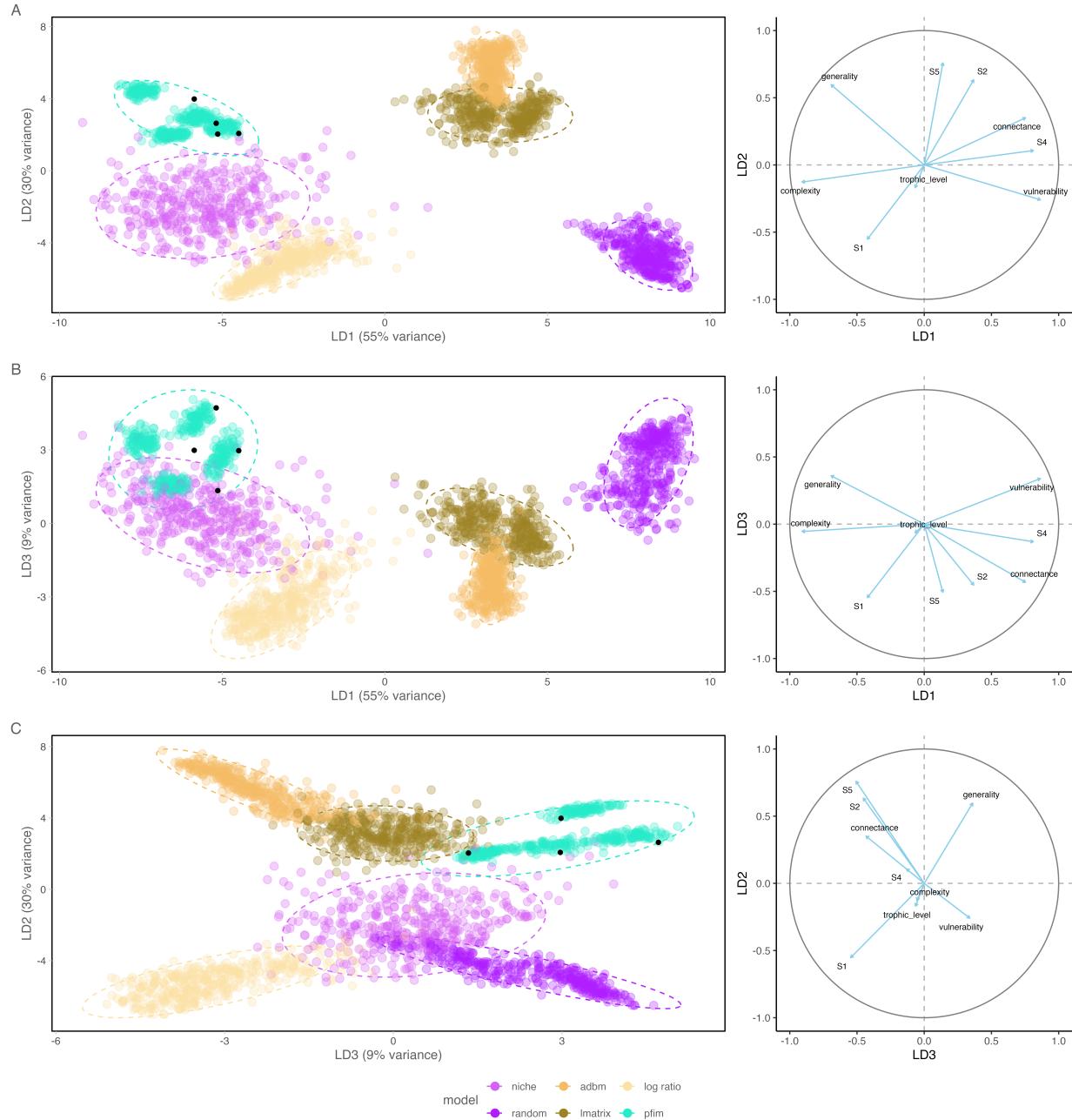


Figure 3: Linear discriminant analysis (LDA) of ecological network metrics for six model types. Each point represents a replicate, and ellipses indicate 95% confidence regions for each model. The second column represents the correlation of the various network metrics with the respective LDA axes.

