

# A Roadmap Toward Predicting Species Interaction Networks (Across Space and Time)

Tanya Strydom<sup>1,2,‡</sup> Michael D. Catchen<sup>3,2,‡</sup> Francis Banville<sup>1,4,2</sup> Dominique Caron<sup>3,2</sup> Gabriel Dansereau<sup>1,2</sup>  
Philippe Desjardins-Proulx<sup>1,2</sup> Norma R. Forero-Muñoz<sup>1,2</sup> Gracielle Higino<sup>5</sup> Benjamin Mercier<sup>4,2</sup>  
Andrew Gonzalez<sup>3,2</sup> Dominique Gravel<sup>4,2</sup> Laura Pollock<sup>3,2</sup> Timothée Poisot<sup>1,2</sup>

<sup>1</sup> Université de Montréal <sup>2</sup> Québec Centre for Biodiversity Sciences <sup>3</sup> McGill University <sup>4</sup> Université de Sherbrooke <sup>5</sup> Universidade Federal de Goiás

<sup>‡</sup> These authors contributed equally to the work

## Correspondance to:

Timothée Poisot — timothee.poisot@umontreal.ca

**Abstract:** Networks of species interactions underpin numerous ecosystem processes, but comprehensively sampling these interactions is difficult. Interactions intrinsically vary across space and time, and given the number of species that compose ecological communities, it can be tough to distinguish between a true negative (where two species never interact) from a false negative (where two species have not been observed interacting even though they actually do). Assessing the likelihood of interactions between species is an imperative for several fields of ecology. This means that to predict interactions between species—and to describe the structure, variation, and change of the ecological networks they form—we need to rely on modeling tools. Here we provide a proof-of-concept, where we show how a simple neural-network model makes accurate predictions about species interactions given limited data. We then assess the challenges and opportunities associated with improving interaction predictions, and provide a conceptual roadmap forward toward predictive models of ecological networks that is explicitly spatial and temporal. We conclude with a brief primer on the relevant methods and tools needed to start building these models, which we hope will guide this research program forward.

## **1 Introduction**

2 Ecosystems are, in large part, constructed by the interactions within them — organisms interact  
3 with one-another and with their environment, either directly or indirectly. Interactions between  
4 individuals, populations, and species create networks of interactions that drive ecological and  
5 evolutionary dynamics and maintain the coexistence, diversity, and functioning of ecosystems  
6 (Delmas et al. 2018; Landi et al. 2018; Albrecht et al. 2018). Species interaction networks  
7 underpin our understanding of numerous ecological processes (Pascual and Dunne 2006; Heleno  
8 et al. 2014). Yet, even basic knowledge of species interactions (like being able to list them, or  
9 guess which ones may exist) remains one of the most severe biodiversity shortfalls (Hortal et  
10 al. 2015), in large part due to the tedious, time-consuming, and expensive process of collecting  
11 species interaction data. Comprehensively sampling every possible interaction is not feasible  
12 given the sheer number of species on Earth, and the data we can collect about interactions tend to  
13 be biased and noisy (de Aguiar et al. 2019). This is then compounded as species interactions are  
14 typically measured as a binary variable (present or absent) even though it is evident interactions  
15 are not all-or-nothing. Empirically we know species interactions occur probabilistically due to  
16 variation in species abundances in space and time (Poisot, Stouffer, and Gravel 2015). Different  
17 types of interactions vary in their intrinsic predictability (e.g. some fungal species engage in  
18 opportunistic saprotrophy (Smith et al. 2017), obligate parasites are more deterministic in their  
19 interactions than facultative parasites (Poisot et al. 2013; Luong and Mathot 2019)). In addition  
20 to this variance in predictability, networks from different systems are structured by different  
21 mechanisms.

22 Still, like all of Earth’s systems, species interaction networks have entered their “long now”  
23 (Carpenter 2002), where anthropogenic change will have long-term, low-predictability conse-  
24 quences (Burkle, Marlin, and Knight 2013) for our planet’s ecology. Therefore, our field needs  
25 a roadmap towards models that enable prediction (for the present) and forecasting (for the fu-  
26 ture) of species interactions and the networks they form, and which accounts for their spatial  
27 and temporal variation (McCann 2007; Seibold et al. 2018). As an example, in disease ecol-  
28 ogy, predicting potential hosts of novel disease (recently notably the search for wildlife hosts  
29 of betacoronaviruses; Becker et al. 2020; Wardeh, Baylis, and Blagrove 2021) has received

30 much attention. Network approaches have been used for the prediction of risk and dynamics of  
31 dengue (Zhao et al. 2020), Chagas disease (Rengifo-Correa et al. 2017), Rickettsiosis (Morand  
32 et al. 2020), Leishmaniasis (Stephens 2009), and a myriad infectious diseases in livestock and  
33 wildlife (Craft 2015). Additionally, prediction of interaction networks is a growing imperative  
34 for next-generation biodiversity monitoring, requiring a conceptual framework and a flexible set  
35 of tools to predict interactions that is explicitly spatial and temporal in perspective (Edwards et  
36 al. 2021; Magioli and Ferraz 2021; Zhang and He 2021). Developing better models for predic-  
37 tion of these interactions will rely on integration of data from many sources, and the sources for  
38 this data may differ depending on the type of interaction we wish to predict (Gibb et al. 2021).

39 Interactions between species can be conceptualised in a multitude of ways (mutualistic vs. antag-  
40 onistic, strong vs. weak, symmetric vs. asymmetric, direct vs. indirect) (Jordano 2016a; Morales-  
41 Castilla et al. 2015). What is common among definitions of species interactions is that *at least*  
42 one of the species is affected by the presence of another (Morales-Castilla et al. 2015). Net-  
43 works can be used to represent a variety of interaction types, including: *unipartite networks*:  
44 where each species can be linked to other species (often food webs), *bipartite networks*: where  
45 there are two pools of species and all interactions occur between species in each pool (typically  
46 used for pairwise interactions; e.g. hosts and parasites), and *k-partite networks*,: which expand  
47 to more than two discrete sets of interacting species (e.g., some parasitoid webs, seed dispersal  
48 networks, and pollination networks (Pocock, Evans, and Memmott 2012)).

49 Methods for predicting interactions between species exist, but at the moment are difficult to  
50 generalise as they are typically based around a single mechanism at a single scale: position in  
51 the trophic niche (Gravel et al. 2013; Petchey et al. 2008), phylogenetic distance (Pomeranz  
52 et al. 2018; Elmasri et al. 2020), functional trait matching (Bartomeus et al. 2016), interac-  
53 tion frequency (Weinstein and Graham 2017; Vázquez, Morris, and Jordano 2005), or other  
54 network properties (Terry and Lewis 2020; Stock et al. 2017). Species interaction networks,  
55 as we observe them on Earth today, are the product of ecological and evolutionary mechanisms  
56 interacting across spatial, temporal and organisational scales. The interwoven nature of these  
57 processes imposes structure on biodiversity data which is invisible when examined only through  
58 the lens of a single scale, however machine learning (ML) methods have enormous potential to

59 find this structure in data (Desjardins-Proulx, Poisot, and Gravel 2019), and have the potential  
60 to be used together with mechanistic models in order to make prediction of ecological dynamics  
61 more robust (Rackauckas et al. 2020).

62 Here we use a case study to show how machine-learning models (specifically a deep neural net-  
63 work) can enable prediction of species interactions: we construct a metaweb of host-parasite  
64 interactions across space, using predictors extracted from empirical data and accounting for the  
65 structure of co-occurrence between species. We use this case study to illustrate a roadmap for  
66 improving predictions using open data and ML methods; specifically, we focus on how emerging  
67 tools from ML can be used to deliver more accurate and more efficient predictions of ecolog-  
68 ical systems, and how the potential of these approaches will be magnified with increased data  
69 access. We then provide a non-exhaustive primer on the literature on interaction prediction, and  
70 identify the tools and methods most suited for the future of interaction network prediction mod-  
71 els, covering the spatial, temporal, and climatic dimensions of network prediction (Burkle and  
72 Alarcon 2011). Both the case study and primer are largely geared towards binary (interactions  
73 are either present or absent) networks; there are limitations in data and tools that make it a more  
74 reasonable starting approach. First, most ecological networks do not have estimates of inter-  
75 action strength, and particularly not estimates that are independent from relative abundances.  
76 Second, the methodological toolkit to analyse the structure of networks is far more developed  
77 for binary interactions (Delmas et al. 2018), meaning that the predictions of binary interactions  
78 can be more readily interpreted.

79 We argue that adopting a more predictive approach to complex ecological systems (like net-  
80 works) will establish a positive feedback loop with our understanding of these systems (Houla-  
81 han et al. 2017): the tasks of understanding and predicting are neither separate nor opposed  
82 (Maris et al. 2017); instead, ML tools have the ability to capture a lot of our understanding into  
83 working assumptions, and comparing predictions to empirical data gives us better insights about  
84 how much we ignore about the systems we model (see for example Borowiec et al. 2021, who  
85 provide an overview of deep learning techniques and concepts in ecology and evolution). Al-  
86 though data on species interaction networks are currently limited in the size and spatial coverage,  
87 machine learning approaches have a demonstrated track record of revealing the “unreasonable

88 effectiveness” of data (Halevy, Norvig, and Pereira 2009); we argue that with a clear roadmap  
89 guiding the use of these methods, the task of predicting species interaction networks will become  
90 more attainable.

91 **A case study: deep learning of spatially sparse host-parasite interac-  
92 tions**

93 The premise of this manuscript is that we can predict interactions between species. In this section  
94 we provide a proof-of-concept, where we use data from Hadfield et al. (2014) describing 51 host-  
95 parasite networks sampled across space. In this data, as in most spatially distributed ecological  
96 networks, not all species co-occur across sites. As a direct consequence there are pairs of species  
97 that may or may not be able to interact for which we have no data; furthermore there are pairs  
98 of species that may interact, but have only been documented in a single location where the  
99 interaction was not detected. In short, there are ecological reasons to believe that a number of  
100 negative associations in the metaweb (*sensu* J. Dunne 2006) are false negatives.

101 Without any species-level information, we resort to using both co-occurrence and known inter-  
102 actions to predict novel interactions. To do this we (i) extract features (equivalent to explanatory  
103 variables in a statistical model) for each species based on co-occurrence, (ii) use these features  
104 to train an artificial neural network to predict interactions, and (iii) apply this classifier (an algo-  
105 rithm that assigns a categorical output based on input features) to the original features to predict  
106 potential interactions across the entire species pool. Machine learning relies on a lexicon that  
107 shares some terms with statistics, albeit with different meaning; we expand on the precise mean-  
108 ings in the “How to validate a predictive model” section below. The outputs of the analysis are  
109 presented in fig. 1, and the code to reproduce it is available at <https://osf.io/6jp4b/>; the  
110 entire example was carried out in *Julia* 1.6.2 (Bezanson et al. 2017), using the *Flux* machine  
111 learning framework (Innes 2018).

112 We first aggregate all species into a co-occurrence matrix  $A$  which represents whether a given  
113 pair of species  $(i, j)$  was observed coexisting across any location. We then transform this co-  
114 occurrence matrix  $A$  via probabilistic PCA (Tipping and Bishop 1999) and use the first 15 values

from this PCA space as the features vector for each species  $i$ . For each pair of (host, parasite) species  $(i, j)$ , we then feed the features vectors  $(v_i, v_j)$  into a neural network. The neural network uses four feed-forward layers (each layer is independent from the one before and after); the first layer uses the RELU activation function (which ignores input below a threshold), the rest use a  $\sigma$  function (which transforms linear activation energies into logistic responses). All layers have appropriate dropout rates (in order to avoid over-fitting, only a fraction of the network is updated on each iteration: 1 – 0.8 for the first layer, 1 – 0.6 for the subsequent ones). This produces an output layer with a single node, which is the probability-score for interaction between species  $i$  and  $j$ .

We then train (equivalent to *fit*) this neural network by dividing the original dataset into testing and training sets (split 80-20 for training and testing respectively). During the training of this neural network (using the ADAM optimiser), the  $5 \times 10^4$  batches of 64 items used for training were constrained to have at least 25% of positive interactions, as Poisot, Ouellet, et al. (2021) show slightly inflating the dataset with positive interactions enables us to counterbalance sampling biases. Furthermore, setting a minimum threshold of response balance is an established approach for datasets with strong biases (Lemaître, Nogueira, and Aridas 2017). Validating this model on the test data shows our model provides highly effective prediction of interactions between pairs of species not present in the training data (fig. 1). The behaviour of the model was, in addition, checked by measuring the training and testing loss (difference between the actual value and the prediction, here using mean-squared error) and stopping well before they diverged (to avoid overfitting).

[Figure 1 about here.]

This case study shows that a simple neural network can be very effective in predicting species interactions even without additional species-level data. Applying this model to the entire dataset (including species pairs never observed to co-occur) identified 1546 new possible interactions – 746 (48%) of which were between pairs of species for which no co-occurrence was observed in the original dataset. This model reaches similar levels of predictive efficacy as previous studies that use far more species-level data and mechanistic assumptions (Gravel et al. 2013), which

143 serves to highlight the potential for including external sources of data for *improving* our predic-  
144 tion of interaction networks even further. For example, Krasnov et al. (2016) collected traits  
145 data for this system that could be added to the model, in addition or in substitution to latent  
146 variables derived from observed interactions.

147 **Predicting species interaction networks across space: challenges and  
148 opportunities**

149 Here we present a conceptual roadmap (fig. 2) which shows a conceptual path from data to  
150 prediction of species interaction networks, incorporating several modelling frameworks. We  
151 envisage this roadmap to be one conceptual path toward incorporating space in to our prediction  
152 of interaction networks, and developing spatially explicit models of networks and their proper-  
153 ties. In the following sections we discuss the challenges and opportunities for this path forward,  
154 and highlight two specific areas where it can have a strong impact: the temporal forecasting of  
155 species interaction networks structure, and the use of predicted networks for applied ecology  
156 and conservation biology.

157 [Figure 2 about here.]

158 **Challenges: constraints on predictions**

159 **Ecological network data are scarce and hard to obtain**

160 At the moment, prediction of species interactions is made difficult by the limited availability of  
161 data. Although we have seen a growth in species occurrence data, this growth is much slower  
162 for ecological interactions because species interactions are challenging to sample comprehen-  
163 sively (Bennett, Evans, and Powell 2019; Jordano 2016b) and sampling methodology has strong  
164 effects on the resulting data (de Aguiar et al. 2019). In turn, the difficulty of sampling interac-  
165 tions can lead to biases in our understanding of network structure (de Aguiar et al. 2019). This  
166 knowledge gap has motivated a variety of approaches to deal with interactions in ecological re-  
167 search based on assumptions that do not always hold, such as the assumption that co-occurrence

168 is equivalent to meaningful interaction strength (Blanchet, Cazelles, and Gravel 2020). Spatial  
169 biases in data coverage are prevalent at the global scale (with South America, Africa and Asia  
170 being under-represented) and different interaction types show biases towards different biomes  
171 (Poisot, Bergeron, et al. 2021). These “spatial gaps” serve as a limitation to our ability to con-  
172 fidently make predictions when accounting for real-world environmental conditions, especially  
173 in environments for which there are no analogous data.

