

# 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 represent the feeding relationships between species and can help infer ecosystem-level processes. Alongside the development of food web theory, methods for constructing food webs have been developed to infer species interactions when empirical data is lacking. Food web construction methods are diverse, each utilising different approaches to infer species interactions —such as the use of traits to infer mechanistic relationships vs using gut content as a proxy for species diets. These methods have distinct theories, mechanisms, and data requirements. In paleoecology, where direct evidence of feeding interactions are rare, food web construction methods are especially valuable and affords us the opportunity to make inferences about paleo communities beyond simply a record of species composition. However, the limitations of paleontological data (e.g., information of species traits is limited to that which can be preserved) restrict which methods can reliably be used. By considering both ecological theory and the constraints of what can be derived from the fossil record, we identify the methods best suited for the construction of paleo food webs. Specifically, we focus on how these methods differ in the networks they produce and what these networks can reveal about species interactions. In doing so we hope to clarify the ecological nuances of network prediction and help prevent the accidental misuse or misinterpretation of paleo food webs.

**Keywords:** food web, network construction

1 There has been a growing interest in understanding community responses to environmental changes across  
2 deep time events as a means to help understand current and future biodiversity changes (Dillon et al., 2022;  
3 Kiessling et al., 2019). Species interactions and the resulting networks have gained popularity in contemporary  
4 settings as a means to help us to understand aspects of community composition and biodiversity (eg Thuiller  
5 et al. (2024) and ??) and so it is perhaps unsurprising that there has been a growing interest in using paleo  
6 food webs in a similar manner (*e.g.*, Dunhill et al., 2024 looked at...; Hao et al., 2025 looked at...; Yeakel et al.,  
7 2014 looked at...). However, one of the core challenges and limitations of being able to effectively *use* food  
8 webs is the challenge of *creating* them (Jordano, 2016), although this is a challenge within contemporary  
9 settings it is compounded in paleo contexts where, in the absence of being able to observe interactions, we  
10 are dependent on the fossil record (and the inherent limitation it imposes) to infer interactions. As a way to  
11 address the challenges with recording species interactions there has been the development of a large number  
12 of models and tools that can be used to infer either species interactions (see *e.g.*, Morales-Castilla et al., 2015;  
13 Pichler & Hartig, 2023; Strydom et al., 2021 for broader reviews) or networks (see *e.g.*, Allesina et al., 2008).  
14 Although there has been the development of models and tools that are specific for inferring paleo food webs  
15 (Fricke et al., 2022; Roopnarine, 2006; *e.g.*, Shaw et al., 2024), it should be noted that these models only  
16 occupy a subset of the broader family of approaches that are used to predict networks, as they typically only  
17 focus on assessing the feasibility of interactions between species. Being able to construct only one ‘type’ of  
18 network means that we are limited in the scope of questions that we can appropriately answer with those  
19 networks [see Strydom in prep; Gauzens et al. (2025)]. However, there is scope that models and tools that  
20 have been developed in contemporary settings have the potential to be used for paleo settings (*e.g.*, Yeakel et  
21 al., 2014), which opens the door for researchers to ask a broader and more complete range of questions about  
22 community responses to environmental change.

23 Here we aim to provide an overview of the different models that can be used to construct food webs using  
24 paleo data. Specifically we focus on identifying a suite of models that are appropriate for use with paleo data  
25 that can feasibly be constructed within the limitations that are imposed by fossil data while still spanning the  
26 larger network space. Additionally we use the data from Dunhill et al. (2024) as a case study to understand  
27 how different models recover different networks, both in terms of structure as well as pairwise interactions and  
28 establish if there are consequences for using networks that are constructed using different models in terms of  
29 making inferences about the behaviour of the system by looking at how the model type influences what we  
30 infer to be the dominant driver of extinctions across a mass extinction event.

<sup>31</sup> **1 Constructing paleo webs**

<sup>32</sup> **2 Challenges specific to building paleo networks**

<sup>33</sup> Although there has been a push for the development of tools and methods that allow us to predict species  
<sup>34</sup> interactions and networks they will not all be suitable for the prediction of paleo communities. This is  
<sup>35</sup> primarily due to limitations that we are faced with in terms of the information that can be inferred from  
<sup>36</sup> the fossil record (such as species traits, abundances, and assemblages), which is needed as input data for the  
<sup>37</sup> different models. The limited information available from the fossil record is compounded by the incomplete  
<sup>38</sup> and biased preservation of species [REF], which part of a species is preserved (part vs whole), the ambiguity of  
<sup>39</sup> the ‘true’ community composition [were communities conserved *in situ* or were they there owing to geological  
<sup>40</sup> processes?; REF], as well as the availability/accessibility of different rock layers (and thus the completeness of  
<sup>41</sup> data we might have for a specific era in time). Additionally there is an increasing degree of ‘fuzziness’ around  
<sup>42</sup> the ecology and life histories of species the further one moves back in geological time [REF]. This is not to say  
<sup>43</sup> that because we have imperfect data we should not be attempting to construct paleo food webs but rather  
<sup>44</sup> that we need to be aware of what the uncertainties are and how these might impact the assumptions that we  
<sup>45</sup> need to make when constructing a network (as well as how this will intersect with the intended end use of  
<sup>46</sup> the network). This will allow us to best identify an approach that minimises the assumption and potential  
<sup>47</sup> uncertainties within the data while still constructing a suitable network. This includes thinking about both  
<sup>48</sup> assumptions you are making about the actual data *e.g.*, trying to extrapolate body size from fossil data but  
<sup>49</sup> also assumptions across time *e.g.*, assuming modern trait-feeding modes are the same or that assumptions  
<sup>50</sup> about network structure will hold across deep time.

<sup>51</sup> **2.1 Understanding the approaches to network construction**

<sup>52</sup> Broadly we can think about network construction as being nested within two different schools of thought (and  
<sup>53</sup> thus methodological approaches, Figure 1), models that focus on assessing the *mechanistic* feasibility of an  
<sup>54</sup> interaction being able to occur between two species or models that are more closely married to specific bodies  
<sup>55</sup> of ecological *theory* - such as niche theory or foraging ecology. The former of which will construct ‘metawebs’  
<sup>56</sup> and the latter ‘realised networks’ [Strydom et al in prep]. Models that have specifically been developed  
<sup>57</sup> in the paleo space tend to be mechanistic in nature in that they focus on using a trait-based approach to  
<sup>58</sup> formalise feeding interactions (*e.g.*, Shaw et al. (2024); Roopnarine (2006)), are assembled by expert opinion  
<sup>59</sup> (*e.g.* Dunne et al. (2014)), or make assumptions based on the evolutionary signals of interactions (*e.g.*,  
60 Fricke et al. (2022)). Thus paleo models typically only construct metawebs, and there is the need for the

61 intentional adoption of theoretical models if we want to realise the full potential of questions and information  
 62 that we can glean from the fossil record. However, there is an argument that the fundamental ‘currencies of  
 63 life’ to have remained constant - *e.g.*, the energetic constraints of foraging or foraging niches, meaning that  
 64 theoretical models that have been developed and tested on contemporary food webs should still hold for paleo  
 65 communities.

66 [Figure 1 about here.]

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

Table 1: A summary of the different families of tools that can be used to generate paleo food webs.

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 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)

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)

### 71 3 Case study: Toarcian mass extinction event

#### 72 3.1 Dataset overview

##### 73 3.1.1 Species occurrence

74 Here we use the fossil occurrence data over an interval extends from the upper Pliensbachian (~185 Ma) to  
 75 the upper Toarcian (~175 Ma) of the Cleveland Basin (see Dunhill et al., 2024 for a more comprehensive  
 76 overview). The data set consists of a subset of four broad time periods (pre-extinction, post-extinction, early  
 77 recovery, and late recovery). The assemblages are treated as communities of interacting organisms. Something  
 78 about the total number of taxa as well as numbers per a time period? Probably also make a comment that  
 79 this is a 'deep time' community we are looking at.