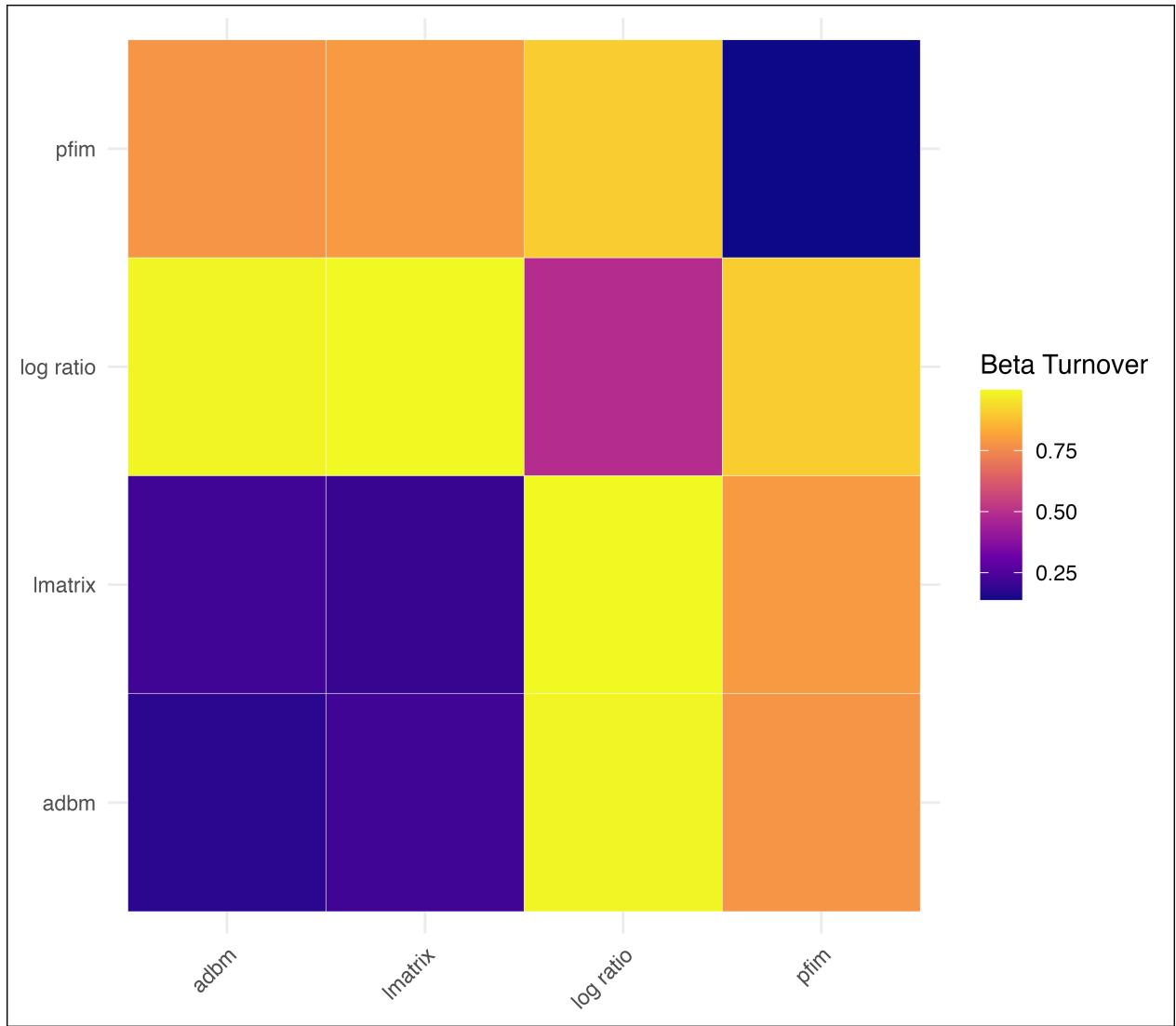
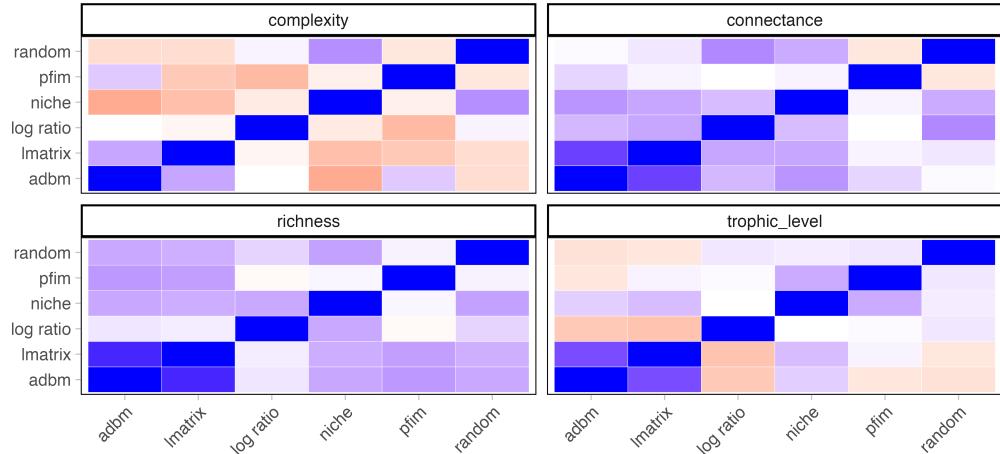
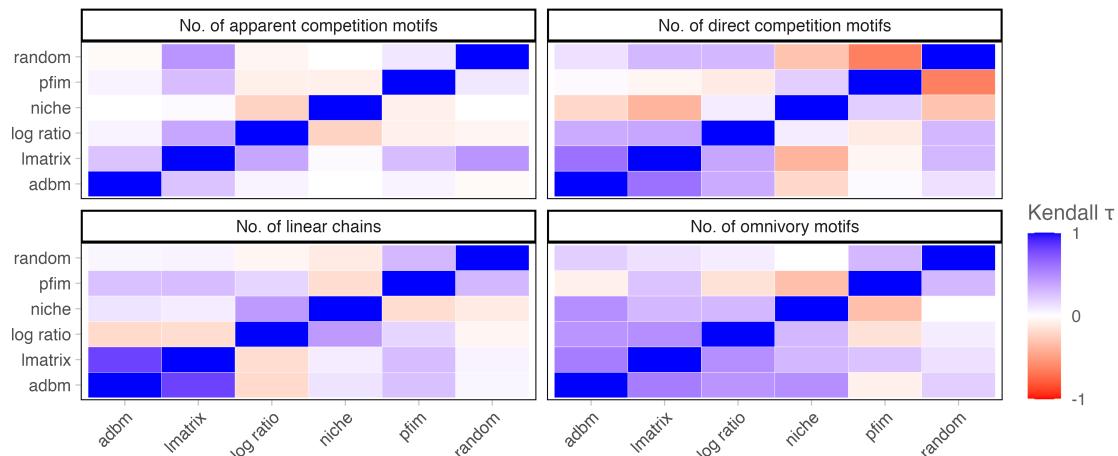