174 Further, empirical estimation of interaction *strength* is highly prone to bias as existing data are  
175 usually summarised at the taxonomic scale of the species or higher, thereby losing informa-  
176 tion that differentiates the strength in per-individual interactions from the strength of a whole  
177 species interaction (Wells and O’Hara 2013). Empirical estimations of interaction strength are  
178 still crucial (Novak and Wootton 2008), but are a hard task to quantify in natural communities  
179 (Wootton 1997; Sala and Graham 2002; Wootton and Emmerson 2005), especially as the num-  
180 ber of species composing communities increases, compounded by the possibility of higher-order  
181 interactions or non-linear responses in interactions (Wootton and Emmerson 2005). Further,  
182 interaction strength is often variable and context dependent and can be influenced by density-  
183 dependence and spatio-temporal variation in community composition (Wootton and Emmerson  
184 2005).

### 185 **Powerful predictive tools work better on large data volumes**

186 This scarcity of data limits the range of computational tools that can be used by network ecolo-  
187 gists. Most deep learning methods, for instance, are very data expensive. The paucity of data is  
188 compounded by a collection of biases in existing datasets. Species interaction data are typically  
189 dominated by food webs, pollination, and host-parasite networks (Ings et al. 2009; Poisot et al.  
190 2020). This could prove to be a limiting factor when trying to understand or predict networks  
191 of underrepresented interaction types or when trying to integrate networks of different types  
192 (Fontaine et al. 2011), especially given their inherent structural variation (Michalska-Smith and  
193 Allesina 2019). This stresses the need for an integrated, flexible, and data-efficient set of com-  
194 putational tools which will allow us to predict ecological networks accurately from existing and  
195 imperfect datasets, but also enable us to perform model validation and comparison with more

flexibility than existing tools. We argue that fig. 1 is an example of the promise of these tools even when facing datasets of small size. The ability to extract and engineer features also serves to bolster our predictive power. Although it may be tempting to rely on approaches like bootstrapping to estimate the consistency of the predictions, we are confronted with the issues of low data volume and data bias—that we are more likely to observe interactions between some pairs of species (i.e. those that co-occur often, e.g. Cazelles et al. (2015), and those with higher relative abundance, e.g. Vazquez et al. (2009)). This introduces risk in training models on pseudo-replicated data. In short, the current lack of massive datasets must not be an obstacle to prediction; it is an ideal testing ground to understand how little data is sufficient to obtain actionable predictions, and how much we can rely on data inflation procedures to reach this minimal amount.

## Scaling-up predictions requires scaled-up data

We are also currently limited by the level of biological organisation at which we can describe ecological networks. For instance, our understanding of individual-based networks (e.g., M. S. Araújo et al. 2008; Tinker et al. 2012) is still in its infancy (Guimarães 2020) and acts as a resolution-limit. Similarly, the resolution of environmental (or landscape) data also limits our ability to predict networks at small scales, although current trends in remote sensing would suggest that this will become less of a hindrance with time (Makiola et al. 2020). Ecosystems are a quintessential complex-adaptive-system (Levin 1998) with a myriad of processes at different spatial, temporal, and organisational scales that influence and respond to one another. Understanding how the product of these different processes drive the properties of ecosystems across different scales remains a central challenge of ecological research, and we should strive to work on methods that will integrate different empirical “snapshots” of this larger system.

219 **Opportunities: an emerging ecosystem of open tools and data**

220 **Data are becoming more interoperable**

221 The acquisition of biodiversity and environmental data has tremendously increased over the past  
222 decades thanks to the rise of citizen science (J. L. Dickinson, Zuckerberg, and Bonter 2010) and  
223 of novel technology (Stephenson 2020), including wireless sensors (Porter et al. 2005), next-  
224 generation DNA sequencing (Creer et al. 2016), and remote sensing (Skidmore and Pettorelli  
225 2015; Lausch et al. 2016). Open access databases, such as [GBIF](#) (for biodiversity data), [NCBI](#)  
226 (for taxonomic and genomics data), [TreeBASE](#) (for phylogenetics data), [CESTE](#) (Jeliazkov et  
227 al. 2020) (for metacommunity ecology and species traits data), and [WorldClim](#) (for bioclimatic  
228 data) contain millions of data points that can be integrated to monitor and model biodiversity at  
229 the global scale. For species interactions data, at the moment [Mangal](#) is the most comprehensive  
230 open database of published ecological networks (Poisot et al. 2016), and [GloBI](#) is an extensive  
231 database of realised and potential species interactions (Poelen, Simons, and Mungall 2014). De-  
232 veloping standard practices in data integration and quality control (Kissling et al. 2018) and in  
233 next-generation biomonitoring (NGB; Makiola et al. 2020) would improve our ability to make  
234 reliable predictions of ecosystem properties on increasing spatial and temporal scales. The ad-  
235 vancement of prediction techniques coupled with a movement towards standardising data col-  
236 lection protocols (e.g. Pérez-Harguindeguy et al. (2013) for plant functional traits) and metadata  
237 (e.g. [DarwinCore](#))—which facilitates interoperability and integration of datasets—as well as a  
238 growing interest at the government level (Scholes et al. 2012) paints a positive picture for data  
239 access and usability in the coming years.

240 **Machine learning tools are becoming more accessible**

241 This effort is also supported by a thriving ecosystem of data sources and novel tools. ML meth-  
242 ods can often be more flexible and perform better than classical statistical methods, and can  
243 achieve a very high level of accuracy in many predictive and classification tasks in a relatively  
244 short amount of time (e.g., Cutler et al. 2007; Krizhevsky, Sutskever, and Hinton 2017). In-  
245 creasing computing power combined with recent advances in machine learning techniques and

246 applications shows promise in ecology and environmental science (see Christin, Hervet, and  
247 Lecomte (2019) for an overview). Moreover, ongoing developments in deep learning are aimed  
248 at improvement in low-data regimes and with unbalanced datasets (Antoniou, Storkey, and Ed-  
249 wards 2018; Chawla 2010). Considering the current biases in network ecology (Poisot, Berg-  
250 eron, et al. 2021) and the scarcity of data of species interactions, the prediction of ecological  
251 networks will undoubtedly benefit from these improvements. Machine learning methods are  
252 emerging as the new standard in computational ecology in general (Olden, Lawler, and Poff  
253 2008; Christin, Hervet, and Lecomte 2019), and in network ecology in particular (Bohan et al.  
254 2017), as long as sufficient, relevant data are available. Many studies have used machine learn-  
255 ing models specifically with ecological interactions. Relevant examples include species traits  
256 used to predict interactions and infer trait-matching rules (Desjardins-Proulx et al. 2017; Pichler  
257 et al. 2020), automated discovery of food webs (Bohan et al. 2011), reconstruction of ecologi-  
258 cal networks using next-generation sequencing data (Bohan et al. 2017), and network inference  
259 from presence-absence data (Sander, Wootton, and Allesina 2017). As many ecological and evo-  
260 lutionary processes underlie species interactions and the structure of their ecological networks  
261 (e.g., Vazquez et al. 2009; Segar et al. 2020), it can be difficult to choose relevant variables and  
262 model species interactions networks explicitly. A promising application of machine learning in  
263 natural sciences is Scientific-Machine Learning (SciML), a framework that combines machine  
264 learning with mechanistic models (Chuang and Keiser 2018; Rackauckas et al. 2020).

## 265 **A primer on predicting ecological networks**

266 Within the constraints outlined in the previous section, we now provide a primer on the back-  
267 ground concepts necessary to build predictive models of species interaction networks, with a  
268 focus on using machine learning approaches in the modelling process. As fig. 2 illustrates, this  
269 involves a variety of numerical and computational approaches; therefore, rather than an exhaus-  
270 tive summary, we aim to convey a high-level understanding that translates the core concepts into  
271 their application to ecological networks.

272 **Models**

273 **What is a predictive model?**

274 Models are used for many purposes, and the term “model” itself embodies a wide variety of  
275 meanings in scientific discourse. All models can be thought of as a function,  $f$ , that takes a set  
276 of inputs  $x$  (also called features, descriptors, or independent variables) and parameters  $\theta$ , and  
277 maps them to predicted output states  $y$  (also called label, response, or dependent variable) based  
278 on the input to the model:  $y = f(x, \theta)$ .

279 A given model  $f$  can be used for either descriptive or predictive purposes. Many forms of sci-  
280 entific inquiry are based around using models *descriptively*, a practice also called inference, the  
281 inverse problem, fitting a model, or training a model (Stouffer 2019). In this context, the goal of  
282 using a model is to estimate the parameters,  $\theta$ , that best explain a set of empirical observations,  
283  $\{\hat{x}, \hat{y}\}$ . In some cases, these parameter values are themselves of interest (e.g., the strength of  
284 selection, intrinsic growth rate, dispersal distance), but in others cases, the goal is to compare a  
285 set of competing models  $f_1, f_2, \dots$  to determine which provides the most parsimonious explana-  
286 tion for a dataset. The quantitative representation of “effects” in these models—the influence of  
287 each input on the output—is often assumed to be linear, and within the frequentist world-view,  
288 the goal is often to determine if the coefficient corresponding with an input is non-zero to deter-  
289 mine its “significance” (often different from its ecological relevance; Martínez-Abraín 2008) in  
290 influencing the outcome.

291 Models designed for inference have utility—descriptive models of networks can reveal under-  
292 lying mechanisms that structure ecological communities, given a proper null model (Connor,  
293 Barberán, and Clauzet 2017). However, in order for ecology to develop as a predictive science  
294 (Evans, Norris, and Benton 2012), interest has grown in developing models that are used not  
295 just for description of data, but also for prediction. Predictive models are based in *the forward*  
296 *problem*, where the aim is to predict new values of the output  $y$  given an input  $x$  and our estimate  
297 value of  $\theta$  (Stouffer 2019). Because the forward problem relies on an estimate of  $\theta$ , then, the  
298 problem of inference is nested within the forward problem (fig. 3): working towards a predictive  
299 view of ecological networks will give us the needed tools to further our understanding of them.

300

[Figure 3 about here.]

301 **What do you need to build a predictive model?**

302 To build a predictive model, one needs the following: first, **data**, split into features  $\hat{x}$  and labels  
 303  $\hat{y}$  (fig. 3). Second, a **model**  $f$ , which maps features  $x$  to labels  $y$  as a function of parameters  $\theta$ ,  
 304 i.e.  $y = f(x, \theta)$ . Third, a **loss function**  $L(\hat{y}, y)$ , which describes how far a model's prediction  
 305  $y$  is from an empirical value  $\hat{y}$ . Lastly, **priors** on parameters,  $P(\theta)$ , which describe the mod-  
 306 eller's *a priori* belief about the value of the parameters; rather than making an analysis implicit,  
 307 specifying priors has the merit of making the modeller's assumptions explicit, which is a most  
 308 desirable feature when communicating predictions to stakeholders (Spiegelhalter et al. 2000).  
 309 Often an important step before fitting a model is feature engineering: adjusting and reworking  
 310 the features to better uncover feature-label relationships (Kuhn and Johnson 2019). This can  
 311 include projecting the features into a lower dimensional space, as we did through a probabilistic  
 312 PCA in the case study, or removing the covariance structure using a Whitening approach. Then,  
 313 when a model is fitted (synonymous with parameter inference or the inverse problem, see fig. 3),  
 314 a fitting algorithm attempts to estimate the values of  $\theta$  that minimises the mean value of loss  
 315 function  $L(\hat{y}, y)$  for all labels  $\hat{y}$  in the provided data  $Y$ . In a Bayesian approach, this typically  
 316 relies on drawing candidate parameter values from priors and applying some form of sampling  
 317 to generate a posterior estimate of parameters,  $P(\theta|\hat{x}, \hat{y})$ . In the training of neural networks,  
 318 this usually involves some form of error back-propagation across the edges in order to tune their  
 319 weights, and the biases of each nodes.

320 **How do we validate a predictive model?**

321 After we fit a model, we inevitably want to see how “good” (meaning, “fit for purpose”) it is.  
 322 This process can be divided into two parts: (i) model selection, where the modeller chooses  
 323 from a set of possible models and (ii) model assessment, where the modeller determines the  
 324 performance characteristics of the chosen model (Hastie, Tibshirani, and Friedman 2009).  
 325 In the context of *model selection*, a naïve initial approach is to simply compute the average error

326 between the model’s prediction and the true data we have, and choose the model with the smallest  
327 error—however this approach inevitably results in *overfitting*. One approach to avoid overfitting  
328 is using information criteria (e.g., AIC, BIC, MDL) based around the heuristic that good mod-  
329 els maximise the ratio of information provided by the model to the number of parameters it has.  
330 However, when the intended use-case of a model is prediction the relevant form of validation  
331 is *predictive accuracy*, which should be tested with *cross-validation*. Cross-validation meth-  
332 ods divide the original dataset into two—one which is used to fit the model (called the *training*  
333 set) and one used to validate its predictive accuracy on the data that it hasn’t “seen” yet (called  
334 the *test* set) (Bishop 2006). This procedure is often repeated across different test and training  
335 subdivisions of the dataset (either picked randomly or stratified by some criteria, like balance  
336 between positive and negative interactions in the case study) to determine the uncertainty asso-  
337 ciated with our measurement due to our choice of test and training sets (Arlot and Celisse 2010),  
338 in the same conceptual vein as data bootstrapping: the mean value of the validation metric gives  
339 an overall estimate of its performance, and the variance around this mean represents the effect of  
340 using different data for training and testing. In a robust model/dataset combination, we expect  
341 this variance to be low, although there are no prescriptive guidelines as to how little variance  
342 is acceptable; the choice of whether to use a model is often left to the best judgement of the  
343 modeller.

344 We still have to define what *predictive accuracy* means in the context of interaction network  
345 prediction. In the proof-of-concept, we used a neural-network to perform binary classification  
346 by predicting the presence/absence of an interaction between any two species. There are two  
347 ways for the model to be right: the model predicts an interaction and there is one (a *true positive*  
348 (TP)), or the model predicts no interaction and there isn’t one (a *true negative* (TN)). Similarly,  
349 there are two ways for the model to be wrong: the model predicts an interaction which does not  
350 exist (*false positive* (FP)), or the model predicts no interaction but it does exist (*false negative*  
351 (FN)).

352 A naïve initial approach to measure how well a model does is *accuracy*, i.e. the proportion  
353 of values it got correct. However, consider what we know about interaction networks: they  
354 are often very sparse, with connectance usually below a third (Cohen, Briand, and Newman

355 1990). If we build a model that always guesses there will be no interaction between two species,  
356 it will be correct in the majority of cases because the majority of potential interactions in a  
357 network typically do not exist. Therefore this “empty-matrix” model would always have an  
358 *accuracy* of  $1 - C$ , where  $C$  is the observed connectance, which would almost always be greater  
359 than 50%. Understanding model performance within sensitivity-specificity space may be more  
360 informative, where sensitivity evaluates how good the model is at predicting true interactions  
361 (True Positive Rate) and specificity refers to the prediction of true “non-interactions” (True  
362 Negative Rate). It must be noted that in ecological networks, there is no guarantee that the “non-  
363 interactions” (assumed true negatives) in the original dataset are indeed true negatives (Jordano  
364 2016a, 2016b). This can result in the positive/negative values, and the false omission/discovery  
365 being artificially worse, and specifically decrease our confidence in predicted interactions.

366 In response to the general problem of biases in classifiers, many metrics have been proposed  
367 to measure binary-classifiers (Gu, Zhu, and Cai 2009; Drummond and Holte 2006) and are  
368 indicative of how well the model performs with regards to some aspect of accuracy, sensitivity,  
369 specificity and/or precision (tbl. 1). Ultimately the choice of metric will depend on the intended  
370 use of the model: there is not a single definition of “success,” but rather different interpretation  
371 of what sources of error are acceptable for a given application.

Table 1: Overview of the validation statistics applied to the case study, alongside the criteria indicating a successful classifier and a guide to interpretation of the values. Taken together, these validation measures indicate that the model performs well, especially considering that it is trained from a small volume of data.