##### 80 3.1.2 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 **3.1.3 Constructing networks**

87 For each paleo community (time bin) we constructed **100** networks for each model (so  $6 * 100$ ) networks.  
88 These networks were simplified so as to remove any disconnected species. In total 2 400 networks were  
89 constructed. When a quantitative measure of body size is needed (ADBM, body size ratio, and l-matrix) we  
90 drew a body mass for each species from a uniform distribution, with ranges being defined by the different  
91 size classes *e.g.*, a species classed as ‘very large’ would have a body mass drawn from  $U(300, 500)$ . This was  
92 repeated for each run in order to add variation to the networks constructed, however the same body sizes were  
93 kept consistent for the relevant models *i.e.*, an ADBM and l-matrix network from the same replicate have  
94 the same bodysizes. For both the random and niche model the desired connectance was randomly selected  
95 between the range 0.07 - 0.15 for each replicate but kept consistent for both models. For each network we  
96 calculated the properties listed in Table 2

97 **3.2 Models capture different network structure but in unexpected ways**

98 Broadly when we talk about quantifying the structure of a network we are interesting in capturing some aspect  
99 of how the links are distributed between nodes, or alternatively about properties of the nodes. Structure  
100 is useful as it is gives information as to how the interactions between species are distributed within the  
101 community, informing us on *e.g.*, energy flows and fluxes [REF], propagation of stress [REF], and something  
102 about trophic levels [REF]. We are also able to glean information on interaction strategies between smaller  
103 interacting units in the bigger community in the form of motifs (Milo et al., 2002; Stouffer et al., 2007).  
104 Motifs allow us to identify *e.g.*, the prevalence of competition, as well as smaller chains within the network.  
105 Node-level properties look at the the number of links coming in to (prey) or out of (predators) a node and are  
106 informative of diet specialisation.

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	

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

Metric	Definition	Scale	Reference (for maths), can make footnotes probs
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>107</sup> In terms of wanting to assess and compare across the different models it is beneficial to approach this task  
<sup>108</sup> by thinking about the different aspects of the network as well as interactions that are being predicted by  
<sup>109</sup> the different models across different ‘scales’ of organisation within the network, namely macro (the entire  
<sup>110</sup> network), meso (smaller interacting units within the network), and micro (species-level attributes). Although  
<sup>111</sup> there are a myriad of possible ways to ‘measure’ and analyse ecological networks (Delmas et al., 2018) we  
<sup>112</sup> have selected those outlined in Table 2 as they span different scales within the network and have been shown  
<sup>113</sup> to be informative of different ecological processes.

<sup>114</sup> Here we used a Multivariate Analysis Of Variance (MANOVA) to assess the differences between networks  
<sup>115</sup> generated by different models based on the combined information of the multiple structural network measures.  
<sup>116</sup> Model defined as `network structure values ~ model` additionally we used a Linear Discriminant Analysis  
<sup>117</sup> (LDA) to determine if different models produced networks with differing structure. In order to do the  
<sup>118</sup> MANOVA and LDA we had to create within model variation for the different networks, with the exception of  
<sup>119</sup> the PFIM model all models have some inherent variation. In order to generate variation within the PFIM  
<sup>120</sup> metawebs we applied a *minimal* degree of downsampling following the protocol described in Roopnarine

121 (2017). This downsampling approach uses a power law distribution to essentially ‘prune’ links from the most  
122 generalist species (See SUPP MATT for a more detailed overview).

123 The multivariate effect of model was statistically significant, Pillai’s Trace = 3.89,  $F(45, 11, 950) = 925.64$ ,  
124  $p < .001$ , indicating systematic differences across multiple ecological or network properties simultaneously.  
125 Follow-up univariate ANOVAs revealed that model type had significant effects on all nine dependent variables  
126 Table 3, show that the network structure differed markedly across the model types on every measured  
127 dimension. Model type accounts for the vast majority of variance in most network metrics (66%–92%),  
128 indicating profound differences in structure between models. The only exception is trophic level ( $\eta^2 = .19$ ),  
129 which still shows a large effect but is much smaller relative to the other metrics.

Table 3: Manova/univariate ANOVA results.

Metric	$F(df = 5, 2394)$	df	partial $\eta^2$
Connectance	2717.8	5	0.85
Complexity	2356.6	5	0.83
Max trophic level	108.86	5	0.19
Generality	5646	5	0.92
Vulnerability	3266.9	5	0.87
S1	1968.5	5	0.80
S2	1527.5	5	0.76
S4	940.79	5	0.66
S5	1919.4	5	0.80

130 Post-hoc pairwise comparisons using Tukey-adjusted estimated marginal means further clarified the differences  
131 among models. The PFIM differed significantly from all other models (all  $p < 0.001$ ). The niche and random  
132 models are similar to each other, and the adbM and lmatrix were also similar to each other Figure 2. This  
133 although there are clear structural difference between the models as a whole we still broadly see the grouping  
134 between the theoretical models (ADBM, l-matrix), structural models (Niche and Random), and a metaweb  
135 (PFIM). Although the bodymass ratio model deviates from this neat grouping it is perhaps not as surprising  
136 since this simplified version of the bodymass ratio model is strongly rooted in the niche-based processes that  
137 are also assumed in the Niche model and so it having some overlap with the other structural models is not  
138 that surprising.

139 [Figure 2 about here.]

<sup>140</sup> From the LDA the first two discriminant functions explained 72% and 18% of the variance, respectively.  
<sup>141</sup> Wilks' indicated that the discriminant functions significantly differentiated among models ( $\lambda = 0.12$ ,  $\chi^2 = 1024$ ,  $p < 0.001$ ). The LDA plot Figure 3 shows clear separation of the pfim model from the others along LD1, with adbm and lmatrix clustering closely together, and niche and random occupying intermediate positions.  
<sup>144</sup> Classification accuracy was 85%, confirming that the combination of dependent variables reliably distinguishes  
<sup>145</sup> model types.

<sup>146</sup> [Figure 3 about here.]

<sup>147</sup> The implications of the above results is that it is clear that different models will recover different structures -  
<sup>148</sup> across all structural measures and highlight how model selection has the potential to strongly shape ecological  
<sup>149</sup> inferences. Using a model that overestimates connectivity could exaggerate our inferences about redundancy  
<sup>150</sup> or disturbance risk, while overly sparse models could underestimate network complexity and functional links.  
<sup>151</sup> Therefore, the choice of model should align with the specific ecological question *e.g.*, in interest in exploring  
<sup>152</sup> *potential* redundancy, robustness, versus trying to understand *realistic* energy flow pathways. Ideally we  
<sup>153</sup> should couple our analyses with sensitivity analyses to assess how conclusions depend on model assumptions.  
<sup>154</sup> It also means that we cannot compare inferences made using different models but any generalisations  
<sup>155</sup> about observed patterns should be standardised across network model *type* at minimum. That is two say it  
<sup>156</sup> may not be completely illogical to make comparisons between two metawebs, however it would be unwise to  
<sup>157</sup> compare a metaweb to a theoretical network.

<sup>158</sup> These structural differences have consequences for predicting species persistence, stability, and  
<sup>159</sup> ecosystem functioning. For example, metrics like generality and vulnerability influence top-down  
<sup>160</sup> and bottom-up control, affecting how energy and biomass flow through trophic levels (Dunne et  
<sup>161</sup> al., 2002). Similarly, connectance and trophic coherence influence stability and resilience; more  
<sup>162</sup> coherent networks tend to resist perturbations, whereas very dense, highly connected webs may  
<sup>163</sup> either buffer or amplify disturbances depending on interaction strengths (Johnson et al., 2014).

### <sup>164</sup> 3.3 Some networks don't share any interactions and some share a lot

<sup>165</sup> In addition to wanting to measure network structure researchers may also be interested in understanding  
<sup>166</sup> aspects about the diets and predators of *specific* species in a community. In this instance the interest should  
<sup>167</sup> be in understanding how the pairwise links predicted between species pairs differ between models. Here we  
<sup>168</sup> look at the interaction turnover both within and between the different models (Poisot et al., 2012). This can  
<sup>169</sup> be thought of as the equivalent of species  $\beta$  turnover and tells us which interactions are 'conserved' (shared)  
<sup>170</sup> across the networks but only between species pairs that are shared - *i.e.*, this turnover is only driven by

171 interaction and not species turnover. Here we only compared networks that we constructed for the same  
172 period (as our interest is only in between model differences) and excluded the random and niche networks  
173 from consideration as these two models are essentially species agnostic.