Figure 4: Pairwise beta turnover in species interactions among four ecological network models (adbm, lmatrix, log-ratio, and pfim). Each cell represents the mean turnover value between a pair of models, with warmer colors indicating greater dissimilarity in inferred interactions. The diagonal is omitted.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

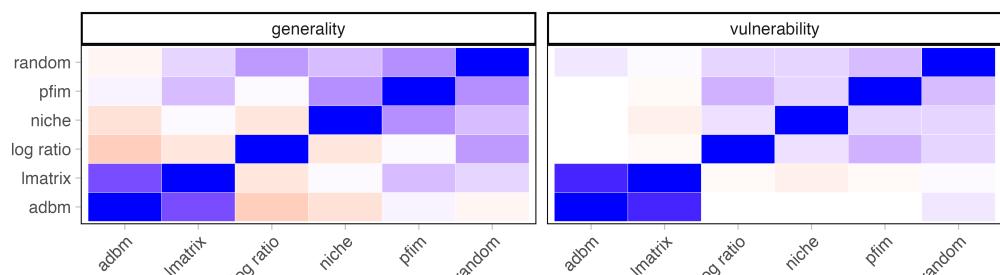
### Macro



### Meso



### Micro



### TSS

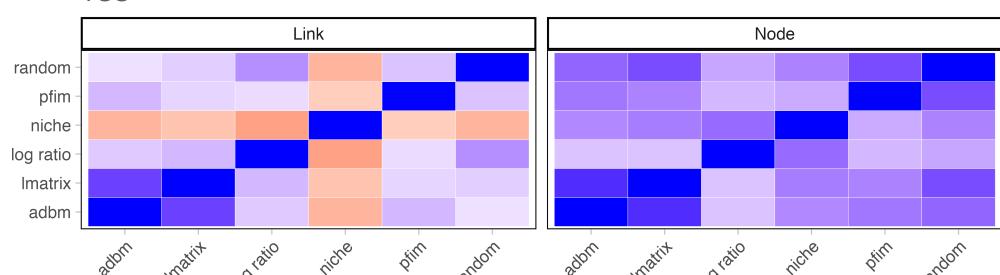


Figure 5: Heatmaps showing pairwise Kendall rank correlation coefficients ( $\tau$ ) between models for each network metric. Each panel corresponds to a different metric and displays the degree of agreement in extinction-scenario rankings across models based on mean absolute differences (MAD) between 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