Name	Value	Success	Description
Random accuracy	0.56		Fraction of correct predictions if the classifier is random
Accuracy	0.81	$\rightarrow 1$	Observed fraction of correct predictions
Balanced accuracy	0.80	$\rightarrow 1$	Average fraction of correct positive and negative predictions
True Positive Rate	0.77	$\rightarrow 1$	Fraction of interactions predicted

Name	Value	Success	Description
True Negative Rate	0.83	→ 1	Fraction of non-interactions predicted
False Positive Rate	0.16	→ 0	Fraction of non-interactions predicted as interactions
False Negative Rate	0.22	→ 0	Fraction of interactions predicted as non-interactions
ROC-AUC	0.86	→ 1	Proximity to a perfect prediction (ROC-AUC=1)
Youden's J	0.60	→ 1	Informedness of predictions (trust in individual prediction)
Cohen's $\kappa$	0.58	≥ 0.5	
Positive Predictive Value	0.66	→ 1	Confidence in predicted interactions
Negative Predictive Value	0.89	→ 1	Confidence in predicted non-interactions
False Omission Rate	0.10	→ 0	Expected proportion of missed interactions
False Discovery Rate	0.33	→ 0	Expected proportion of wrongly imputed interactions

<sup>372</sup> In the machine learning literature, a common way of visualising this extensive list of possible  
<sup>373</sup> metrics is through the use of ROC (receiver-operating-characteristic; False Positive Rate on the  
<sup>374</sup> x-axis, and True Positive Rate on the y-axis) and PR (precision-recall; True-Positive-Rate on  
<sup>375</sup> the x-axis, Positive-predictive-value on the y-axis) curves (see fig. 1). These curves are gen-  
<sup>376</sup> erated by considering a continuum of thresholds of classifier acceptance, and computing the  
<sup>377</sup> values of ROC/PR metrics for each value of the threshold. The area-under-the-curve (AUC)  
<sup>378</sup> is then used as a validation metric and are typically called AUC-ROC (Area-Under-the-Curve

379 Receiver-Operator-Curve) and AUC-PR (Area-Under-the-Curve Precision-Recall) (e.g. ROC-  
380 AUC in [tbl. 1](#)). These measures have the unstated assumption that the training and testing set are  
381 “correct,” or at least correct enough that the number of true/false positive/negatives are meaning-  
382 ful; although should this assumption be true, there would be no need for any predictive approach  
383 – but it is a well established fact that machine learning systems are resilient to even relatively  
384 high uncertainties in the data (Halevy, Norvig, and Pereira 2009).

### 385 Networks and interactions as predictable objects

#### 386 Why predict networks and interactions at the same time?

387 Ecological networks are quite sparse, and larger networks tend to get sparser (MacDonald,  
388 Banville, and Poisot 2020); in other words, although networks are composed of a set of in-  
389 teractions between species pairs, they also form a much larger set of species pairs that do not  
390 interact. If we aim to predict the structure of networks from the “bottom-up”— by consider-  
391 ing each pairwise combination of  $S$  different species—we are left with  $S^2$  interaction values  
392 to estimate, a majority of which will be 0. Instead, we can use our existing understanding of  
393 the mechanisms that structure ecological networks to whittle down the set of feasible adjacency  
394 matrices, thereby reducing the amount of information we must predict, and making the problem  
395 of predicting interactions less daunting. The processes that structure ecological networks do not  
396 only occur at the scale of interactions—there are also processes at the network level which limit  
397 what interactions (or how many) are realistic. The realised structure of a network is the synthesis  
398 of the interactions forming the basis for network structure, and the network structure refining the  
399 possible interactions—“Part makes whole, and whole makes part” (Levins and Lewontin 1987).

400 Another argument for the joint prediction of networks and interactions is to reduce circularity  
401 and biases in the predictions. As an example, models like linear filtering (Stock et al. 2017) gen-  
402 erate probabilities of non-observed interactions existing, but do so based on measured network  
403 properties. Some recent models make interaction-level predictions (e.g. Gravel et al. 2019);  
404 these are not unlike stacked species distribution models, which are individually fit, but collec-  
405 tively outperformed by joint models or rule-based models (Zurell et al. 2020). By relying on

406 adequate testing of model performance of biases (i.e. optimising not only accuracy, but pay-  
407 ing attention to measures like false discovery and false omission rates), and developing models  
408 around a feedback loop between network and interaction prediction, it is likely that the quality  
409 of the predicted networks will be greatly improved compared to current models.

410 **What network properties should we use to inform our predictions of interactions?**

411 There are many dimensions of network structure (Delmas et al. 2018), yet there are two argu-  
412 ments to support basing network prediction around a single property: *connectance* (the ratio of  
413 actual edges to possible edges in the network). First, connectance is ecologically informative—  
414 it relates to resilience to invasion (Baiser, Russell, and Lockwood 2010; Smith-Ramesh, Moore,  
415 and Schmitz 2016), can increase robustness to extinction in food webs (J. Dunne, Williams, and  
416 Martinez 2002), while decreasing it in mutualistic networks (Vieira and Almeida-Neto 2015),  
417 and connectance relates to network stability (Landi et al. 2018). Second, most (if not all) net-  
418 work properties covary with connectance (Poisot and Gravel 2014; J. A. Dunne, Williams, and  
419 Martinez 2002).

420 Within the network science literature, there are numerous methods for predicting edges based  
421 on network properties (e.g., block models (Yen and Larremore 2020) based on modularity, hi-  
422 erarchical models (Kawakatsu et al. 2021) based on embedding, etc.). However, in the context  
423 of species interaction networks, these properties often covary with connectance. As a result we  
424 suggest that using connectance as the primary property of interest is most likely to be practical  
425 to formulate at the moment. We have models to estimate species richness over space (Jenkins,  
426 Pimm, and Joppa 2013), and because we can predict connectance from species richness alone  
427 (MacDonald, Banville, and Poisot 2020), we can then derive distributions of network properties  
428 from richness estimates, that can serve to penalise further models that formulate their predictions  
429 at the scale of each possible interaction.

430 **How do we predict how species that we have never observed together will interact?**

431 A neutral approach to ecological interactions would assume the probability of an interaction to  
432 mirror the relative abundance of both species, and would be unaffected by trait variation (Poisot,  
433 Stouffer, and Gravel 2015; Pichler et al. 2020); more accurately, a neutral assumption states  
434 that the relative abundances are sufficient to predict the structure of networks, and this view is  
435 rather well supported in empirical and theoretical systems (Canard et al. 2012, 2014). However,  
436 functional-trait based proxies could enable better predictions of ecological interactions (Cirtwill  
437 and Eklöf 2018; Cirtwill et al. 2019; Bartomeus et al. 2016; Bartomeus 2013). Selection  
438 on functional traits could cause interactions to be conserved at some evolutionary scales, and  
439 therefore predictions of interaction could be informed by phylogenetic analyses (Davies 2021;  
440 Elmasri et al. 2020; Gómez, Verdú, and Perfectti 2010). Phylogenetic matching in bipartite  
441 networks is consistent across scales (Poisot and Stouffer 2018), even in the absence of strong  
442 selective pressure (Coelho, Rodrigues, and Rangel 2017).

443 A separate family of methods are based on network embedding (as in the proof-of-concept). A  
444 network embedding projects each node of the network into a lower-dimensional latent space.  
445 Previous explorations of the dimensionality of food webs have revealed that a reduced number  
446 of dimensions (7) was sufficient to capture most of their structure (Eklöf et al. 2013); however,  
447 recent quantifications of the complexity of the embedding space of bipartite ecological networks  
448 found a consistent high complexity (Strydom, Dalla Riva, and Poisot 2021), suggesting that  
449 the precise depth of embedding required may vary considerably across systems. Embeddings  
450 enables us to represent the structure of a network, which previously required the  $S^2$  dimensions  
451 of an adjacency matrix, with a smaller number of dimensions. The position of each node in this  
452 lower dimensional space is then treated as a latent measurement corresponding to the role of that  
453 species in the network (e.g. Poisot, Ouellet, et al. 2021, where a network of about 1500 species  
454 was most accurately described using 12 dimensions). Species close together in the latent space  
455 should interact with similar set of species (Rossberg et al. 2006; Rohr et al. 2010). However,  
456 these models are sensitive to sampling biases as they are limited to species for which there is  
457 already interaction data, and as a result a methodological breakthrough is needed to extend these  
458 models to species for which there is little or no interaction data.

459 **How do we quantify interaction strength?**

460 Species interaction networks can also be used as a means to quantify and understand *interaction*  
461 *strength*. Interaction strength, unlike the qualitative presence or absence of an interaction, is a  
462 continuous measurement which attempts to quantify the effect of one species on another. This  
463 results in weighted networks representing different patterns of ‘flows’ between nodes – which  
464 can be modelled in a variety of ways (Borrett and Scharler 2019). Interaction strength can gen-  
465 erally be divided into two main categories (as suggested by Berlow et al. (2004)): 1) the strength  
466 of an interaction between individuals of each species, or 2) the effect that changes in one species  
467 population has on the dynamics of the other species. It can be measured as the effect over a pe-  
468 riod of time (in the units of biomass or energy flux (Barnes et al. 2018; Brown et al. 2004)) or the  
469 relative importance of one species on another (Heleno et al. 2014; Berlow et al. 2004; Wootton  
470 and Emmerson 2005). One recurring observation is that networks are often composed of many  
471 weak interactions and few strong interactions (Berlow et al. 2004). The distribution of interac-  
472 tion strength within a network effects its stability (Neutel 2002; Ruiter, Neutel, and Moore 1995)  
473 and functioning (Duffy 2002; José M. Montoya, Rodríguez, and Hawkins 2003), and serves to  
474 benefit multi-species models (Wootton and Emmerson 2005). Alternatively, understanding flow  
475 in modules within networks can aid in understanding the organisation of networks (Farage et al.  
476 2021; Jose M. Montoya and Solé 2002) or the cascading effects of perturbations (Gaiarsa and  
477 Guimarães 2019).

478 In some systems, quantifying interaction strength is relatively straightforward; this includes a  
479 lot of host-parasite systems. For example, freshwater cyprinid fish can be divided in micro-  
480 habitats (fins, skin, digestive system, gill subsections) and the parasites counted in each of these  
481 micro-habitats, giving within-host resolution (Simková et al. 2002); marine sparids and labrids  
482 have similarly been studied this way, see notably (Sasal, Niquil, and Bartoli 1999; Desdevives  
483 2006; Morand et al. 2002). In some cases, within-host assessments of interaction strengths  
484 can reveal macro-ecological events, like in the conservatism of micro-habitat use in amphibian  
485 hosts by helminths (Badets et al. 2011). Even ectoparasites can provide reliable assessments  
486 of interaction strength; for example, when rodent hosts are minimally disturbed during capture,  
487 fine combing of their fur will result in exhaustive ectoparasites inventories (Hadfield et al. 2014;

488 Karbowiak et al. 2019; Matthee et al. 2020; Sánchez et al. 2014; E. R. Dickinson, Millins, and  
489 Biek 2020). Parasites have the desirable property of usually remaining intact within their host  
490 during the interaction, as opposed to prey items as can be recovered through *e.g.* gut content  
491 analysis or stable isotopes (Macías-Hernández et al. 2018; Schmid-Araya et al. 2016). As  
492 network ecology is starting to explore the use of predictive models, leading up to forecasting,  
493 we argue that host-parasite systems can provide data that are reliable and trustworthy enough  
494 that they can become the foundations for methodological development and benchmark studies,  
495 thereby providing more information about host-parasite systems and supporting the technical  
496 development of the field.

497 Yet in most situations, much like quantifying the occurrence of an interaction, quantifying in-  
498 teraction *strength* in the field is challenging in the majority of systems, and one must often  
499 rely on proxies. In some contexts, interaction strength can be estimated via functional foraging  
500 (Portalier et al. 2019), where the primary basis for inferring interaction is foraging behaviour  
501 like searching, capture and handling times. In food-webs, metabolic based models use body  
502 mass, metabolic demands, and energy loss to infer energy fluxes between organisms (Yodzis  
503 and Innes 1992; Berlow et al. 2009). In addition, food-web energetics models can be incorpo-  
504 rated at various resolutions for a specific network, ranging from individual-based data to more  
505 lumped data at the species level or trophic group, depending on data availability (Barnes et al.  
506 2018; Berlow et al. 2009). Taken together, these considerations impose too many constraints  
507 on predicting continuous interaction strength at the moment, resulting in our primary focus in  
508 binary present/absent interactions within this manuscript.

## 509 **How do we determine what interaction networks are feasible?**

510 For several decades, ecologists have aimed to understand how networks of many interacting  
511 species persist through time. The diversity-stability paradox, first explored by May (1974),  
512 shows that under a neutral set of assumptions ecological networks should become decreasingly  
513 stable as the number of species increases. Yet, in the natural world we observe networks of  
514 interactions that consist of far more species than May's model predicts (Albouy et al. 2019).  
515 As a result, understanding what aspects of the neutral assumptions of May's model are incor-

rect has branched many investigations into the relationship between ecological network structure and persistence (Allesina and Tang 2012). These assumptions can be split into dynamical assumptions and topological assumptions. Topologically, we know that ecological networks are not structured randomly. Some properties, like the aforementioned connectance, are highly predictable (MacDonald, Banville, and Poisot 2020). Generative models of food-webs (based on network embeddings) fit empirical networks more effectively than random models (Allesina, Alonso, and Pascual 2008). These models have long used allometry as a single-dimensional niche space—naturally we want to extend this to traits in general. The second approach to stability is through *dynamics*. Early models of community dynamics rely on the assumption of linear interaction effects, but in recent years models of bioenergetic community dynamics have shown promise in basing our understanding of energy flow in food-webs in the understood relationship between allometry and metabolism (Delmas et al. 2017). An additional consideration is the multidimensional nature of “stability” and “feasibility” (e.g. resilience to environmental change vs extinctions) (Domínguez-García, Dakos, and Kéfi 2019) and how different disturbances propagate across levels of biological organisation (Kéfi et al. 2019; Gravel, Massol, and Leibold 2016). Recent approaches such as structural stability (Saavedra et al. 2017; Ferrera, Pascual-García, and Bastolla 2016) allow us to think of network feasibility in rigorous mathematical terms, which may end up as usable parameters to penalise network predictions.

#### 534 **What taxonomic scales are suitable for the prediction of species interactions?**

535 If we use different trait-based proxies to predict potential interactions between species the choice  
536 of such proxies should be theoretically linked to the taxonomic and spatial scale we are using in  
537 our prediction (Wiens 1989). At some scales we can use morphological traits of co-occurring  
538 species to assess the probability of interaction between them (Bartomeus et al. 2016). On  
539 broader taxonomic scales we can infer interaction probability through the phylogenetic distance,  
540 assuming that functional traits themselves are conserved (Gómez, Verdú, and Perfectti 2010).  
541 In this case, we can think of the probability that one species will interact with another as the  
542 distance between them in niche-space (Desjardins-Proulx et al. 2017), and this can be modelled  
543 by simulating neutral expectations of trait variation on phylogenetic trees (Davies 2021). At the

544 narrowest scales, we may be interested in predicting behavioural traits like foraging behaviour  
545 (Bartomeus et al. 2016), and at this scale we may need to consider abundance's effect on the  
546 probability of an encounter (Wells and O'Hara 2013).