174 Across the four network models, turnover in species interactions varied substantially, revealing clear differences  
175 in how each model approaches determining the presence of links between species pairs Figure 4. The log-  
176 ratio model consistently showed high turnover relative to all other approaches, indicating that it produces  
177 interaction pairs that are the most distinct from other models. In contrast the ADBM and l-matrix exhibited  
178 the lowest turnover and suggests strong agreement between how pairwise interactions are determined. This is  
179 unsurprising given the underlying inference mechanisms of the models. The PFIM displays an intermediate  
180 turnover, aligning most closely with log-ratio and least with ADBM and l-matrix. Although this result is  
181 unsurprising as the mechanisms that determine interactions in ADBM and l-matrix (a single trait (bodysize) +  
182 parameterisation of links by ecological theory) is very different from the PFIM model that makes assumptions  
183 on a trait-based, mechanistic hierarchy. Taken together, these results demonstrate that model choice strongly  
184 influences inferred pairwise interactions.

185 [Figure 4 about here.]

186 In terms of how model choice will influence our inference - this will have the biggest consequence when thinking  
187 about diet related questions. In Figure 4 we can see that the ADBM and PFIM are recovering (almost) totally  
188 different pairwise links and so will have very different answers when we want to start interrogating the specific  
189 interactions that may be established by a specific species within the network. Pragmatically when it comes to  
190 deciding which model is the most appropriate choice when it comes to diet-based questions it makes sense to  
191 contextualise this in the feasible vs realised interaction spectrum. Specifically that from a ‘philosophical’ basis  
192 if you are asking questions about possible diets of species then it makes sense to use models that fall firmly in  
193 the ‘feasible’ space *e.g.*, PFIM model or even something like the Fricke et al. (2022) model.

### 194 3.4 Model choice changes the narrative

195 As shown in both Figure 3 and Figure 4 different models do indeed create networks that look different despite  
196 being created using the same species pool (or at least richness in the case of the Niche and Random models).  
197 As we have alluded to multiple times these underlying differences could have consequences on inferences that  
198 we make about the behaviour of the system. Here we will look at how model choice alters the inference we  
199 make about the underlying mechanisms that were driving extinctions in the Toarcian extinction event. Here  
200 we again we use we followed the approach outlined in Dunhill et al. (2024) when simulating extinctions as  
201 well as assessing which extinction mechanism results in a simulated network most closely match the real post

202 extinction network. Extinction simulations were only run on the pre extinction networks whereby species  
203 were removed until they reached the ‘target richness’, which is the richness of the post extinction community.

204 **TODO** The whole shebang about the different extinction simulation scenarios?

205 Extinctions were simulated using different plausible mechanisms based on species traits (size, motility), their  
206 position within the network (generality, vulnerability), as well as randomly. Each network was subjected  
207 to **50** extinction runs for each extinction mechanism. The extinctions themselves were cascading in nature  
208 meaning that after the target species was removed all species that no longer had any prey were also deemed  
209 as extinct (secondary extinction), checking for secondary extinctions was then repeated until there were no  
210 longer any species without prey. This represents one extinction event and only then would the proceeding  
211 target species be removed from the network and cascading extinctions assessed again. Note that for extinction  
212 simulations which use the network position of a species to determine extinction order we follow the protocol  
213 from Curtsdotter et al. (2011) and reassess the vulnerability/generality of each species after each extinction  
214 event to ‘redetermine’ the extinction order based on their position in the newly generated network.

215 **TODO** not sure if we also want to unpack/showcase robustness  $R_{50}$  (Jonsson et al., 2015)

216 [Figure 5 about here.]

### 217 3.4.1 Trends over time

218 Generalized additive mixed modelling showed strong differences in both the magnitude and temporal dynamics  
219 of network structure across models. Model results indicated that both parametric and smooth terms  
220 significantly influenced network metrics across models, with some exceptions. For connectance, all parametric  
221 coefficients were negative except the intercept, and most smooth terms showed highly significant effects  
222 ( $p < 0.05$ ), except for the Niche ( $p = 0.48$ ) and Random ( $p = 0.016$ ) models. Complexity and generality  
223 similarly exhibited significant parametric and smooth effects, whereas for vulnerability, the smooth term the  
224 Random model was non-significant ( $p = 0.73$ ). Trophic level effects varied among models: modellmatrix  
225 had a non-significant parametric effect ( $p = 0.89$ ), while others were strongly significant. Across the motifs  
226 (S1, S2, S4, S5), most parametric and smooth effects were significant ( $p < 0.05$ ), though a few, such as the  
227 intercepts of S1 and S5 were non-significant. Overall, these results indicate model- and metric-specific variation  
228 in both baseline (parametric) and temporal (smooth) effects, highlighting the importance of model structure  
229 in influences the observed temporal trends.

230 **3.4.2 Inferred extinction drivers**

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

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

245 For each network metric, we treated each model as an independent evaluator of scenario performance. MAD  
246 and TSS values were converted to within-model rankings, with rank 1 assigned to the scenario with the  
247 smallest MAD (i.e., the closest match to the empirical value) or highest TSS score. Ranking was performed  
248 independently for each combination of model and network metric to avoid assumptions about comparability  
249 across metrics. To evaluate whether different models produced consistent rankings of extinction scenarios,  
250 we quantified rank correlation among models separately for each network metric. Agreement among model  
251 rankings was assessed using Kendall's rank correlation coefficient ( ), which measures the degree of concordance  
252 between two ordinal rankings. Kendall's was selected because it is robust for small sample sizes, handles  
253 tied ranks appropriately, and provides a direct measure of the probability that model pairs agree or disagree  
254 on the relative ordering of scenarios. Kendall's ranges from  $-1$  to  $+1$ , where  $+1$  indicates perfect agreement  
255 between rankings,  $0$  reflects no relationship, and  $-1$  represents complete disagreement such that one ranking is  
256 the exact reverse of the other.  
257 When we look at Kendall's for the MAD across network structure and models Figure 6 we see that generally  
258 there is a positive correlation between the different different models. This implies that different models

259 are often recovering a similar ranking of extinction mechanisms (as in the ‘signal’ as to which extinction  
260 mechanisms may be the most plausible are the same). Although there is not a strong agreement between  
261 models as values tend to be low it is positive to observe that it is not often that we have a completely different  
262 ranking of extinction mechanisms, with the exception of complexity and to a lesser degree connectance.  
263 Interestingly we once again see the strong similarity between the l-matrix and the ADBM. Broadly this is  
264 positive, as despite the fact that we show that these models construct very different networks structurally  
265 Figure 3 they are still (broadly) capturing a similar extinction mechanism signal.

266 [Figure 6 about here.]

267 [Figure 7 about here.]

268 When looking at the node-level TSS scores (Figure 7 panel 2) we see that in general the signal of the extinction  
269 mechanism is maintained across the different models. However as many of the extinction mechanisms are  
270 determine by the *traits* of the node it is not surprising that we see a similar signal as the species are being  
271 removed in the exact same order. The only two extinction mechanisms that are driven by The link-level TSS  
272 scores (Figure 7 panel 1) do not show the same signal within extinction mechanism. We see that the random  
273 and PFIM models have high TSS scores (*i.e.*, have a ‘good fit’), however in the case of the PFIM this is to be  
274 expected as the links are deterministic and so if you have the same two species pools you will recover the  
275 same links. The ‘stochastic’ element of the theoretical models (ADBM, l-matrix, and log-ratio) means that  
276 they create a degree of ‘noise’ at the link-level and thus they are probably inappropriate to use for the type of  
277 extinction mechanism question we are asking here - specifically does the real and the simulated network look  
278 the same. Link-level TSS is perhaps also not an appropriate approach to determine the ‘best fit’ extinction  
279 mechanism if used in isolation and we advocate that the node-level TSS score (or alternatively some measure  
280 of  $\beta$  diversity is used). Finally, if we were to focus only on node level TSS we do not observe any strong  
281 differences between the models and it suggests that node-level driven (topological) extinction processes are  
282 insensitive to model type.

## 283 4 Discussion (need a catchier heading)

284 I want this section to be more about contextualising model choice within the bigger research question discussion  
285 - i.e. mapping question and model choice more tightly...

286 Points to discuss:

- 287 • Guidelines - as a box? Can we give something concrete?? Something like Figure 8

- 288 • How to we synthesise these results? As in should we give clear directives or is it enough to do a bit  
289 more handwaving and have the bigger message be that model choice matters?

290 [Figure 8 about here.]