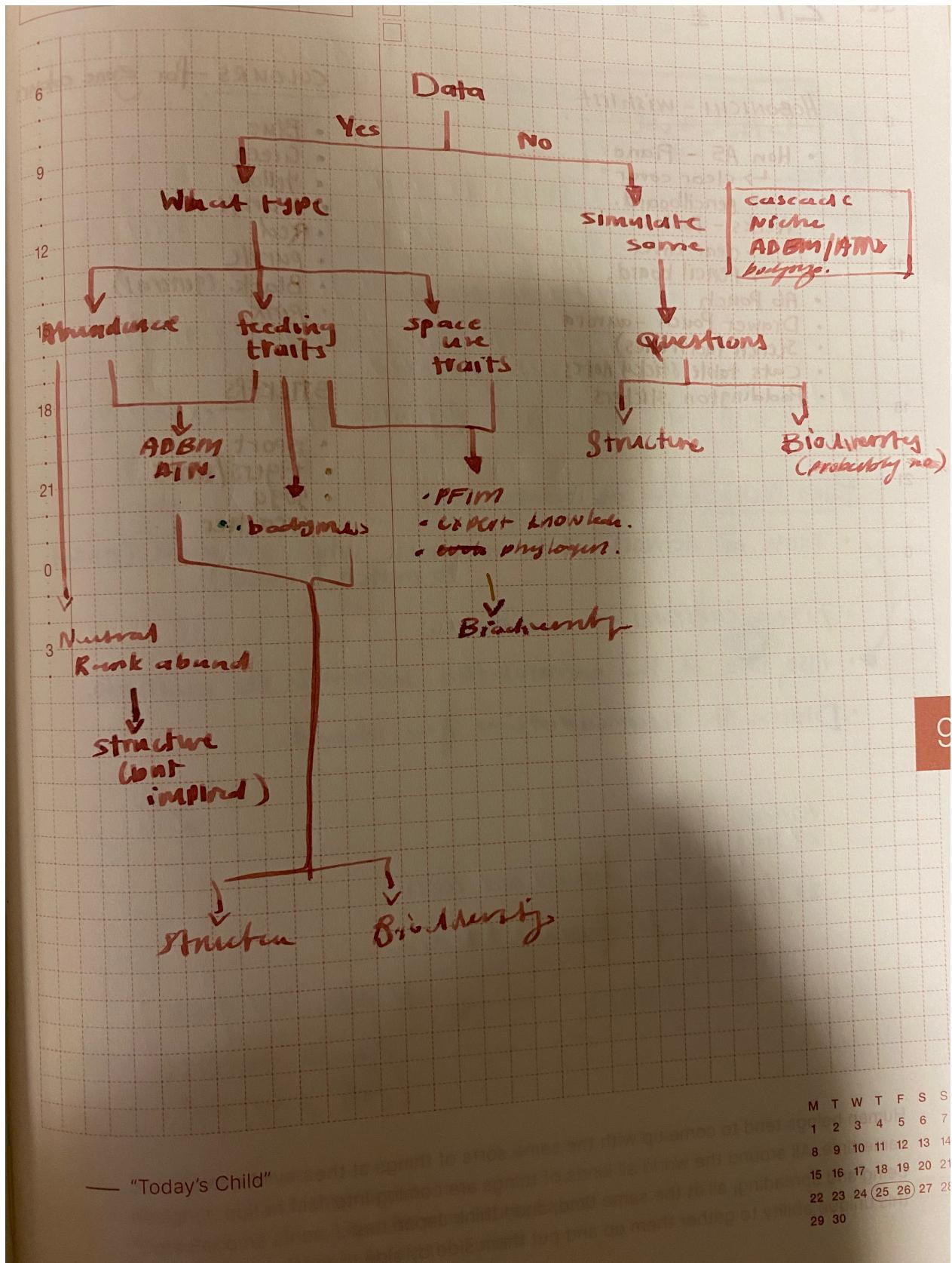


Figure 6: TODO.