547 **What about indirect and higher-order interactions?**

548 Although network ecology often assumes that interactions go strictly from one node to the other,  
549 the web of life is made up of a variety of interactions. Indirect interactions—either higher-order  
550 interactions between species, or interaction strengths that themselves interact — have gained  
551 interest in recent years (Golubski et al. 2016; Golubski and Abrams 2011). One mathemat-  
552 ical tool to describe these situations is hypergraphs: hypergraphs are the generalisation of a  
553 graph, allowing a broad yet manageable approach to complex interactions (Carletti, Fanelli, and  
554 Nicoletti 2020), by allowing for particular interactions to occur beyond a pair of nodes. An ad-  
555 dditional degree of complexity is introduced by multi-layer networks (Hutchinson et al. 2019).

556 Multi-layer networks include edges across “variants” of the networks (timepoints, locations, or  
557 environments). These can be particularly useful to account for the metacommunity structure  
558 (Gross et al. 2020), or to understand how dispersal can inform conservation action (Albert et  
559 al. 2017). Ecological networks are intrinsically multi-layered (Pilosof et al. 2017). However,  
560 *prima facie*, increasing the dimensionality of the object we need to predict (the multiple layers  
561 rather than a single network) makes the problem more complicated. Yet, multi-layer approaches  
562 improve prediction in social networks (Jalili et al. 2017; Najari et al. 2019; Yasami and Safaei  
563 2018), and they may prove useful in network ecology going forward.

564 **Space**

565 Although networks were initially used to describe the interactions *within* a community, interest  
566 in the last decade has shifted towards understanding their structure and variation over space  
567 (Trøjelsgaard and Olesen 2016; Baiser et al. 2019), and has established network ecology as an  
568 important emerging component of biogeography and macroecology.

569 **How much do networks vary over space?**

570 Networks can vary across space either in their structural properties (e.g. connectance or degree  
571 distribution) or in their composition (identity of nodes and edges). Interestingly, variation in  
572 the structural properties of ecological networks primarily responds to changes in the size of the  
573 network. The number of links in ecological networks scales with the number of species (Mac-  
574 Donald, Banville, and Poisot 2020; Brose et al. 2004), and connectance and size drive the rest  
575 of network structure (Poisot and Gravel 2014; J. A. Dunne, Williams, and Martinez 2002; Riede  
576 et al. 2010). Species turnover in space results in changes in the composition of ecological net-  
577 works. But, this is not the only reason network composition varies (Poisot, Stouffer, and Gravel  
578 2015). Intraspecific variation can result in interaction turnovers without changes in species com-  
579 position (Bolnick et al. 2011). Similarly, changes in species abundances can lead to variation in  
580 interaction strengths (Canard et al. 2014; Vázquez et al. 2007). Variation in the abiotic environ-  
581 ment and indirect interactions (Golubski et al. 2016) could modify the occurrence and strength  
582 of individual interactions. Despite this, empirical networks tend to share a common backbone  
583 (Mora et al. 2018) and functional composition (Dehling et al. 2020) across space.

584 **How do we predict what the species pool at a particular location is?**

585 As the species pool forms the basis for network structure, predicting which species are present  
586 at a particular location is essential to predict networks across space. Species distribution mod-  
587 els (SDMs) are increasingly ubiquitous in macroecology—these models predict the range of a  
588 species based on known occurrences and environmental conditions, such as climate and land  
589 cover (Guisan and Thuiller 2005; Elith et al. 2006). Including interactions or co-occurrences  
590 in SDMs generally improves predictive performance (Wisz et al. 2013). Several approaches  
591 exist to combine multiple SDMs: community assemblage at a particular site can be predicted  
592 either by combining independent single-species SDMs (stacked-SDMs, SSDMs) or by directly  
593 modelling the entire species assemblage and multiple species at the same time (joint SDMs,  
594 JSMDs) (Norberg et al. 2019). Building on the JSMD framework, hierarchical modelling of  
595 species communities (Ovaskainen et al. 2017) has the advantage of capturing processes that  
596 structure communities. Spatially Explicit Species Assemblage Modelling (SESAM) constrains

597 SDM predictions using macro-ecological models (Guisan and Rahbek 2011) — for example,  
598 variation in species richness across space can constrain assemblage predictions (D’Amen et al.  
599 2015).

600 The next step is to constrain distribution predictions using network properties. This builds on  
601 previous calls to adopt a probabilistic view: a probabilistic species pool (Karger et al. 2016),  
602 and probabilistic interactions through Bayesian networks (Staniczenko et al. 2017). Blanchet,  
603 Cazelles, and Gravel (2020) argue that the probabilistic view avoids confusion between inter-  
604 actions and co-occurrences, but that it requires prior knowledge of interactions. This could  
605 potentially be solved through our framework of predicting networks first, interactions next, and  
606 finally the realised species pool.

## 607 **How do we combine spatial and network predictions?**

608 In order to predict networks across space, we need to combine multiple models—one which  
609 predicts what the species pool will be at a given location, and one to predict what interaction  
610 networks composed from this species pool are likely to be (see fig. 2). Both of these models  
611 contain uncertainty, and when we combine them the uncertainty from each model should be  
612 propagated into the combined model. The Bayesian paradigm provides a convenient solution to  
613 this—if we have a chain of models where each model feeds into the next, we can sample from  
614 the posterior of the input models. A different approach is *ensemble modelling* which combines  
615 the predictions made by several models, where each model is predicting the same thing (Parker  
616 2013). Error propagation, an important step in building any ecological model, describes the  
617 effect of the uncertainty of input variables on the uncertainty of output variables (Draper 1995;  
618 Parysow, Gertner, and Westervelt 2000). Benke et al. (2018) identifies two broad approaches to  
619 model error propagation: analytically using differential equations or stochastically using Monte-  
620 Carlo simulation methods. Errors induced by the spatial or temporal extrapolation of data also  
621 need to be taken into account when estimating the uncertainty of a model’s output (Peters and  
622 Herrick 2004).

623 **Time**

624 **Why should we forecast species interaction networks?**

625 Forecasting species interactions are critical for informing ecosystem management (Harvey et  
626 al. 2017) and systematic conservation prioritisation (Pollock et al. 2020), and for anticipating  
627 extinctions and their consequences (McDonald-Madden et al. 2016; McWilliams et al. 2019).  
628 Ecological interactions shape species distributions at both local and broad spatial scales, and  
629 including interactions in SDM models typically improves predictive performance (M. B. Araújo  
630 and Luoto 2007; Wisz et al. 2013; Pigot and Tobias 2013). However, these tend to rely on ap-  
631 proaches involving estimating pairwise dependencies based on co-occurrence, using surrogates  
632 for biotic-interaction gradients, and hybridising SDMs with dynamic models (Wisz et al. 2013).  
633 Most existing models to predict the future distribution of species ignore interactions (Urban et  
634 al. 2016). Changes in species ranges and phenology will inevitably create spatiotemporal mis-  
635 matches and affect encounter rates between species (Gilman et al. 2010), which will further  
636 shift the distribution of species across space. New interactions will also appear between species  
637 that are not currently co-occurring (Gilman et al. 2010). Only by forecasting how species will  
638 interact can we hope to have an accurate portrait of how biodiversity will be distributed under  
639 the future climate.

640 Forecasting how climate change will alter biodiversity is also crucial for maximising conserva-  
641 tion outcomes. Improving SDMs through interactions is crucial for conservation, as nearly 30%  
642 of models in SDM studies are used to assess population declines or landscape ability to support  
643 populations (M. B. Araújo et al. 2019). Reliable predictions about how ecological networks  
644 will change over time will give us critical information that could be communicated to decision-  
645 makers and the scientific community about what future environmental risks we are awaiting and  
646 how to mitigate them (Kindsvater et al. 2018). Not only this, but how biodiversity is struc-  
647 tured influences the functioning of the whole ecosystem, community stability and persistence  
648 (Thompson et al. 2012; Stouffer and Bascompte 2010). Will climate change impact the distri-  
649 bution of network properties (e.g. connectance)? If so, which regions or species groups need  
650 special conservation efforts? These overarching questions are yet to be answered (but see Albouy

651 et al. 2013; Kortsch et al. 2015; Hattab et al. 2016). We believe that the path toward forecasting  
652 ecological networks provides useful guidelines to ultimately better predict how climate change  
653 will affect the different dimensions of biodiversity and ecosystem functioning.

654 **How do we turn a predictive model into a forecasting model?**

655 On some scales, empirical time-series encode enough information about ecological processes  
656 for machine-learning approaches to make accurate forecasts. However, there is an intrinsic limit  
657 to the predictability of ecological time-series (Pennekamp et al. 2019). A forecast inherently has  
658 a *resolution limit* in space, time, and organisation. For example, one could never hope to predict  
659 the precise abundance of every species on Earth on every day hundreds of years into the future.  
660 There is often a trade-off between the resolution and horizon of forecast, e.g., a lower resolution  
661 forecast, like primary production will be at a maximum in the summer, is likely to be true much  
662 further into the future than a higher resolution forecast. If we want to forecast the structure  
663 of ecological networks beyond the forecasting horizon of time-series based methods, we need  
664 forecasts of our predictive model’s inputs—a forecast of the distribution of both environmental  
665 conditions and the potential species pool across space (fig. 3).

666 **How can we validate a forecasting model?**

667 Often the purpose of building a forecasting model is to inform *present* action (Dietze et al.  
668 2018). Yet, the nature of forecasting—trying to predict the future—is that you can only know  
669 if a forecast is “right” once it is too late to change it. If we want to maximise the chance that  
670 reality falls within a forecasting model’s predictions, there are two directions to approach this  
671 problem: the first is to extend model validation techniques to a forecasting context, and the  
672 second is to attempt to maximise the amount of uncertainty in the forecast without compromising  
673 its resolution. Cross-validation (see *How do we validate a predictive model?*) can be used to  
674 test the efficacy of a forecasting model. Given a time-series of  $N$  observations, a model can  
675 iteratively be trained on the first  $n$  time-points of data, and the forecasting model’s accuracy can  
676 be evaluated on the remaining time-points it hasn’t “seen” (Bishop 2006). This enables us to  
677 understand both how much temporal data is required for a model to be robust, and also enables

678 us to explore the *forecasting horizon* of a process. Further, this approach can also be applied in  
679 the opposite temporal direction—if we have reliable data from the past, “hindcasting” can also  
680 be used to test a forecast’s robustness.

681 However, these methods inevitably bump into a hard-limitation on what is feasible for a forecast-  
682 ing model. The future is uncertain. Any empirical time-series we use to validate a model was  
683 collected in past conditions that may not persist into the future. Any system we wish to forecast  
684 will undergo only one of many possible scenarios, yet we can only observe the realised outcome  
685 of the system under the scenario that actually unfolds. It is therefore impossible to assess the  
686 quality of a forecasting model in scenarios that remain hypothetical. If the goal is to maximise  
687 the probability that reality will fall within the forecast’s estimates, forecasts should incorporate  
688 as much uncertainty about the future scenario as possible—one way to do this is ensemble mod-  
689 elling (Parker 2013). However, as we increase the amount of uncertainty we incorporate into a  
690 forecasting model, the resolution of the forecast’s predictions could shrink (Lei and Whitaker  
691 2017), and therefore the modeller should be mindful of the trade-off between resolution and ac-  
692 curacy when developing any forecast. Finally, ensemble models are not guaranteed to give more  
693 accurate results: for example, Becker et al. (2020) noted that the ensemble model outperforms  
694 the best-in-class models, which should be taken as an indication that careful model building and  
695 selection is of the utmost importance when dealing with a problem as complex as the prediction  
696 of species interactions.

## 697 Conclusion: why should we predict species interaction networks?

698 Because we almost can, and because we definitely should.  
699 A better understanding of species interactions, and the networks they form, would help unify the  
700 fields of community, network, and spatial ecology; improve the quantification of the functional  
701 relationships between species (Dehling and Stouffer 2018; O’Connor et al. 2020); re-evaluate  
702 metacommunities in light of network structure (Guzman et al. 2019); and enable a new line of  
703 research into the biogeography of species interactions (Massol et al. 2017; Braga et al. 2019)  
704 which incorporates a synthesis of both Eltonian and Grinnellian niche (Gravel et al. 2019). Fur-

ther, the ability to reliably predict and forecast species interactions would inform conservation efforts for protecting species, communities, and ecosystems. Integration of species interactions into the assessment of vulnerability to climate change is a needed methodological advancement (Foden and Young 2016). International panels draw on models to establish scientific consensus (M. B. Araújo et al. 2019), and they can be improved through more effective prediction of species distributions and interactions (Syfert et al. 2014). Further, recent studies argue for a shift in focus from species to interaction networks for biodiversity conservation to better understand ecosystem processes (Harvey et al. 2017).

We should invest in network prediction because the right conditions to do so reliably and rapidly are beginning to emerge. Given the possible benefits to a variety of ecological disciplines that would result from an increased ability to predict networks, we feel strongly that the research agenda we outline here should be picked up by the community. Although novel technologies are bringing massive amounts of data to some parts of ecology (primarily environmental DNA and remote sensing, but now more commonly image analysis and bioacoustics), it is even more important to be intentional about *reconciling* data. This involves not only the work of understanding the processes encoded within data, but also the groundwork of developing pipelines to bridge the ever-expanding gap between “high-throughput” and “low-throughput” sampling methods. An overall increase in the volume of data will not result in an increase of our predictive capacity as long as this data increase is limited to specific aspects of the problem. In the areas we highlight in fig. 2, many data steps are still limiting: documenting empirical interactions is natural history work that doesn’t lend itself to systematic automation; expert knowledge is by design a social process that may be slightly accelerated by text mining and natural language processing (but is not yet, or not routinely or at scale). These limitations are affecting our ability to reconstruct networks.

But the tools to which we feed these data, incomplete as they may be, are gradually getting better; that is, they can do predictions faster, they handle uncertainty and propagate it well, and they can accommodate data volumes that are lower than we may expect (Pichler et al. 2020). It is clear attempting to predict the structure of ecological networks at any scale is a methodological and ecological challenge; yet it will result in qualitative changes in our understanding of com-

plex adaptive systems, as well as changes to our ability to leverage information about network structure for conservation decision. It is perhaps even more important to forecast the structure of ecological networks because it is commonly neglected as a facet of biodiversity that can (and should) be managed. In fact, none of the Aichi targets mention biostructure or its protection, despite this being recognised as an important task (McCann 2007), either implicitly or explicitly. Being able to generate reliable datasets on networks in space or time will make this information more actionable.

**Acknowledgements:** We acknowledge that this study was conducted on land within the traditional unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwininiwak nations. TS, NF, TP are funded by a donation from the Courtois Foundation; FB, NF, and TP are funded by IVADO; BM is funded by the NSERC Alexander Graham Bell Canada Graduate Scholarship and the FRQNT master's scholarship; FB, GD, NF, and GH are funded by the NSERC BIOS<sup>2</sup> CREATE program; GD is funded by the FRQNT doctoral scholarship; DC, TS, LP, and TP are funded by the Canadian Institute of Ecology & Evolution; this research was enabled in part by support provided by Calcul Québec ([www.calculquebec.ca](http://www.calculquebec.ca)) and Compute Canada ([www.computecanada.ca](http://www.computecanada.ca)). This work was supported by funding to the Viral Emergence Research Initiative (VERENA) consortium including NSF BII 2021909. AG and MDC are supported in part by the Liber Ero Chair.