## 291 References

- 292 Allesina, S., Alonso, D., & Pascual, M. (2008). A general model for food web structure. *Science*, 320(5876),  
293 658–661. <https://doi.org/10.1126/science.1156269>
- 294 Bambach, R. K., Bush, A. M., & Erwin, D. H. (2007). Autecology and the Filling of Ecospace: Key Metazoan  
295 Radiations. *Palaeontology*, 50(1), 1–22. <https://doi.org/10.1111/j.1475-4983.2006.00611.x>
- 296 Curtsdotter, A., Binzer, A., Brose, U., De Castro, F., Ebenman, B., Eklöf, A., Riede, J. O., Thierry, A., & Rall,  
297 B. C. (2011). Robustness to secondary extinctions: Comparing trait-based sequential deletions in static and  
298 dynamic food webs. *Basic and Applied Ecology*, 12(7), 571–580. <https://doi.org/10.1016/j.baae.2011.09.008>
- 299 Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Dalla Riva, G. V., Fortin, M.-J., Gravel, D., Guimarães, P.  
300 R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2018). Analysing  
301 ecological networks of species interactions. *Biological Reviews*, 112540. <https://doi.org/10.1111/brv.12433>
- 302 Dillon, E. M., Pier, J. Q., Smith, J. A., Raja, N. B., Dimitrijević, D., Austin, E. L., Cybulski, J. D., De  
303 Entrambasaguas, J., Durham, S. R., Grether, C. M., Haldar, H. S., Kocáková, K., Lin, C.-H., Mazzini, I.,  
304 Mychajliw, A. M., Ollendorf, A. L., Pimiento, C., Regalado Fernández, O. R., Smith, I. E., & Dietl, G. P.  
305 (2022). What is conservation paleobiology? Tracking 20 years of research and development. *Frontiers in  
306 Ecology and Evolution*, 10. <https://doi.org/10.3389/fevo.2022.1031483>
- 307 Dunhill, A. M., Zarzyczny, K., Shaw, J. O., Atkinson, J. W., Little, C. T. S., & Beckerman, A. P. (2024).  
308 Extinction cascades, community collapse, and recovery across a Mesozoic hyperthermal event. *Nature  
309 Communications*, 15(1), 8599. <https://doi.org/10.1038/s41467-024-53000-2>
- 310 Dunne, J. A., Labandeira, C. C., & Williams, R. J. (2014). Highly resolved early eocene food webs show  
311 development of modern trophic structure after the end-cretaceous extinction. *Proceedings of the Royal  
312 Society B: Biological Sciences*, 281(1782), 20133280. <https://doi.org/10.1098/rspb.2013.3280>
- 313 Erdős, P., & Rényi, A. (1959). On random graphs. i. *Publicationes Mathematicae Debrecen*, 6(3-4), 290–297.  
314 <https://doi.org/10.5486/pmd.1959.6.3-4.12>
- 315 Fricke, E. C., Hsieh, C., Middleton, O., Gorczynski, D., Cappello, C. D., Sanisidro, O., Rowan, J., Svenning,  
316 J.-C., & Beaudrot, L. (2022). Collapse of terrestrial mammal food webs since the Late Pleistocene. *Science*,  
317 377(6609), 1008–1011. <https://doi.org/10.1126/science.abn4012>
- 318 Gauzens, B., Thouvenot, L., Srivastava, D. S., Kratina, P., Romero, G. Q., Berti, E., O’Gorman, E. J.,

- 319 González, A. L., Dézerald, O., Eisenhauer, N., Pires, M., Ryser, R., Farjalla, V. F., Rogy, P., Brose, U.,  
320 Petermann, J. S., Geslin, B., & Hines, J. (2025). Tailoring interaction network types to answer different  
321 ecological questions. *Nature Reviews Biodiversity*, 1–10. <https://doi.org/10.1038/s44358-025-00056-7>
- 322 Gupta, A., Furrer, R., & Petchey, O. L. (2022). Simultaneously estimating food web connectance and structure  
323 with uncertainty. *Ecology and Evolution*, 12(3), e8643. <https://doi.org/10.1002/ece3.8643>
- 324 Hao, X., Holyoak, M., Zhang, Z., & Yan, C. (2025). Global Projection of Terrestrial Vertebrate Food  
325 Webs Under Future Climate and Land-Use Changes. *Global Change Biology*, 31(2), e70061. <https://doi.org/10.1111/gcb.70061>
- 327 Jonsson, T., Berg, S., Pimenov, A., Palmer, C., & Emmerson, M. (2015). The reliability of R50 as a  
328 measure of vulnerability of food webs to sequential species deletions. *Oikos*, 124(4), 446–457. <https://doi.org/10.1111/oik.01588>
- 330 Jordano, P. (2016). Chasing Ecological Interactions. *PLOS Biology*, 14(9), e1002559. <https://doi.org/10.1371/journal.pbio.1002559>
- 332 Kiessling, W., Raja, N. B., Roden, V. J., Turvey, S. T., & Saupe, E. E. (2019). Addressing priority questions  
333 of conservation science with palaeontological data. *Philosophical Transactions of the Royal Society B:  
334 Biological Sciences*, 374(1788), 20190222. <https://doi.org/10.1098/rstb.2019.0222>
- 335 Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple  
336 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/>337 5594.824)
- 338 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from  
339 proxies. *Trends in Ecology & Evolution*, 30(6), 347–356. <https://doi.org/10.1016/j.tree.2015.03.014>
- 340 Petchey, O. L., Beckerman, A. P., Riede, J. O., & Warren, P. H. (2008). Size, foraging, and food web structure.  
341 *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/>342 0710672105)
- 343 Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in  
344 Ecology and Evolution*, 14(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- 345 Poisot, T., Canard, E., Mouquet, N., & Hochberg, M. E. (2012). A comparative study of ecological  
346 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/>347 210x.2011.00174.x)
- 348 Rohr, R., Scherer, H., Kehrli, P., Mazza, C., & Bersier, L.-F. (2010). Modeling food webs: Exploring  
349 unexplained structure using latent traits. *The American Naturalist*, 176(2), 170–177. [https://doi.org/10.1086/653667](https://doi.org/10.<br/>350 1086/653667)
- 351 Roopnarine, P. D. (2006). Extinction cascades and catastrophe in ancient food webs. *Paleobiology*, 32(1),

- 352 1–19. <https://www.jstor.org/stable/4096814>
- 353 Roopnarine, P. D. (2017). *Ecological Modelling of Paleocommunity Food Webs* (pp. 201–226). University of  
354 Chicago Press.
- 355 Schneider, F. D., Brose, U., Rall, B. C., & Guill, C. (2016). Animal diversity and ecosystem functioning in  
356 dynamic food webs. *Nature Communications*, 7(1), 12718. <https://doi.org/10.1038/ncomms12718>
- 357 Shaw, J. O., Dunhill, A. M., Beckerman, A. P., Dunne, J. A., & Hull, P. M. (2024). *A framework for  
358 reconstructing ancient food webs using functional trait data* (p. 2024.01.30.578036). bioRxiv. <https://doi.org/10.1101/2024.01.30.578036>
- 360 Stouffer, D. B., Camacho, J., Jiang, W., & Nunes Amaral, L. A. (2007). Evidence for the existence of a robust  
361 pattern of prey selection in food webs. *Proceedings of the Royal Society B: Biological Sciences*, 274(1621),  
362 1931–1940. <https://doi.org/10.1098/rspb.2007.0571>
- 363 Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz,  
364 N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap  
365 towards predicting species interaction networks (across space and time). *Philosophical Transactions of the  
366 Royal Society B: Biological Sciences*, 376(1837), 20210063. <https://doi.org/10.1098/rstb.2021.0063>
- 367 Thuiller, W., Calderón-Sanou, I., Chalmandrier, L., Gaüzère, P., O'Connor, L. M. J., Ohlmann, M., Poggiato,  
368 G., & Münkemüller, T. (2024). Navigating the integration of biotic interactions in biogeography. *Journal  
369 of Biogeography*, 51(4), 550–559. <https://doi.org/10.1111/jbi.14734>
- 370 Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774), 180–183.  
371 <https://doi.org/10.1038/35004572>
- 372 Williams, R. J., & Martinez, N. D. (2004). Stabilization of chaotic and non-permanent food-web dynamics.  
373 *The European Physical Journal B - Condensed Matter*, 38(2), 297–303. <https://doi.org/10.1140/epjb/e2004-00122-1>
- 375 Williams, R. J., & Martinez, N. D. (2008). Success and its limits among structural models of complex food  
376 webs. *The Journal of Animal Ecology*, 77(3), 512–519. <https://doi.org/10.1111/j.1365-2656.2008.01362.x>
- 377 Yeakel, J. D., Pires, M. M., Rudolf, L., Dominy, N. J., Koch, P. L., Guimarães, P. R., & Gross, T.  
378 (2014). Collapse of an ecological network in ancient egypt. *PNAS*, 111(40), 14472–14477. <https://doi.org/10.1073/pnas.1408471111>

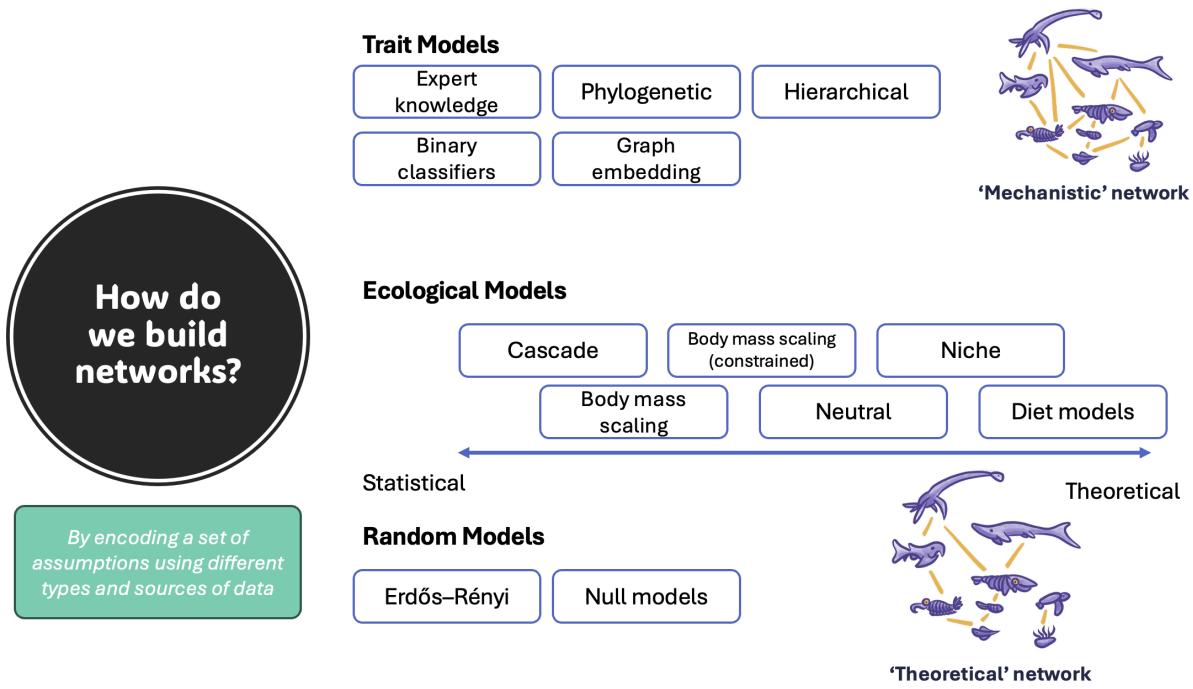


Figure 1: This obviously needs work but a variation on this to try and articulate the different approaches and broadly how they may differ.

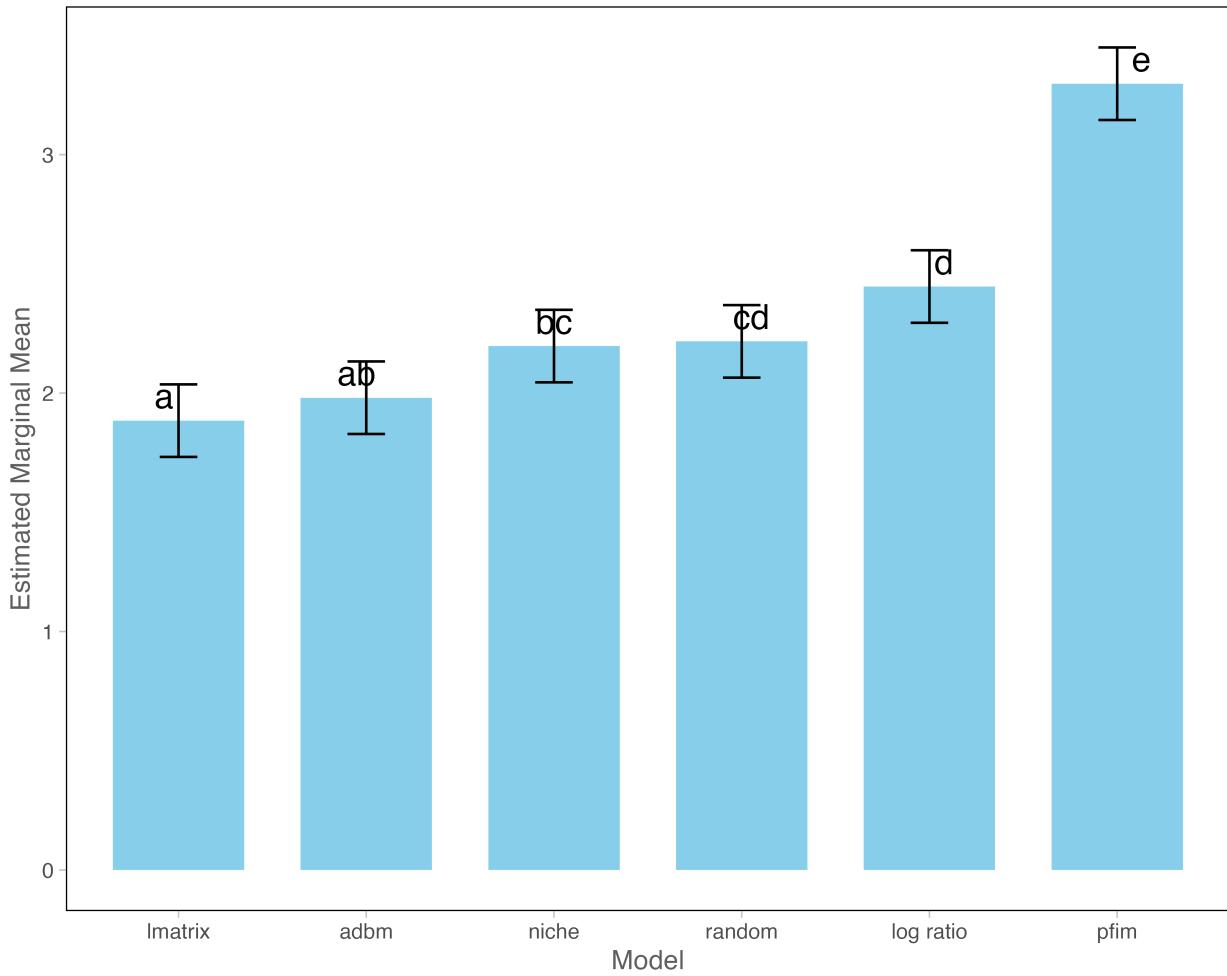


Figure 2: Estimated marginal means (EMMs) of ecological network metrics across six model types with 95% confidence intervals. Bars represent the predicted values for each model, and error bars indicate the 95% confidence limits. Letters above each bar denote Tukey-adjusted pairwise significance: models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