## References

- Aguiar, Marcus A. M. de, Erica A. Newman, Mathias M. Pires, Justin D. Yeakel, Carl Boettiger, Laura A. Burkle, Dominique Gravel, et al. 2019. “Revealing Biases in the Sampling of Ecological Interaction Networks.” *PeerJ* 7 (September): e7566. <https://doi.org/10.7717/peerj.7566>.
- Albert, Cécile H., Bronwyn Rayfield, Maria Dumitru, and Andrew Gonzalez. 2017. “Applying Network Theory to Prioritize Multispecies Habitat Networks That Are Robust to Climate and Land-Use Change: Prioritizing a Network for Biodiversity.” *Conservation Biology* 31 (6): 1383–96. <https://doi.org/10.1111/cobi.12943>.

- 761 Albouy, Camille, Philippe Archambault, Ward Appeltans, Miguel B. Araújo, David Beauchesne,  
762 Kevin Cazelles, Alyssa R. Cirtwill, et al. 2019. “The Marine Fish Food Web Is Globally  
763 Connected.” *Nature Ecology & Evolution* 3 (8): 1153–61. <https://doi.org/10.1038/s41559-019-0950-y>.
- 764
- 765 Albouy, Camille, François Guilhaumon, Fabien Leprieur, Frida Ben Rais Lasram, Samuel So-  
766 mot, Roland Aznar, Laure Velez, François Le Loc'h, and David Mouillot. 2013. “Projected  
767 Climate Change and the Changing Biogeography of Coastal Mediterranean Fishes.” Edited  
768 by Peter Pearman. *Journal of Biogeography* 40 (3): 534–47. <https://doi.org/10.1111/jbi.12013>.
- 769
- 770 Albrecht, Jörg, Alice Classen, Maximilian G. R. Vollständt, Antonia Mayr, Neduvoto P. Mollel,  
771 David Schellenberger Costa, Hamadi I. Dulle, et al. 2018. “Plant and Animal Functional  
772 Diversity Drive Mutualistic Network Assembly Across an Elevational Gradient.” *Nature  
773 Communications* 9 (1): 3177. <https://doi.org/10.1038/s41467-018-05610-w>.
- 774
- 775 Allesina, Stefano, David Alonso, and Mercedes Pascual. 2008. “A General Model for Food Web  
776 Structure.” *Science* 320 (5876): 658–61. <https://doi.org/10.1126/science.1156269>.
- 777
- 778 Allesina, Stefano, and Si Tang. 2012. “Stability Criteria for Complex Ecosystems.” *Nature* 483  
779 (7388): 205–8. <https://doi.org/10.1038/nature10832>.
- 780
- 781 Antoniou, Antreas, Amos Storkey, and Harrison Edwards. 2018. “Data Augmentation Gen-  
782 erative Adversarial Networks.” *arXiv:1711.04340 [cs, Stat]*, March. <http://arxiv.org/abs/1711.04340>.
- 783
- 784 Araújo, Márcio S., Paulo R. Guimarães, Richard Svanbäck, Aluisio Pinheiro, Paulo Guimarães,  
785 Sérgio F. dos Reis, and Daniel I. Bolnick. 2008. “Network Analysis Reveals Contrasting  
786 Effects of Intraspecific Competition on Individual Vs. Population Diets.” *Ecology* 89 (7):  
787 1981–93. <https://doi.org/10.1890/07-0630.1>.
- 788
- 789 Araújo, Miguel B., Robert P. Anderson, A. Márcia Barbosa, Colin M. Beale, Carsten F. Dor-  
790 man, Regan Early, Raquel A. Garcia, et al. 2019. “Standards for Distribution Models  
791 in Biodiversity Assessments.” *Science Advances* 5 (1): eaat4858. <https://doi.org/10.1126/sciadv.aat4858>.

- 789 Araújo, Miguel B., and Miska Luoto. 2007. "The Importance of Biotic Interactions for Mod-  
790 elling Species Distributions Under Climate Change." *Global Ecology and Biogeography* 16  
791 (6): 743–53. <https://doi.org/10.1111/j.1466-8238.2007.00359.x>.
- 792 Arlot, Sylvain, and Alain Celisse. 2010. "A Survey of Cross-Validation Procedures for Model  
793 Selection." *Statistics Surveys* 4: 40–79. <https://doi.org/10.1214/09-SS054>.
- 794 Badets, Mathieu, Ian Whittington, Fabrice Lalubin, Jean-Francois Allienne, Jean-Luc Maspimby,  
795 Sophie Bentz, Louis H. Du Preez, et al. 2011. "Correlating Early Evolution of Para-  
796 sitic Platyhelminths to Gondwana Breakup." *Systematic Biology* 60 (6): 762–81. <https://doi.org/10.1093/sysbio/syr078>.
- 797
- 798 Baiser, Benjamin, Dominique Gravel, Alyssa R. Cirtwill, Jennifer A. Dunne, Ashkaan K. Fahimipour,  
799 Luis J. Gilarranz, Joshua A. Grochow, et al. 2019. "Ecogeographical Rules and the Macroe-  
800 cology of Food Webs." *Global Ecology and Biogeography* 28 (9): 1204–18. <https://doi.org/10.1111/geb.12925>.
- 801
- 802 Baiser, Benjamin, Gareth J. Russell, and Julie L. Lockwood. 2010. "Connectance Determines  
803 Invasion Success via Trophic Interactions in Model Food Webs." *Oikos* 119 (12): 1970–76.  
804 <https://doi.org/10.1111/j.1600-0706.2010.18557.x>.
- 805 Barnes, Andrew D., Malte Jochum, Jonathan S. Lefcheck, Nico Eisenhauer, Christoph Scher-  
806 ber, Mary I. O'Connor, Peter de Ruiter, and Ulrich Brose. 2018. "Energy Flux: The Link  
807 Between Multitrophic Biodiversity and Ecosystem Functioning." *Trends in Ecology & Evo-*  
808 *lution* 33 (3): 186–97. <https://doi.org/10.1016/j.tree.2017.12.007>.
- 809 Bartomeus, Ignasi. 2013. "Understanding Linkage Rules in Plant-Pollinator Networks by Using  
810 Hierarchical Models That Incorporate Pollinator Detectability and Plant Traits." *PLoS ONE*  
811 8 (7): e69200. <https://doi.org/10.1371/journal.pone.0069200>.
- 812 Bartomeus, Ignasi, Dominique Gravel, Jason M. Tylianakis, Marcelo A. Aizen, Ian A. Dickie,  
813 and Maud Bernard-Verdier. 2016. "A Common Framework for Identifying Linkage Rules  
814 Across Different Types of Interactions." *Functional Ecology* 30 (12): 1894–1903. <https://doi.org/10.1111/1365-2435.12666>.
- 815
- 816 Becker, Daniel J., Gregory F. Albery, Anna R. Sjodin, Timothée Poisot, Tad A. Dallas, Evan

- 817 A. Eskew, Maxwell J. Farrell, et al. 2020. "Predicting Wildlife Hosts of Betacoronaviruses  
818 for SARS-CoV-2 Sampling Prioritization: A Modeling Study." Preprint. Ecology. <https://doi.org/10.1101/2020.05.22.111344>.
- 820 Benke, K. K., S. Norgn, N. J. Robinson, L. R. Benke, and T. J. Peterson. 2018. "Error Prop-  
821 agation in Computer Models: Analytic Approaches, Advantages, Disadvantages and Con-  
822 straints." *Stochastic Environmental Research and Risk Assessment* 32 (10): 2971–85. <https://doi.org/10.1007/s00477-018-1555-8>.
- 824 Bennett, Alison E., Darren M. Evans, and Jeff R. Powell. 2019. "Potentials and Pitfalls in the  
825 Analysis of Bipartite Networks to Understand Plantmicrobe Interactions in Changing Envi-  
826 ronments." *Functional Ecology* 33 (1): 107–17. <https://doi.org/10.1111/1365-2435.13223>.
- 828 Berlow, Eric L., Jennifer Dunne, Neo D. Martinez, Philip B. Stark, Richard J. Williams, and  
829 Ulrich Brose. 2009. "Simple Prediction of Interaction Strengths in Complex Food Webs."  
830 *Proceedings of the National Academy of Sciences* 106 (1): 187–91. <https://doi.org/10.1073/pnas.0806823106>.
- 832 Berlow, Eric L., Anje-Margiet Neutel, Joel E. Cohen, Peter C. de Ruiter, Bo Ebenman, Mark  
833 Emmerson, Jeremy W. Fox, et al. 2004. "Interaction Strengths in Food Webs: Issues and  
834 Opportunities." *Journal of Animal Ecology* 73 (3): 585–98. <https://doi.org/10.1111/j.0021-8790.2004.00833.x>.
- 836 Bezanson, J., A. Edelman, S. Karpinski, and V. Shah. 2017. "Julia: A Fresh Approach to Numer-  
837 ical Computing." *SIAM Review* 59 (1): 65–98. <https://doi.org/10.1137/141000671>.
- 838 Bishop, Christopher M. 2006. *Pattern Recognition and Machine Learning*. Information Science  
839 and Statistics. New York: Springer.
- 840 Blanchet, F. Guillaume, Kevin Cazelles, and Dominique Gravel. 2020. "Co-Occurrence Is Not  
841 Evidence of Ecological Interactions." *Ecology Letters* 23 (7): 1050–63. <https://doi.org/10.1111/ele.13525>.
- 843 Bohan, David A., Geoffrey Caron-Lormier, Stephen Muggleton, Alan Raybould, and Alireza  
844 Tamaddoni-Nezhad. 2011. "Automated Discovery of Food Webs from Ecological Data

- 845 Using Logic-Based Machine Learning.” *PLOS ONE* 6 (12): e29028. <https://doi.org/10.1371/journal.pone.0029028>.
- 846
- 847 Bohan, David A., Corinne Vacher, Alireza Tamaddoni-Nezhad, Alan Raybould, Alex J. Dum-  
848 brell, and Guy Woodward. 2017. “Next-Generation Global Biomonitoring: Large-Scale,  
849 Automated Reconstruction of Ecological Networks.” *Trends in Ecology & Evolution* 32 (7):  
850 477–87. <https://doi.org/10.1016/j.tree.2017.03.001>.
- 851 Bolnick, Daniel I., Priyanga Amarasekare, Márcio S. Araújo, Reinhard Bürger, Jonathan M.  
852 Levine, Mark Novak, Volker H. W. Rudolf, Sebastian J. Schreiber, Mark C. Urban, and  
853 David A. Vasseur. 2011. “Why Intraspecific Trait Variation Matters in Community Ecol-  
854 ogy.” *Trends in Ecology & Evolution* 26 (4): 183–92. <https://doi.org/10.1016/j.tree.2011.01.009>.
- 855
- 856 Borowiec, Marek L., Paul Frandsen, Rebecca Dikow, Alexander McKeeken, Gabriele Valentini,  
857 and Alexander E. White. 2021. “Deep Learning as a Tool for Ecology and Evolution,” June.  
858 <https://doi.org/10.32942/osf.io/nt3as>.
- 859
- 860 Borrett, Stuart R., and Ursula M. Scharler. 2019. “Walk Partitions of Flow in Ecological Net-  
861 work Analysis: Review and Synthesis of Methods and Indicators.” *Ecological Indicators*  
106 (November): 105451. <https://doi.org/10.1016/j.ecolind.2019.105451>.
- 862
- 863 Braga, J, Laura J. Pollock, C. Barros, J. Montoya, Dominique Gravel, Luigi Maiorano, Gentile  
Francesco Ficetola, Stéphane Dray, and Wilfried Thuiller. 2019. “Spatial Analyses of Multi-  
864 Trophic Terrestrial Vertebrate Assemblages in Europe.” *Global Ecology and Biogeography*  
28 (11): 1636–48.
- 865
- 866 Brose, Ulrich, Annette Ostling, Kateri Harrison, and Neo D. Martinez. 2004. “Unified Spatial  
867 Scaling of Species and Their Trophic Interactions.” *Nature* 428 (6979): 167–71. <https://doi.org/10.1038/nature02297>.
- 868
- 869 Brown, James H., James F. Gillooly, Andrew P. Allen, Van M. Savage, and Geoffrey B. West.  
870 2004. “Toward a Metabolic Theory of Ecology.” *Ecology* 85 (7): 1771–89. <https://doi.org/10.1890/03-9000>.
- 871
- 872 Burkle, L. A., and R. Alarcon. 2011. “The Future of Plant-Pollinator Diversity: Understanding

- 873      Interaction Networks Across Time, Space, and Global Change.” *American Journal of Botany*  
874      98 (3): 528–38. <https://doi.org/10.3732/ajb.1000391>.
- 875      Burkle, L. A., J. C. Marlin, and T. M. Knight. 2013. “Plant-Pollinator Interactions over 120  
876      Years: Loss of Species, Co-Occurrence, and Function.” *Science* 339 (6127): 1611–15.  
877      <https://doi.org/10.1126/science.1232728>.
- 878      Canard, Elsa, N. Mouquet, D. Mouillot, M. Stanko, D. Miklisova, and D. Gravel. 2014. “Empir-  
879      ical Evaluation of Neutral Interactions in Host-Parasite Networks.” *The American Naturalist*  
880      183 (4): 468–79. <https://doi.org/10.1086/675363>.
- 881      Canard, Elsa, Nicolas Mouquet, Lucile Marescot, Kevin J. Gaston, Dominique Gravel, and  
882      David Mouillot. 2012. “Emergence of Structural Patterns in Neutral Trophic Networks.”  
883      *PLOS ONE* 7 (8): e38295. <https://doi.org/10.1371/journal.pone.0038295>.
- 884      Carletti, Timoteo, Duccio Fanelli, and Sara Nicoletti. 2020. “Dynamical Systems on Hyper-  
885      graphs.” *Journal of Physics: Complexity* 1 (3): 035006. <https://doi.org/10.1088/2632-072X/aba8e1>.
- 887      Carpenter, Stephen R. 2002. “Ecological Futures: Building an Ecology of the Long Now.”  
888      *Ecology* 83 (8): 2069–83.
- 889      Cazelles, Kévin, Miguel B. Araújo, Nicolas Mouquet, and Dominique Gravel. 2015. “A Theory  
890      for Species Co-Occurrence in Interaction Networks.” *Theoretical Ecology* 9 (1): 39–48.  
891      <https://doi.org/10.1007/s12080-015-0281-9>.
- 892      Chawla, Nitesh V. 2010. “Data Mining for Imbalanced Datasets: An Overview.” In *Data Mining*  
893      and *Knowledge Discovery Handbook*, edited by Oded Maimon and Lior Rokach, 875–86.  
894      Boston, MA: Springer US. [https://doi.org/10.1007/978-0-387-09823-4\\_45](https://doi.org/10.1007/978-0-387-09823-4_45).
- 895      Christin, Sylvain, Éric Hervet, and Nicolas Lecomte. 2019. “Applications for Deep Learning  
896      in Ecology.” *Methods in Ecology and Evolution* 10 (10): 1632–44. <https://doi.org/10.1111/2041-210X.13256>.
- 898      Chuang, Kangway V., and Michael J. Keiser. 2018. “Adversarial Controls for Scientific Ma-  
899      chine Learning.” *Acs Chemical Biology* 13 (10): 2819–21. <https://doi.org/10.1021/acscchembio.8b00881>.