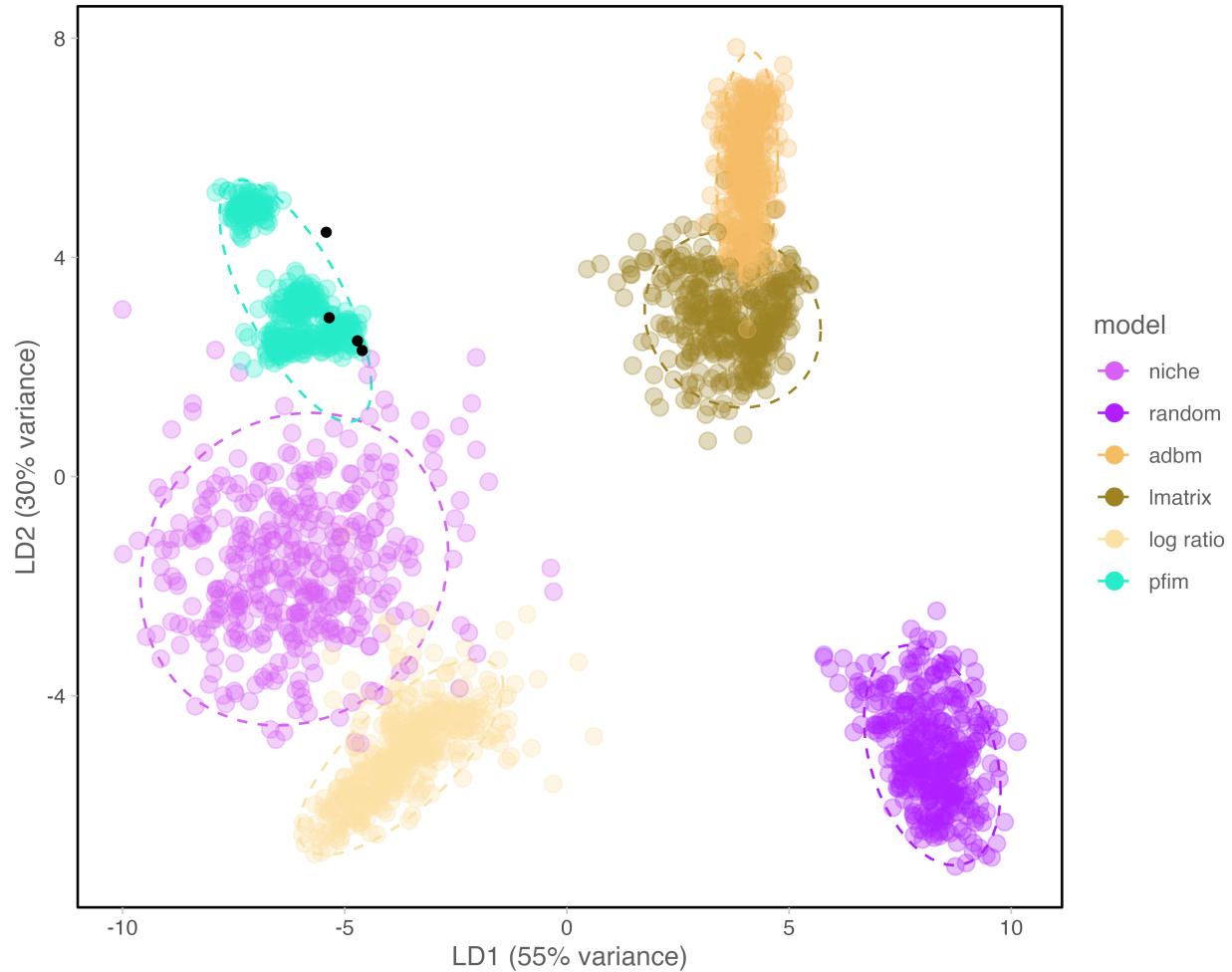


Figure 3: Linear discriminant analysis (LDA) of ecological network metrics for six model types. The first two discriminant functions (LD1 and LD2) explain 72% and 18% of the variance, respectively. Each point represents a replicate, and ellipses indicate 95% confidence regions for each model. The PFIM model is strongly separated along LD1, reflecting the highest values of network metrics, while adb and lmatrix cluster closely together, indicating similar, lower metric values. The niche and random models occupy intermediate positions. Classification accuracy of the LDA was 85%, demonstrating that the combination of dependent variables effectively discriminates among model types.

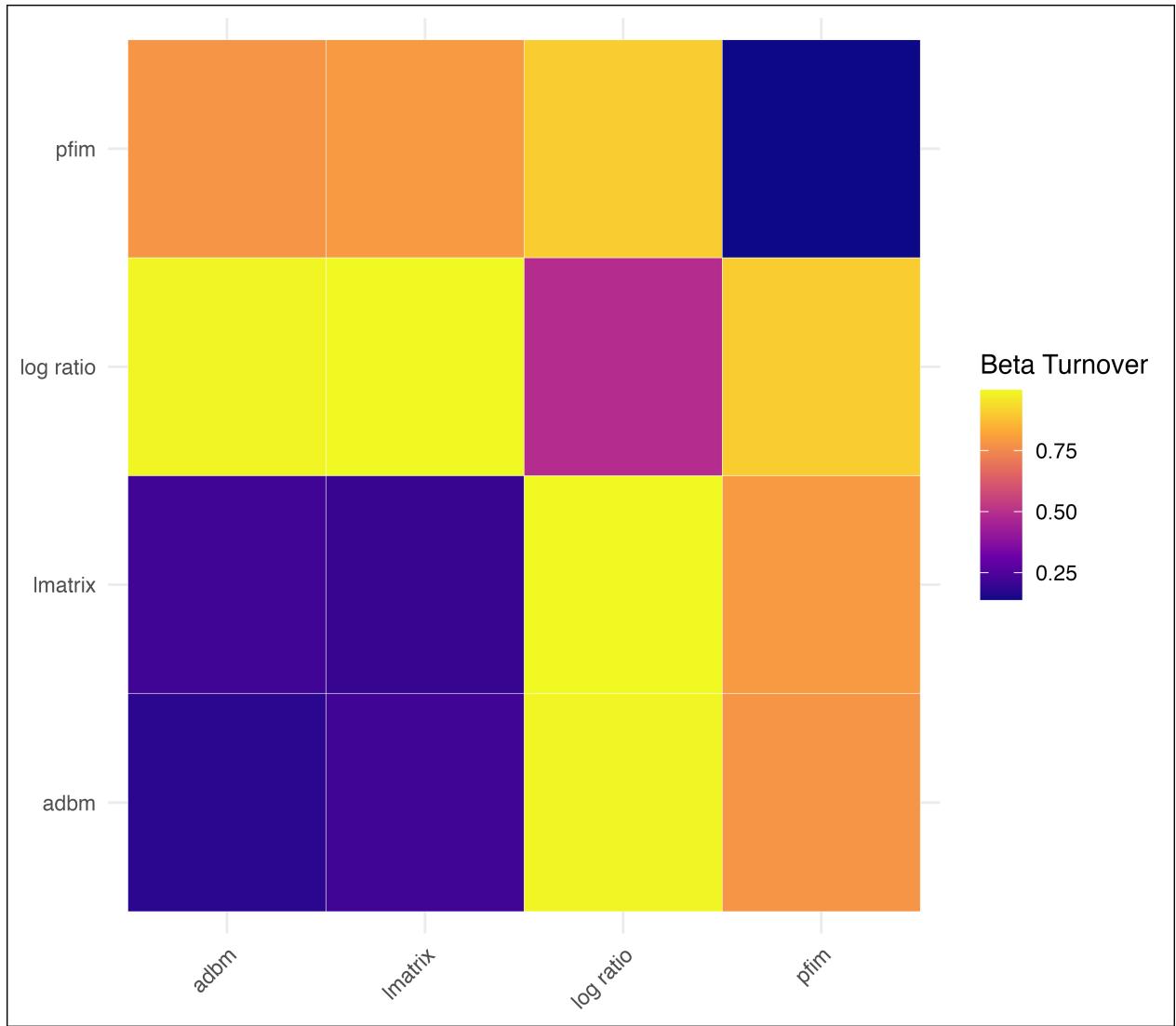


Figure 4: Pairwise beta turnover in species interactions among four ecological network models (adbm, lmatrix, log-ratio, and pfim). Each cell represents the mean turnover value between a pair of models, with warmer colors indicating greater dissimilarity in inferred interactions. The diagonal is omitted.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

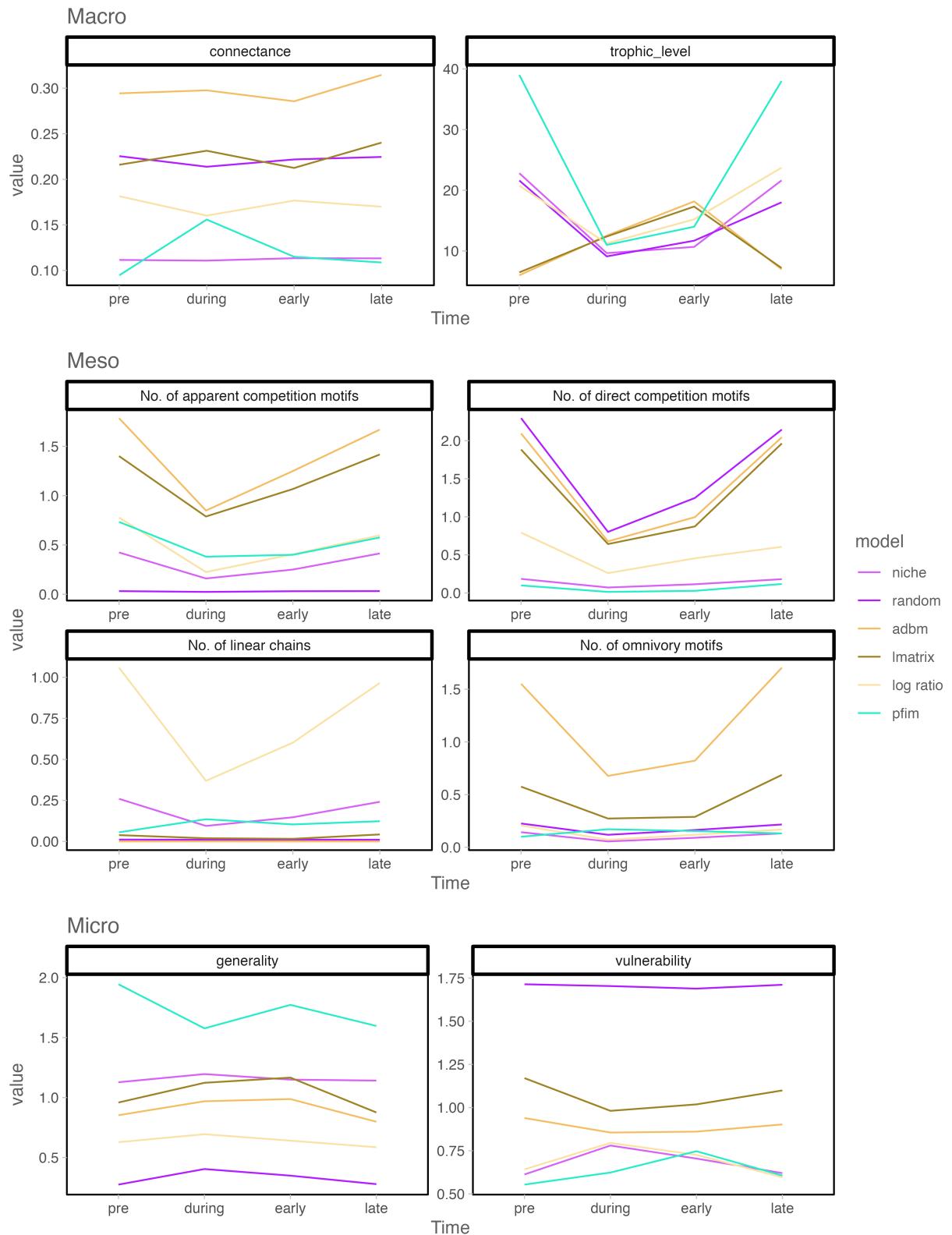
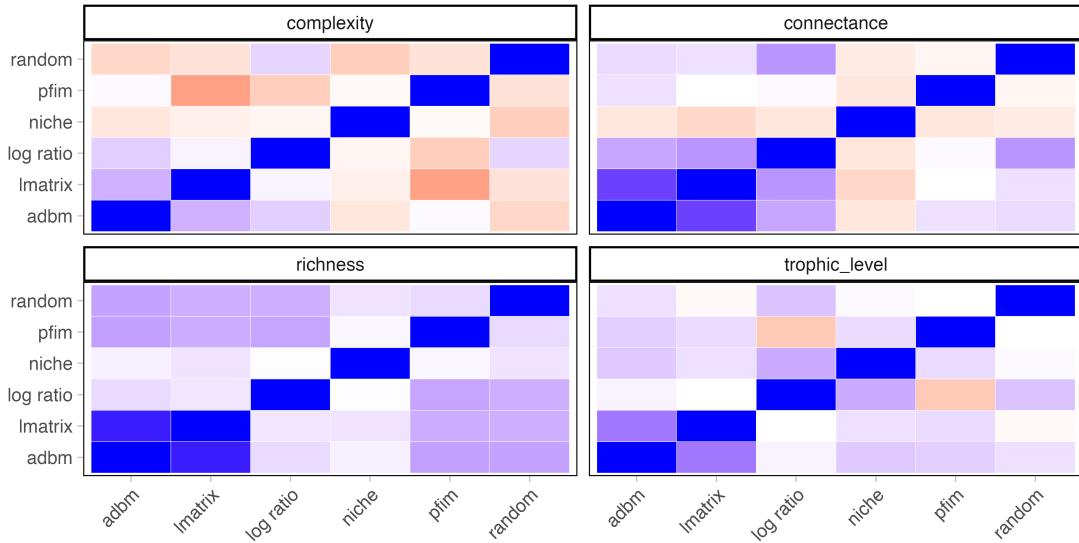
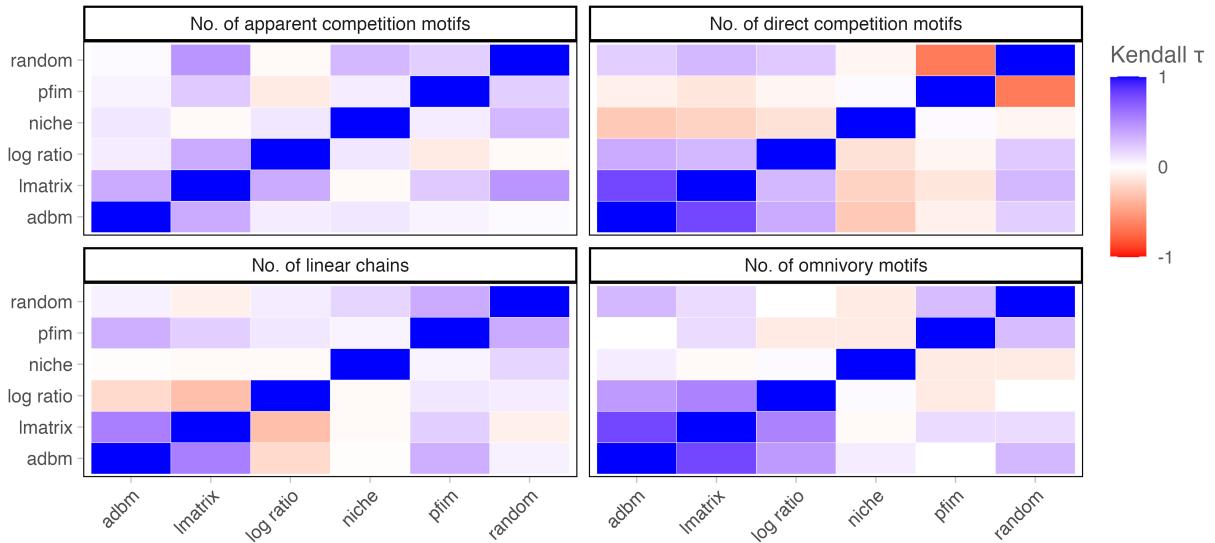


Figure 5: stuff... Recreation of the figure from Dunhill 2024. I think it makes sense to split and divide into the below sections (depending on potential figure limitations of target journal)

### Macro



### Meso



### Micro

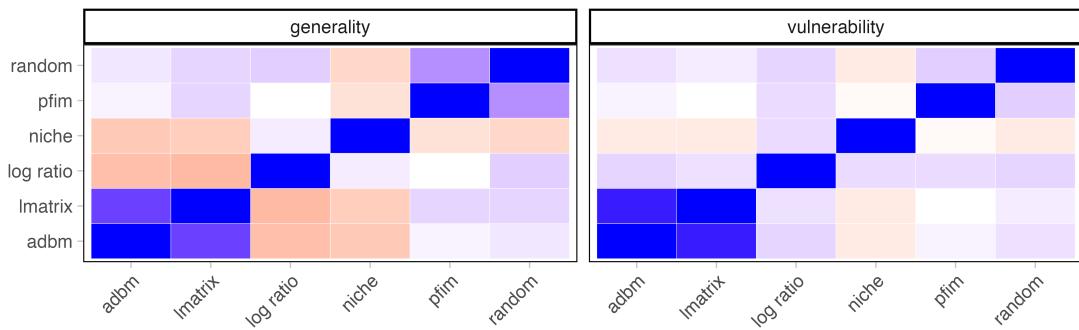


Figure 6: Heatmaps showing pairwise Kendall rank correlation coefficients ( $\tau$ ) between models for each network metric. Each panel corresponds to a different metric and displays the degree of agreement in extinction-scenario rankings across models based on mean absolute differences (MAD) between 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 illustrate how consistently different modelling approaches evaluate the relative realism of extinction scenarios across multiple network properties.