- 901 Cirtwill, Alyssa R., Anna Eklf, Tomas Roslin, Kate Wootton, and Dominique Gravel. 2019. “A  
902 Quantitative Framework for Investigating the Reliability of Empirical Network Construc-  
903 tion.” *Methods in Ecology and Evolution* 0 (ja). <https://doi.org/10.1111/2041-210X.13180>.
- 905 Cirtwill, Alyssa R., and Anna Eklöf. 2018. “Feeding Environment and Other Traits Shape  
906 Species’ Roles in Marine Food Webs.” *Ecology Letters* 21 (6): 875–84. <https://doi.org/10.1111/ele.12955>.
- 908 Coelho, Marco Túlio Pacheco, João Fabrício Mota Rodrigues, and Thiago F. Rangel. 2017.  
909 “Neutral Biogeography of Phylogenetically Structured Interaction Networks.” *Ecography*  
910 40 (12): 1467–74. <https://doi.org/10.1111/ecog.02780>.
- 911 Cohen, Joel E, Frederic Briand, and Charles Newman. 1990. *Community Food Webs: Data and*  
912 *Theory*. Biomathematics. Berlin Heidelberg: Springer-Verlag.
- 913 Connor, Nora, Albert Barberán, and Aaron Clauset. 2017. “Using Null Models to Infer Mi-  
914 crobial Co-Occurrence Networks.” *PLOS ONE* 12 (5): e0176751. <https://doi.org/10.1371/journal.pone.0176751>.
- 916 Craft, Meggan E. 2015. “Infectious Disease Transmission and Contact Networks in Wildlife  
917 and Livestock.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 370  
918 (1669): 20140107. <https://doi.org/10.1098/rstb.2014.0107>.
- 919 Creer, Simon, Kristy Deiner, Serita Frey, Dorota Porazinska, Pierre Taberlet, W. Kelley Thomas,  
920 Caitlin Potter, and Holly M. Bik. 2016. “The Ecologist’s Field Guide to Sequence-Based  
921 Identification of Biodiversity.” *Methods in Ecology and Evolution* 7 (9): 1008–18. <https://doi.org/10.1111/2041-210X.12574>.
- 923 Cutler, D. Richard, Thomas C. Edwards, Karen H. Beard, Adele Cutler, and Kyle T. Hess. 2007.  
924 “Random Forests for Classification in Ecology.” *Ecology* 88 (11): 2783–92. <https://doi.org/10.1890/07-0539.1>.
- 926 D’Amen, Manuela, Anne Dubuis, Rui F. Fernandes, Julien Pottier, Loïc Pellissier, and Antoine  
927 Guisan. 2015. “Using Species Richness and Functional Traits Predictions to Constrain As-  
928 semblage Predictions from Stacked Species Distribution Models.” *Journal of Biogeography*

- 929 42 (7): 1255–66. <https://doi.org/10.1111/jbi.12485>.
- 930 Davies, T. Jonathan. 2021. “Ecophylogenetics Redux.” *Ecology Letters* n/a. <https://doi.org/10.1111/ele.13682>.
- 932 Dehling, D. Matthias, Guadalupe Peralta, Irene M. A. Bender, Pedro G. Blendinger, Katrin  
933 Böhning-Gaese, Marcia C. Muñoz, Eike Lena Neuschulz, et al. 2020. “Similar Compo-  
934 sition of Functional Roles in Andean Seed-Dispersal Networks, Despite High Species and  
935 Interaction Turnover.” *Ecology* 101 (7). <https://doi.org/10.1002/ecy.3028>.
- 936 Dehling, D. Matthias, and Daniel B. Stouffer. 2018. “Bringing the Eltonian Niche into Func-  
937 tional Diversity.” *Oikos* 127 (12): 1711–23. <https://doi.org/10.1111/oik.05415>.
- 938 Delmas, Eva, Mathilde Besson, Marie-Hélène Brice, Laura A. Burkle, Giulio V. Dalla Riva,  
939 Marie-Josée Fortin, Dominique Gravel, et al. 2018. “Analysing Ecological Networks of  
940 Species Interactions.” *Biological Reviews*, June, 112540. <https://doi.org/10.1111/brv.12433>.
- 942 Delmas, Eva, Ulrich Brose, Dominique Gravel, Daniel B. Stouffer, and Timothée Poisot. 2017.  
943 “Simulations of Biomass Dynamics in Community Food Webs.” *Methods in Ecology and  
944 Evolution* 8 (7): 881–86. <https://doi.org/10.1111/2041-210X.12713>.
- 945 Desdevises, Y. 2006. “Determinants of Parasite Species Richness on Small Taxonomical and  
946 Geographical Scales: Lamellodiscus Monogeneans of Northwestern Mediterranean Sparid  
947 Fish.” *Journal of Helminthology* 80 (3): 235–41.
- 948 Desjardins-Proulx, Philippe, Idaline Laigle, Timothée Poisot, and Dominique Gravel. 2017.  
949 “Ecological Interactions and the Netflix Problem.” *PeerJ* 5 (August): e3644. <https://doi.org/10.7717/peerj.3644>.
- 951 Desjardins-Proulx, Philippe, Timothée Poisot, and Dominique Gravel. 2019. “Artificial Intel-  
952 ligence for Ecological and Evolutionary Synthesis.” *Frontiers in Ecology and Evolution* 7.  
953 <https://doi.org/10.3389/fevo.2019.00402>.
- 954 Dickinson, Eleanor R., Caroline Millins, and Roman Biek. 2020. “Sampling Scale and Season  
955 Influence the Observed Relationship Between the Density of Deer and Questing Ixodes Ricin-  
956 nus Nymphs.” *Parasites & Vectors* 13 (1): 493. <https://doi.org/10.1186/s13071-020-04369-8>.

- 957 Dickinson, Janis L., Benjamin Zuckerberg, and David N. Bonter. 2010. "Citizen Science as an  
958 Ecological Research Tool: Challenges and Benefits." *Annual Review of Ecology, Evolution,  
959 and Systematics* 41 (1): 149–72. <https://doi.org/10.1146/annurev-ecolsys-102209-144636>.
- 960 Dietze, Michael C., Andrew Fox, Lindsay M. Beck-Johnson, Julio L. Betancourt, Mevin B.  
961 Hooten, Catherine S. Jarnevich, Timothy H. Keitt, et al. 2018. "Iterative Near-Term Eco-  
962 logical Forecasting: Needs, Opportunities, and Challenges." *Proceedings of the National  
963 Academy of Sciences* 115 (7): 1424–32. <https://doi.org/10.1073/pnas.1710231115>.
- 964 Domínguez-García, Virginia, Vasilis Dakos, and Sonia Kéfi. 2019. "Unveiling Dimensions  
965 of Stability in Complex Ecological Networks." *Proceedings of the National Academy of  
966 Sciences* 116 (51): 25714–20. <https://doi.org/10.1073/pnas.1904470116>.
- 967 Draper, D. 1995. "Assessment and Propagation of Model Uncertainty." *Journal of the Royal  
968 Statistical Society Series B-Statistical Methodology* 57 (1): 45–97. <https://doi.org/10.1111/j.2517-6161.1995.tb02015.x>.
- 970 Drummond, Chris, and Robert C. Holte. 2006. "Cost Curves: An Improved Method for Visual-  
971 izing Classifier Performance." *Machine Learning* 65 (1): 95–130. <https://doi.org/10.1007/s10994-006-8199-5>.
- 973 Duffy, J. Emmett. 2002. "Biodiversity and Ecosystem Function: The Consumer Connection."  
974 *Oikos* 99 (2): 201–19. <https://doi.org/10.1034/j.1600-0706.2002.990201.x>.
- 975 Dunne, Jennifer. 2006. "The Network Structure of Food Webs." In *Ecological Networks: Link-  
976 ing Structure to Dynamics in Food Webs*, 27–86.
- 977 Dunne, Jennifer A., Richard J. Williams, and Neo D. Martinez. 2002. "Food-Web Structure and  
978 Network Theory: The Role of Connectance and Size." *Proceedings of the National Academy  
979 of Sciences* 99 (20): 12917–22. <https://doi.org/10.1073/pnas.192407699>.
- 980 Dunne, Jennifer, Richard J Williams, and Néo D Martinez. 2002. "Network Structure and  
981 Biodiversity Loss in Food Webs: Robustness Increases with Connectance." *Ecol. Lett.* 5  
982 (4): 558–67.
- 983 Edwards, Felicity A., David P. Edwards, Keith C. Hamer, and Tom M. Fayle. 2021. "Tropical  
984 Land-Use Change Alters Trait-Based Community Assembly Rules for Dung Beetles and

- 985 Birds.” *Oecologia*, February. <https://doi.org/10.1007/s00442-020-04829-z>.
- 986 Eklöf, Anna, Ute Jacob, Jason Kopp, Jordi Bosch, Rocío Castro-Urgal, Natacha P. Chacoff, Bo  
987 Dalsgaard, et al. 2013. “The Dimensionality of Ecological Networks.” *Ecology Letters* 16  
988 (5): 577–83. <https://doi.org/10.1111/ele.12081>.
- 989 Elith, Jane, Catherine H. Graham, Robert P. Anderson, Miroslav Dudík, Simon Ferrier, Antoine  
990 Guisan, Robert J. Hijmans, et al. 2006. “Novel Methods Improve Prediction of Species’  
991 Distributions from Occurrence Data.” *Ecography* 29 (2): 129–51. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>.
- 992 Elmasri, Mohamad, Maxwell J. Farrell, T. Jonathan Davies, and David A. Stephens. 2020. “A  
993 Hierarchical Bayesian Model for Predicting Ecological Interactions Using Scaled Evolution-  
994 ary Relationships.” *The Annals of Applied Statistics* 14 (1): 221–40. <https://doi.org/10.1214/19-AOAS1296>.
- 995 Evans, Matthew R., Ken J. Norris, and Tim G. Benton. 2012. “Predictive Ecology: Systems  
996 Approaches.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 367  
997 (1586): 163–69. <https://doi.org/10.1098/rstb.2011.0191>.
- 1000 Farage, Carmel, Daniel Edler, Anna Eklöf, Martin Rosvall, and Shai Pilosof. 2021. “Identifying  
1001 Flow Modules in Ecological Networks Using Infomap.” *Methods in Ecology and Evolution*  
1002 12 (5): 778–86. <https://doi.org/10.1111/2041-210X.13569>.
- 1003 Ferrera, Antonio, Alberto Pascual-García, and Ugo Bastolla. 2016. “Effective Competition  
1004 Determines the Global Stability of Model Ecosystems.” *Theoretical Ecology*, November,  
1005 1–11. <https://doi.org/10.1007/s12080-016-0322-z>.
- 1006 Foden, Wendy B., and Bruce E. Young. 2016. *IUCN SSC Guidelines for Assessing Species’  
1007 Vulnerability to Climate Change*. IUCN Cambridge, England and Gland, Switzerland.
- 1008 Fontaine, Colin, Paulo R. Guimarães, Sonia Kéfi, Nicolas Loeuille, Jane Memmott, Wim H.  
1009 van der Putten, Frank J. F. van Veen, and Elisa Thébaud. 2011. “The Ecological and Evo-  
1010 lutionary Implications of Merging Different Types of Networks.” *Ecology Letters* 14 (11):  
1011 1170–81. <https://doi.org/10.1111/j.1461-0248.2011.01688.x>.
- 1012 Gaiarsa, Marília P., and Paulo R. Guimarães. 2019. “Interaction Strength Promotes Robustness

- 1013      Against Cascading Effects in Mutualistic Networks.” *Scientific Reports* 9 (1): 676. <https://doi.org/10.1038/s41598-018-35803-8>.
- 1014
- 1015      Gibb, Rory, Gregory F. Albery, Daniel J. Becker, Liam Brierley, Ryan Connor, Tad A. Dallas,  
1016      Evan A. Eskew, et al. 2021. “Data Proliferation, Reconciliation, and Synthesis in Viral  
1017      Ecology.” Preprint. *Ecology*. <https://doi.org/10.1101/2021.01.14.426572>.
- 1018      Gilman, Sarah E., Mark C. Urban, Joshua Tewksbury, George W. Gilchrist, and Robert D. Holt.  
1019      2010. “A Framework for Community Interactions Under Climate Change.” *Trends in Ecol-*  
1020      *ogy & Evolution* 25 (6): 325–31. <https://doi.org/10.1016/j.tree.2010.03.002>.
- 1021      Golubski, Antonio J., and Peter A. Abrams. 2011. “Modifying Modifiers: What Happens When  
1022      Interspecific Interactions Interact?” *Journal of Animal Ecology* 80 (5): 1097–1108. <https://doi.org/10.1111/j.1365-2656.2011.01852.x>.
- 1023
- 1024      Golubski, Antonio J., Erik E. Westlund, John Vandermeer, and Mercedes Pascual. 2016. “Eco-  
1025      logical Networks over the Edge: Hypergraph Trait-Mediated Indirect Interaction (TMII)  
1026      Structure.” *Trends in Ecology & Evolution* 31 (5): 344–54. <https://doi.org/10.1016/j.tree.2016.02.006>.
- 1027
- 1028      Gómez, José M., Miguel Verdú, and Francisco Perfectti. 2010. “Ecological Interactions Are  
1029      Evolutionarily Conserved Across the Entire Tree of Life.” *Nature* 465 (7300): 918–21.  
1030      <https://doi.org/10.1038/nature09113>.
- 1031      Gravel, Dominique, Benjamin Baiser, Jennifer A. Dunne, Jens-Peter Kopalke, Neo D. Mar-  
1032      tinez, Tommi Nyman, Timothée Poisot, et al. 2019. “Bringing Elton and Grinnell Together:  
1033      A Quantitative Framework to Represent the Biogeography of Ecological Interaction Net-  
1034      works.” *Ecography* 42 (3): 401–15. <https://doi.org/10.1111/ecog.04006>.
- 1035      Gravel, Dominique, François Massol, and Mathew A. Leibold. 2016. “Stability and Complexity  
1036      in Model Meta-Ecosystems.” *Nature Communications* 7 (1): 12457. <https://doi.org/10.1038/ncomms12457>.
- 1037
- 1038      Gravel, Dominique, Timothée Poisot, Camille Albouy, Laure Velez, and David Mouillot. 2013.  
1039      “Inferring Food Web Structure from Predatorprey Body Size Relationships.” *Methods in  
1040      Ecology and Evolution* 4 (11): 1083–90. <https://doi.org/10.1111/2041-210X.12103>.

- 1041 Gross, Thilo, Korinna T. Allhoff, Bernd Blasius, Ulrich Brose, Barbara Drossel, Ashkan K.  
1042 Fahimipour, Christian Guill, Justin D. Yeakel, and Fanqi Zeng. 2020. “Modern Models of  
1043 Trophic Meta-Communities.” *Philosophical Transactions of the Royal Society B: Biological*  
1044 *Sciences* 375 (1814): 20190455. <https://doi.org/10.1098/rstb.2019.0455>.
- 1045 Gu, Qiong, Li Zhu, and Zhihua Cai. 2009. “Evaluation Measures of the Classification Perfor-  
1046 mance of Imbalanced Data Sets.” In *Computational Intelligence and Intelligent Systems*,  
1047 edited by Zhihua Cai, Zhenhua Li, Zhuo Kang, and Yong Liu, 51:461–71. Berlin, Heidel-  
1048 berg: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-04962-0\\_53](https://doi.org/10.1007/978-3-642-04962-0_53).
- 1049 Guimarães, Paulo R. 2020. “The Structure of Ecological Networks Across Levels of Organi-  
1050 zation.” *Annual Review of Ecology, Evolution, and Systematics* 51 (1): 433–60. <https://doi.org/10.1146/annurev-ecolsys-012220-120819>.
- 1051 Guisan, Antoine, and Carsten Rahbek. 2011. “SESAM a New Framework Integrating Macroe-  
1052 cological and Species Distribution Models for Predicting Spatio-Temporal Patterns of Species  
1053 Assemblages.” *Journal of Biogeography* 38 (8): 1433–44. <https://doi.org/10.1111/j.1365-2699.2011.02550.x>.
- 1054 Guisan, Antoine, and Wilfried Thuiller. 2005. “Predicting Species Distribution: Offering More  
1055 Than Simple Habitat Models.” *Ecology Letters* 8 (9): 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>.
- 1056 Guzman, Laura Melissa, Rachel M. Germain, Coreen Forbes, Samantha Straus, Mary I. O’Connor,  
1057 Dominique Gravel, Diane S. Srivastava, and Patrick L. Thompson. 2019. “Towards a  
1058 Multi-Trophic Extension of Metacommunity Ecology.” *Ecology Letters* 22 (1): 19–33.  
1059 <https://doi.org/10.1111/ele.13162>.
- 1060 Hadfield, Jarrod D., Boris R. Krasnov, Robert Poulin, and Shinichi Nakagawa. 2014. “A Tale of  
1061 Two Phylogenies: Comparative Analyses of Ecological Interactions.” *The American Natu-  
1062 ralist* 183 (2): 174–87. <https://doi.org/10.1086/674445>.
- 1063 Halevy, A., P. Norvig, and F. Pereira. 2009. “The Unreasonable Effectiveness of Data.” *IEEE*  
1064 *Intelligent Systems* 24 (2): 8–12. <https://doi.org/10.1109/MIS.2009.36>.
- 1065 Harvey, Eric, Isabelle Gounand, Colette L. Ward, and Florian Altermatt. 2017. “Bridging Ecol-

- 1069        ogy and Conservation: From Ecological Networks to Ecosystem Function.” *Journal of Applied Ecology* 54 (2): 371–79. <https://doi.org/10.1111/1365-2664.12769>.
- 1070
- 1071        Hastie, Trevor, Robert Tibshirani, and J. H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer Series in Statistics. New York, NY: Springer.
- 1072
- 1073
- 1074        Hattab, Tarek, Fabien Leprieur, Frida Ben Rais Lasram, Dominique Gravel, François Le Loc'h, and Camille Albouy. 2016. “Forecasting Fine-Scale Changes in the Food-Web Structure of Coastal Marine Communities Under Climate Change.” *Ecography* 39 (12): 1227–37. <https://doi.org/10.1111/ecog.01937>.
- 1075
- 1076
- 1077
- 1078        Heleno, Ruben, Cristina Garcia, Pedro Jordano, Anna Traveset, José María Gómez, Nico Blüthgen, Jane Memmott, et al. 2014. “Ecological Networks: Delving into the Architecture of Biodiversity.” *Biology Letters* 10 (1). <https://doi.org/10.1098/rsbl.2013.1000>.
- 1079
- 1080
- 1081        Hortal, Joaquín, Francesco de Bello, José Alexandre F. Diniz-Filho, Thomas M. Lewinsohn, Jorge M. Lobo, and Richard J. Ladle. 2015. “Seven Shortfalls That Beset Large-Scale Knowledge of Biodiversity.” *Annual Review of Ecology, Evolution, and Systematics* 46 (1): 523–49. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>.
- 1082
- 1083
- 1084
- 1085        Houlahan, Jeff E., Shawn T. McKinney, T. Michael Anderson, and Brian J. McGill. 2017. “The Priority of Prediction in Ecological Understanding.” *Oikos* 126 (1): 1–7. <https://doi.org/10.1111/oik.03726>.
- 1086
- 1087
- 1088        Hutchinson, Matthew C., Bernat Bramon Mora, Shai Pilosof, Allison K. Barner, Sonia Kéfi, Elisa Thébault, Pedro Jordano, and Daniel B. Stouffer. 2019. “Seeing the Forest for the Trees: Putting Multilayer Networks to Work for Community Ecology.” Edited by Oscar Godoy. *Functional Ecology* 33 (2): 206–17. <https://doi.org/10.1111/1365-2435.13237>.
- 1089
- 1090
- 1091
- 1092
- 1093        Ings, Thomas C., José M. Montoya, Jordi Bascompte, Nico Blüthgen, Lee Brown, Carsten F. Dormann, François Edwards, et al. 2009. “Ecological Networks—Beyond Food Webs.” *The Journal of Animal Ecology* 78 (1): 253–69. <https://doi.org/10.1111/j.1365-2656.2008.01460.x>.
- 1094
- 1095
- 1096

- 1097 Innes, Mike. 2018. “Flux: Elegant Machine Learning with Julia.” *Journal of Open Source*  
1098 *Software* 3 (25): 602. <https://doi.org/10.21105/joss.00602>.
- 1099 Jalili, Mahdi, Yasin Orouskhani, Milad Asgari, Nazanin Alipourfard, and Matjaž Perc. 2017.  
1100 “Link Prediction in Multiplex Online Social Networks.” *Royal Society Open Science* 4 (2):  
1101 160863. <https://doi.org/10.1098/rsos.160863>.
- 1102 Jeliazkov, Alienor, Darko Mijatovic, Stéphane Chantepie, Nigel Andrew, Raphaël Arlettaz, Luc  
1103 Barbaro, Nadia Barsoum, et al. 2020. “A Global Database for Metacommunity Ecology,  
1104 Integrating Species, Traits, Environment and Space.” *Scientific Data* 7 (1): 6. <https://doi.org/10.1038/s41597-019-0344-7>.
- 1105 Jenkins, Clinton N., Stuart L. Pimm, and Lucas N. Joppa. 2013. “Global Patterns of Terrestrial  
1106 Vertebrate Diversity and Conservation.” *Proceedings of the National Academy of Sciences*  
1107 110 (28): E2602–10. <https://doi.org/10.1073/pnas.1302251110>.
- 1108 Jordano, Pedro. 2016a. “Chasing Ecological Interactions.” *PLOS Biology* 14 (9): e1002559.  
1109 <https://doi.org/10.1371/journal.pbio.1002559>.
- 1110 ———. 2016b. “Sampling Networks of Ecological Interactions.” *Functional Ecology*, Septem-  
1111 ber. <https://doi.org/10.1111/1365-2435.12763>.
- 1112 Karbowiak, Grzegorz, Dana Miklisová, Michal Stanko, Joanna Werszko, Marta Hajdul-Marwicz,  
1113 Tomasz Szewczyk, and Leszek Rychlik. 2019. “The Competition Between Immatures of  
1114 Ixodes Ricinus and Dermacentor Reticulatus (Ixodida: Ixodidae) Ticks for Rodent Hosts.”  
1115 *Journal of Medical Entomology* 56 (2): 448–52. <https://doi.org/10.1093/jme/tjy188>.
- 1116 Karger, Dirk Nikolaus, Anna F. Cord, Michael Kessler, Holger Kreft, Ingolf Kühn, Sven Pompe,  
1117 Brody Sandel, et al. 2016. “Delineating Probabilistic Species Pools in Ecology and Biogeog-  
1118 raphy.” *Global Ecology and Biogeography* 25 (4): 489–501. <https://doi.org/10.1111/geb.12422>.
- 1119 Kawakatsu, Mari, Philip S. Chodrow, Nicole Eikmeier, and Daniel B. Larremore. 2021. “Emer-  
1120 gence of Hierarchy in Networked Endorsement Dynamics.” *arXiv:2007.04448 [nlin, Physics:physics]*,  
1121 May. <http://arxiv.org/abs/2007.04448>.
- 1122 Kéfi, Sonia, Virginia Domínguez-García, Ian Donohue, Colin Fontaine, Elisa Thébault, and  
1123

- 1125 Vasilis Dakos. 2019. "Advancing Our Understanding of Ecological Stability." *Ecology*  
1126 *Letters* 22 (9): 1349–56. <https://doi.org/10.1111/ele.13340>.
- 1127 Kindvater, Holly K., Nicholas K. Dulvy, Cat Horswill, Maria-José Juan-Jordá, Marc Mangel,  
1128 and Jason Matthiopoulos. 2018. "Overcoming the Data Crisis in Biodiversity Conserva-  
1129 tion." *Trends in Ecology & Evolution* 33 (9): 676–88. <https://doi.org/10.1016/j.tree.2018.06.004>.
- 1130 Kissling, W. Daniel, Jorge A. Ahumada, Anne Bowser, Miguel Fernandez, Nestor Fernandez,  
1131 Enrique Alonso Garcia, Robert P. Guralnick, et al. 2018. "Building Essential Biodiversity  
1132 Variables (EBVs) of Species Distribution and Abundance at a Global Scale." *Biological*  
1133 *Reviews* 93 (1): 600–625. <https://doi.org/10.1111/brv.12359>.
- 1134 Kortsch, Susanne, Raul Primicerio, Maria Fossheim, Andrey V. Dolgov, and Michaela Aschan.  
1135 2015. "Climate Change Alters the Structure of Arctic Marine Food Webs Due to Poleward  
1136 Shifts of Boreal Generalists." *Proceedings of the Royal Society B: Biological Sciences* 282  
1137 (1814): 20151546. <https://doi.org/10.1098/rspb.2015.1546>.
- 1138 Krasnov, Boris R., Georgy I. Shenbrot, Irina S. Khokhlova, and A. Allan Degen. 2016. "Trait-  
1139 Based and Phylogenetic Associations Between Parasites and Their Hosts: A Case Study with  
1140 Small Mammals and Fleas in the Palearctic." *Oikos* 125 (1): 29–38. <https://doi.org/10.1111/oik.02178>.
- 1141 Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2017. "ImageNet Classification  
1142 with Deep Convolutional Neural Networks." *Communications of the ACM* 60 (6): 84–90.  
1143 <https://doi.org/10.1145/3065386>.
- 1144 Kuhn, Max, and Kjell Johnson. 2019. *Feature Engineering and Selection : A Practical Ap-  
1145 proach for Predictive Models*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781315108230>.
- 1146 Landi, Pietro, Henintsoa O. Minoarivelo, Åke Brännström, Cang Hui, and Ulf Dieckmann.  
1147 2018. "Complexity and Stability of Ecological Networks: A Review of the Theory." *Popu-  
1148 lation Ecology* 60 (4): 319–45. <https://doi.org/10.1007/s10144-018-0628-3>.
- 1149 Lausch, A., L. Bannehr, M. Beckmann, C. Boehm, H. Feilhauer, J. M. Hacker, M. Heurich, et al.  
1150 2016. "Linking Earth Observation and Taxonomic, Structural and Functional Biodiversity:

- 1153 Local to Ecosystem Perspectives.” *Ecological Indicators*, Navigating Urban Complexity:  
1154 Advancing Understanding of Urban Social Ecological Systems for Transformation and Re-  
1155 silience, 70 (November): 317–39. <https://doi.org/10.1016/j.ecolind.2016.06.022>.
- 1156 Lei, Lili, and Jeffrey S. Whitaker. 2017. “Evaluating the Trade-Offs Between Ensemble Size  
1157 and Ensemble Resolution in an Ensemble-Variational Data Assimilation System.” *Jour-*  
1158 *nal of Advances in Modeling Earth Systems* 9 (2): 781–89. <https://doi.org/10.1002/2016MS000864>.
- 1160 Lemaître, Guillaume, Fernando Nogueira, and Christos K. Aridas. 2017. “Imbalanced-Learn:  
1161 A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning.” *Jour-*  
1162 *nal of Machine Learning Research* 18 (17): 1–5.
- 1163 Levin, Simon A. 1998. “Ecosystems and the Biosphere as Complex Adaptive Systems.” *Ecosys-*  
1164 *tems* 1 (5): 431–36. <https://doi.org/10.1007/s100219900037>.
- 1165 Levins, Richard, and Richard Charles Lewontin. 1987. *The Dialectical Biologist*. Cambridge,  
1166 Mass.: Harvard University Press.
- 1167 Luong, Lien T., and Kimberley J. Mathot. 2019. “Facultative Parasites as Evolutionary Stepping-  
1168 Stones Towards Parasitic Lifestyles.” *Biology Letters* 15 (4): 20190058. <https://doi.org/10.1098/rsbl.2019.0058>.
- 1170 MacDonald, Arthur Andrew Meahan, Francis Banville, and Timothée Poisot. 2020. “Revisiting  
1171 the Links-Species Scaling Relationship in Food Webs.” *Patterns* 1 (7). <https://doi.org/10.1016/j.patter.2020.100079>.
- 1173 Macías-Hernández, Nuria, Kacie Athey, Vanina Tonzo, Owen S. Wangensteen, Miquel Arnedo,  
1174 and James D. Harwood. 2018. “Molecular Gut Content Analysis of Different Spider Body  
1175 Parts.” *PloS One* 13 (5): e0196589. <https://doi.org/10.1371/journal.pone.0196589>.
- 1176 Magioli, Marcelo, and Katia Maria Paschoaletto Micchi de Barros Ferraz. 2021. “Deforestation  
1177 Leads to Prey Shrinkage for an Apex Predator in a Biodiversity Hotspot.” *Mammal Research*,  
1178 February. <https://doi.org/10.1007/s13364-021-00556-9>.
- 1179 Makiola, Andreas, Zacchaeus G. Compson, Donald J. Baird, Matthew A. Barnes, Sam P. Boer-  
1180 lijst, Agnès Bouchez, Georgina Brennan, et al. 2020. “Key Questions for Next-Generation

- 1181 Biomonitoring.” *Frontiers in Environmental Science* 7. <https://doi.org/10.3389/fenvs.2019.00197>.
- 1183 Maris, Virginie, Philippe Huneman, Audrey Coreau, Sonia Kéfi, Roger Pradel, and Vincent Devictor. 2017. “Prediction in Ecology: Promises, Obstacles and Clarifications.” *Oikos*, December, n/a–. <https://doi.org/10.1111/oik.04655>.
- 1186 Martínez-Abraín, Alejandro. 2008. “Statistical Significance and Biological Relevance: A Call for a More Cautious Interpretation of Results in Ecology.” *Acta Oecologica* 34 (1): 9–11. <https://doi.org/10.1016/j.actao.2008.02.004>.
- 1189 Massol, F., M. Dubart, V. Calcagno, K. Cazelles, C. Jacquet, S. Kéfi, and D. Gravel. 2017. “Chapter Four - Island Biogeography of Food Webs.” In *Advances in Ecological Research*, edited by David A. Bohan, Alex J. Dumbrell, and François Massol, 56:183–262. Networks of Invasion: A Synthesis of Concepts. Academic Press. <https://doi.org/10.1016/bs.aecr.2016.10.004>.
- 1194 Matthee, Sonja, Alexandr A. Stekolnikov, Luther van der Mescht, Götz Froeschke, and Serge Morand. 2020. “The Diversity and Distribution of Chigger Mites Associated with Rodents in the South African Savanna.” *Parasitology* 147 (9): 1038–47. <https://doi.org/10.1017/S0031182020000748>.
- 1198 May, Robert M. 1974. *Stability and Complexity in Model Ecosystems*. Vol. 1. Princeton University Press. <https://doi.org/10.2307/j.ctvs32rq4>.
- 1200 McCann, Kevin. 2007. “Protecting Biostructure.” *Nature* 446 (7131): 29–29. <https://doi.org/10.1038/446029a>.
- 1202 McDonald-Madden, E., R. Sabbadin, E. T. Game, P. W. J. Baxter, I. Chadès, and H. P. Possingham. 2016. “Using Food-Web Theory to Conserve Ecosystems.” *Nature Communications* 7 (1): 10245. <https://doi.org/10.1038/ncomms10245>.
- 1205 McWilliams, Chris, Miguel Lurgi, Jose M. Montoya, Alix Sauve, and Daniel Montoya. 2019. “The Stability of Multitrophic Communities Under Habitat Loss.” *Nature Communications* 10 (1): 2322. <https://doi.org/10.1038/s41467-019-10370-2>.
- 1208 Michalska-Smith, Matthew J., and Stefano Allesina. 2019. “Telling Ecological Networks Apart