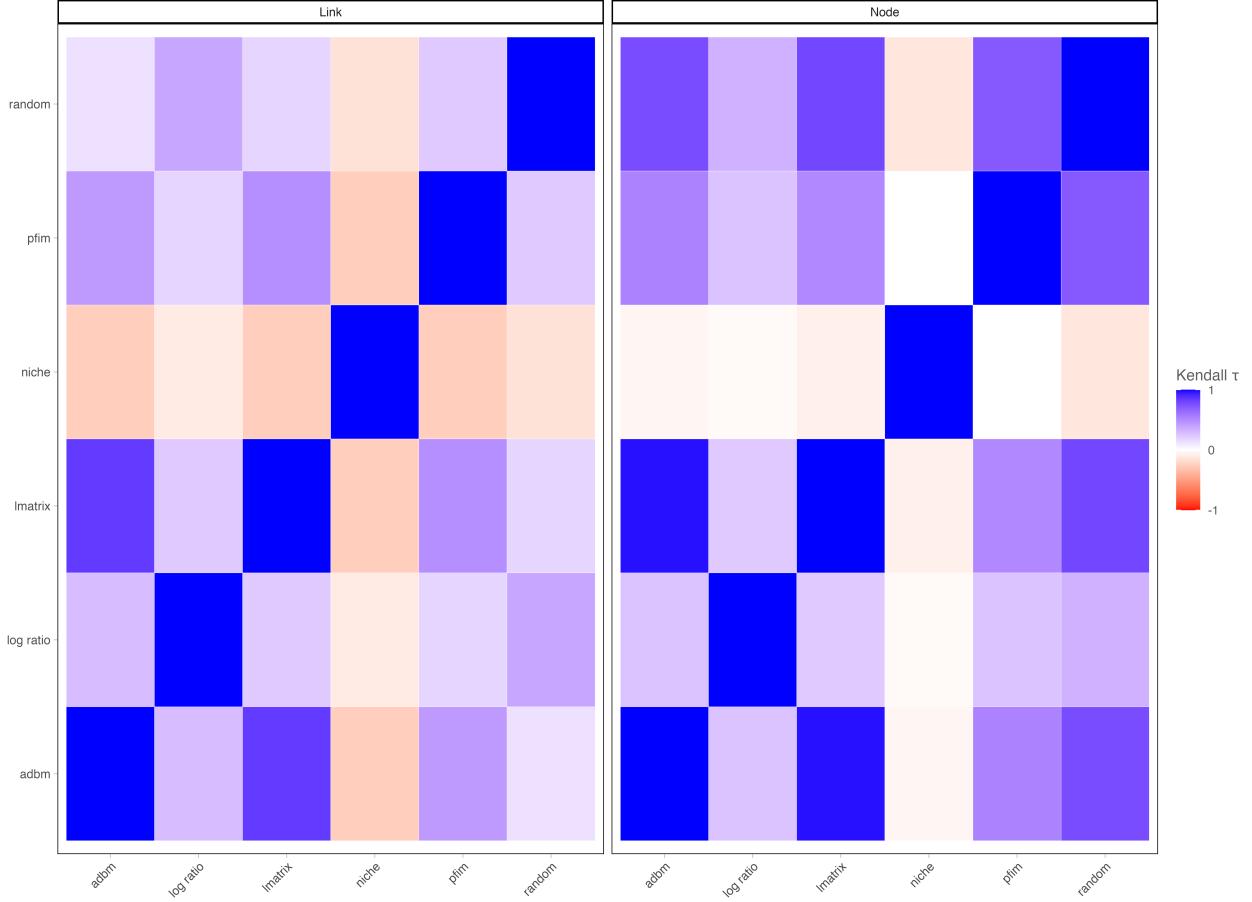


Figure 7: Heatmaps showing pairwise Kendall rank correlation coefficients ( $\tau$ ) between models for link and node level TSS. Each panel corresponds to a different metric and displays the degree of agreement in extinction-scenario rankings across models based on mean TSS scores between observed and 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 illustrate how consistently different modelling approaches evaluate the relative realism of extinction scenarios across multiple network properties.

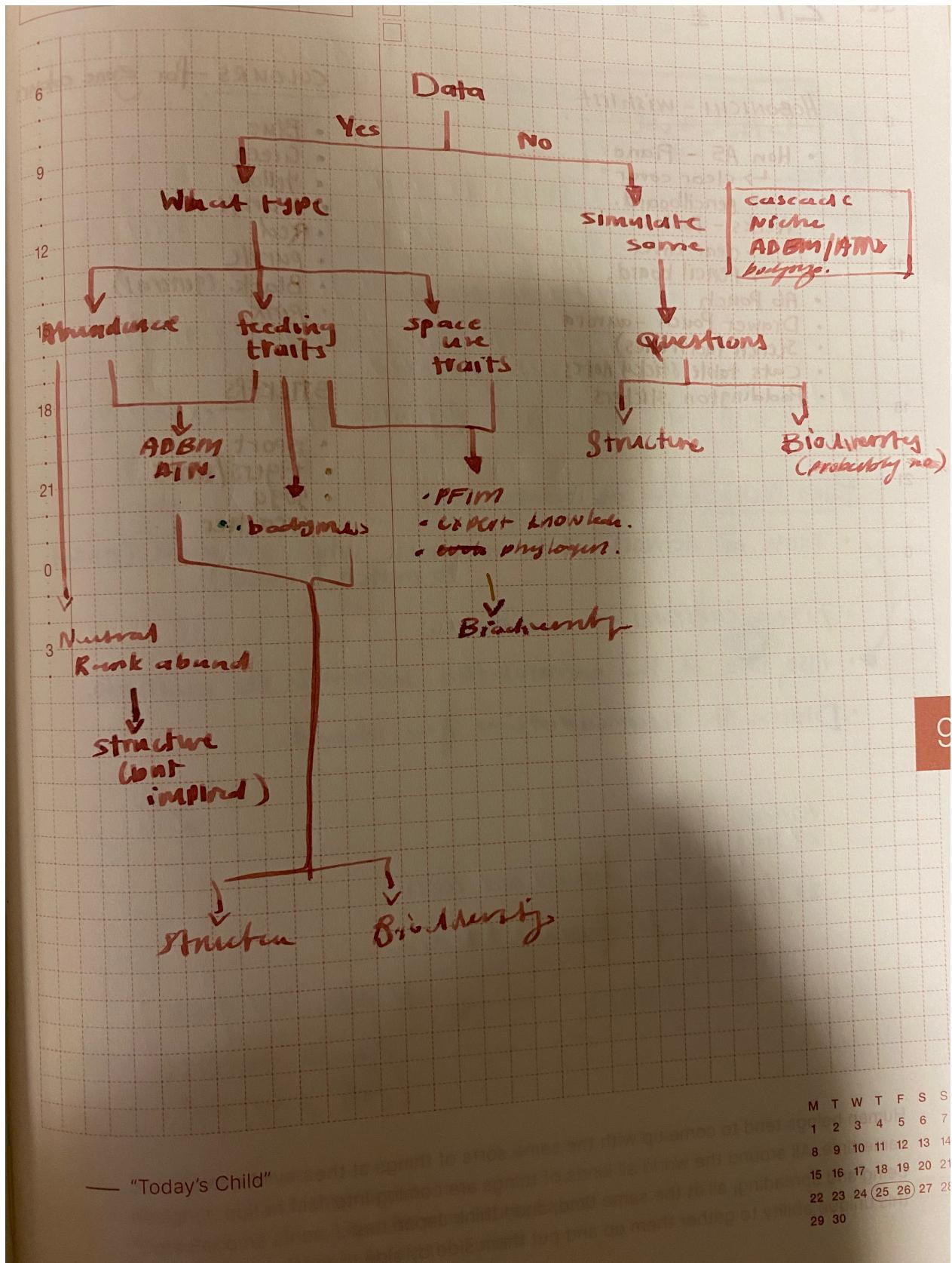


Figure 8: TODO.