- 1209 by Their Structure: A Computational Challenge.” *PLOS Computational Biology* 15 (6):  
1210 e1007076. <https://doi.org/10.1371/journal.pcbi.1007076>.
- 1211 Montoya, Jose M., and Ricard V. Solé. 2002. “Small World Patterns in Food Webs.” *Journal*  
1212 *of Theoretical Biology* 214 (3): 405–12. <https://doi.org/10.1006/jtbi.2001.2460>.
- 1213 Montoya, José M., Miguel A. Rodríguez, and Bradford A. Hawkins. 2003. “Food Web Com-  
1214 plexity and Higher-Level Ecosystem Services.” *Ecology Letters* 6 (7): 587–93. <https://doi.org/10.1046/j.1461-0248.2003.00469.x>.
- 1215
- 1216 Mora, Bernat Bramon, Dominique Gravel, Luis J. Gilarranz, Timothée Poisot, and Daniel B.  
1217 Stouffer. 2018. “Identifying a Common Backbone of Interactions Underlying Food Webs  
1218 from Different Ecosystems.” *Nature Communications* 9 (1): 1–8. <https://doi.org/10.1038/s41467-018-05056-0>.
- 1219
- 1220 Morales-Castilla, Ignacio, Miguel G. Matias, Dominique Gravel, and Miguel B. Araújo. 2015.  
1221 “Inferring Biotic Interactions from Proxies.” *Trends in Ecology & Evolution* 30 (6): 347–56.  
1222 <https://doi.org/10.1016/j.tree.2015.03.014>.
- 1223
- 1224 Morand, Serge, Kittipong Chaisiri, Anamika Kritiyakan, and Rawadee Kumlert. 2020. “Dis-  
1225 ease Ecology of Rickettsial Species: A Data Science Approach.” *Tropical Medicine and*  
*Infectious Disease* 5 (2): 64. <https://doi.org/10.3390/tropicalmed5020064>.
- 1226
- 1227 Morand, Serge, Andrea Simková, Iveta Matejusová, Laetitia Plaisance, Olivier Verneau, and  
1228 Yves Desdevises. 2002. “Investigating Patterns May Reveal Processes: Evolutionary Ecol-  
1229 ogy of Ectoparasitic Monogeneans.” *International Journal for Parasitology* 32 (2): 111–19.  
[https://doi.org/10.1016/s0020-7519\(01\)00347-2](https://doi.org/10.1016/s0020-7519(01)00347-2).
- 1230
- 1231 Najari, Shaghayegh, Mostafa Salehi, Vahid Ranjbar, and Mahdi Jalili. 2019. “Link Prediction in  
1232 Multiplex Networks Based on Interlayer Similarity.” *Physica A: Statistical Mechanics and*  
*Its Applications* 536 (December): 120978. <https://doi.org/10.1016/j.physa.2019.04.214>.
- 1233
- 1234 Neutel, A.-M. 2002. “Stability in Real Food Webs: Weak Links in Long Loops.” *Science* 296  
1235 (5570): 1120–23. <https://doi.org/10.1126/science.1068326>.
- 1236 Norberg, Anna, Nerea Abrego, F. Guillaume Blanchet, Frederick R. Adler, Barbara J. Anderson,

- 1237 Jani Anttila, Miguel B. Araújo, et al. 2019. "A Comprehensive Evaluation of Predictive Per-  
1238 formance of 33 Species Distribution Models at Species and Community Levels." *Ecological*  
1239 *Monographs* 89 (3): e01370. <https://doi.org/10.1002/ecm.1370>.
- 1240 Novak, Mark, and J. Timothy Wootton. 2008. "Estimating Nonlinear Interaction Strengths:  
1241 An Observation-Based Method for Species-Rich Food Webs." *Ecology* 89 (8): 2083–89.  
1242 <https://doi.org/10.1890/08-0033.1>.
- 1243 O'Connor, Louise M. J., Laura J. Pollock, João Braga, Gentile Francesco Ficetola, Luigi Maiorano,  
1244 Camille Martinez-Almoyna, Alessandro Montemaggiori, Marc Ohlmann, and Wilfried Thuiller. 2020. "Unveiling the Food Webs of Tetrapods Across Europe Through the Prism  
1245 of the Eltonian Niche." *Journal of Biogeography* 47 (1): 181–92. <https://doi.org/10.1111/jbi.13773>.
- 1246 Olden, Julian D., Joshua J. Lawler, and N. LeRoy Poff. 2008. "Machine Learning Methods  
1247 Without Tears: A Primer for Ecologists." *The Quarterly Review of Biology* 83 (2): 171–93.  
<https://doi.org/10.1086/587826>.
- 1248 Ovaskainen, Otso, Gleb Tikhonov, Anna Norberg, F. Guillaume Blanchet, Leo Duan, David  
1249 Dunson, Tomas Roslin, and Nerea Abrego. 2017. "How to Make More Out of Community  
1250 Data? A Conceptual Framework and Its Implementation as Models and Software." *Ecology*  
1251 *Letters* 20 (5): 561–76. <https://doi.org/10.1111/ele.12757>.
- 1252 Parker, Wendy S. 2013. "Ensemble Modeling, Uncertainty and Robust Predictions." *WIREs  
1253 Climate Change* 4 (3): 213–23. <https://doi.org/10.1002/wcc.220>.
- 1254 Parysow, Pablo, George Gertner, and James Westervelt. 2000. "Efficient Approximation for  
1255 Building Error Budgets for Process Models." *Ecological Modelling* 135 (2): 111–25. [https://doi.org/10.1016/S0304-3800\(00\)00347-1](https://doi.org/10.1016/S0304-3800(00)00347-1).
- 1256 Pascual, Mercedes, and Jennifer A. Dunne. 2006. *Ecological Networks: Linking Structure to  
1257 Dynamics in Food Webs*. Oxford University Press, USA.
- 1258 Pennekamp, Frank, Alison C. Iles, Joshua Garland, Georgina Brennan, Ulrich Brose, Ursula  
1259 Gaedke, Ute Jacob, et al. 2019. "The Intrinsic Predictability of Ecological Time Series  
1260 and Its Potential to Guide Forecasting." *Ecological Monographs* 89 (2): e01359. <https://doi.org/10.1002/ecm.1359>.

- 1265 //doi.org/10.1002/ecm.1359.
- 1266 Petchey, O. L., A. P. Beckerman, J. O. Riede, and P. H. Warren. 2008. "Size, Foraging, and  
1267 Food Web Structure." *Proceedings of the National Academy of Sciences* 105 (11): 4191–  
1268 96. <https://doi.org/10.1073/pnas.0710672105>.
- 1269 Peters, Debra P. C., and Jeffrey E. Herrick. 2004. "Strategies for Ecological Extrapolation."  
1270 *Oikos* 106 (3): 627–36. <https://doi.org/10.1111/j.0030-1299.2004.12869.x>.
- 1271 Pérez-Harguindeguy, N., S. Díaz, E. Garnier, S. Lavorel, H. Poorter, P. Jaureguiberry, M. S.  
1272 Bret-Harte, et al. 2013. "New Handbook for Standardised Measurement of Plant Functional  
1273 Traits Worldwide." *Australian Journal of Botany* 61 (3): 167. <https://doi.org/10.1071/BT12225>.
- 1275 Pichler, Maximilian, Virginie Boreux, Alexandra-Maria Klein, Matthias Schleuning, and Flo-  
1276 rian Hartig. 2020. "Machine Learning Algorithms to Infer Trait-Matching and Predict  
1277 Species Interactions in Ecological Networks." *Methods in Ecology and Evolution* 11 (2):  
1278 281–93. <https://doi.org/10.1111/2041-210X.13329>.
- 1279 Pigot, Alex L., and Joseph A. Tobias. 2013. "Species Interactions Constrain Geographic Range  
1280 Expansion over Evolutionary Time." Edited by Colleen Webb. *Ecology Letters* 16 (3): 330–  
1281 38. <https://doi.org/10.1111/ele.12043>.
- 1282 Pilosof, Shai, Mason A. Porter, Mercedes Pascual, and Sonia Kéfi. 2017. "The Multilayer  
1283 Nature of Ecological Networks." *Nature Ecology & Evolution* 1 (4): 101. <https://doi.org/10.1038/s41559-017-0101>.
- 1285 Pocock, Michael J. O., Darren M. Evans, and Jane Memmott. 2012. "The Robustness and  
1286 Restoration of a Network of Ecological Networks." *Science* 335 (6071): 973–77. <https://doi.org/10.1126/science.1214915>.
- 1288 Poelen, Jorrit H., James D. Simons, and Chris J. Mungall. 2014. "Global Biotic Interactions:  
1289 An Open Infrastructure to Share and Analyze Species-Interaction Datasets." *Ecological  
1290 Informatics* 24 (November): 148–59. <https://doi.org/10.1016/j.ecoinf.2014.08.005>.

- 1292 Poisot, Timothée, Benjamin Baiser, Jennifer Dunne, Sonia Kéfi, François Massol, Nicolas Mou-  
1293 quet, Tamara N. Romanuk, Daniel B. Stouffer, Spencer A. Wood, and Dominique Gravel.  
1294 2016. “Mangal Making Ecological Network Analysis Simple.” *Ecography* 39 (4): 384–90.  
1295 <https://doi.org/10.1111/ecog.00976>.
- 1296 Poisot, Timothée, Gabriel Bergeron, Kevin Cazelles, Tad Dallas, Dominique Gravel, Andrew  
1297 MacDonald, Benjamin Mercier, Clément Violet, and Steve Vissault. 2021. “Global Knowl-  
1298 edge Gaps in Species Interaction Networks Data.” *Journal of Biogeography*, April, jbi.14127.  
1299 <https://doi.org/10.1111/jbi.14127>.
- 1300 Poisot, Timothée, Gabriel Bergeron, Kevin Cazelles, Tad Dallas, Dominique Gravel, Andrew  
1301 Macdonald, Benjamin Mercier, Clément Violet, and Steve Vissault. 2020. “Environmen-  
1302 tal Biases in the Study of Ecological Networks at the Planetary Scale.” *bioRxiv*, January,  
1303 2020.01.27.921429. <https://doi.org/10.1101/2020.01.27.921429>.
- 1304 Poisot, Timothée, and Dominique Gravel. 2014. “When Is an Ecological Network Complex?  
1305 Connectance Drives Degree Distribution and Emerging Network Properties.” *PeerJ* 2 (Febru-  
1306 ary): e251. <https://doi.org/10.7717/peerj.251>.
- 1307 Poisot, Timothée, Marie-Andrée Ouellet, Nardus Mollentze, Maxwell J. Farrell, Daniel J. Becker,  
1308 Gregory F. Albery, Rory J. Gibb, Stephanie N. Seifert, and Colin J. Carlson. 2021. “Im-  
1309 puting the Mammalian Virome with Linear Filtering and Singular Value Decomposition.”  
1310 *arXiv:2105.14973 [q-Bio]*, May. <http://arxiv.org/abs/2105.14973>.
- 1311 Poisot, Timothée, Michal Stanko, Dana Miklisová, and Serge Morand. 2013. “Facultative and  
1312 Obligate Parasite Communities Exhibit Different Network Properties.” *Parasitology* 140  
1313 (11): 1340–45. <https://doi.org/10.1017/S0031182013000851>.
- 1314 Poisot, Timothée, and Daniel B. Stouffer. 2018. “Interactions Retain the Co-Phylogenetic  
1315 Matching That Communities Lost.” *Oikos* 127 (2): 230–38. <https://doi.org/10.1111/oik.03788>.
- 1317 Poisot, Timothée, Daniel B. Stouffer, and Dominique Gravel. 2015. “Beyond Species: Why  
1318 Ecological Interaction Networks Vary Through Space and Time.” *Oikos* 124 (3): 243–51.  
1319 <https://doi.org/10.1111/oik.01719>.

- 1320 Pollock, Laura J., Louise M. J. O'Connor, Karel Mokany, Dan F. Rosauer, Matthew V. Talluto,  
1321 and Wilfried Thuiller. 2020. “Protecting Biodiversity (in All Its Complexity): New Models  
1322 and Methods.” *Trends in Ecology & Evolution* 35 (12): 1119–28. [https://doi.org/10.](https://doi.org/10.1016/j.tree.2020.08.015)  
1323 [10.1016/j.tree.2020.08.015](https://doi.org/10.1016/j.tree.2020.08.015).
- 1324 Pomeranz, Justin PF, Ross M. Thompson, Timothée Poisot, and Jon S. Harding. 2018. “Inferring  
1325 Predator-Prey Interactions in Food Webs.” *Methods in Ecology and Evolution* 0 (ja). [https:](https://doi.org/10.1111/2041-210X.13125)  
1326 [//doi.org/10.1111/2041-210X.13125](https://doi.org/10.1111/2041-210X.13125).
- 1327 Portalier, Sébastien M. J., Gregor F. Fussmann, Michel Loreau, and Mehdi Cherif. 2019. “The  
1328 Mechanics of Predatorprey Interactions: First Principles of Physics Predict Predatorprey  
1329 Size Ratios.” *Functional Ecology* 33 (2): 323–34. [https://doi.org/10.1111/1365-2435.](https://doi.org/10.1111/1365-2435.13254)  
1330 [13254](https://doi.org/10.1111/1365-2435.13254).
- 1331 Porter, John, Peter Arzberger, Hans-Werner Braun, Pablo Bryant, Stuart Gage, Todd Hansen,  
1332 Paul Hanson, et al. 2005. “Wireless Sensor Networks for Ecology.” *BioScience* 55 (7): 561–  
1333 72. [https://doi.org/10.1641/0006-3568\(2005\)055%5B0561:WSNFE%5D2.0.CO;2](https://doi.org/10.1641/0006-3568(2005)055%5B0561:WSNFE%5D2.0.CO;2).
- 1334 Rackauckas, Christopher, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Su-  
1335 pekar, Dominic Skinner, Ali Ramadhan, and Alan Edelman. 2020. “Universal Differential  
1336 Equations for Scientific Machine Learning,” January.
- 1337 Rengifo-Correa, Laura, Christopher R Stephens, Juan J Morrone, Juan Luis Téllez-Rendón, and  
1338 Constantino Gonzalez-Salazar. 2017. “Understanding Transmissibility Patterns of Chagas  
1339 Disease Through Complex Vector-Host Networks.” *Parasitology* 144 (6): 760.
- 1340 Riede, Jens O., Björn C. Rall, Carolin Banasek-Richter, Sergio A. Navarrete, Evie A. Wieters,  
1341 Mark C. Emmerson, Ute Jacob, and Ulrich Brose. 2010. “Scaling of Food-Web Properties  
1342 with Diversity and Complexity Across Ecosystems.” In *Advances in Ecological Research*,  
1343 42:139–70. Elsevier. <https://doi.org/10.1016/B978-0-12-381363-3.00003-4>.
- 1344 Rohr, Rudolf Philippe, Heike Scherer, Patrik Kehrli, Christian Mazza, and Louis-Félix Bersier.  
1345 2010. “Modeling Food Webs: Exploring Unexplained Structure Using Latent Traits.” *The*  
1346 *American Naturalist* 176 (2): 170–77. <https://doi.org/10.1086/653667>.
- 1347 Rossberg, A. G., H. Matsuda, T. Amemiya, and K. Itoh. 2006. “Food Webs: Experts Consuming

- 1348 Families of Experts.” *Journal of Theoretical Biology* 241 (3): 552–63. <https://doi.org/10.1016/j.jtbi.2005.12.021>.
- 1349
- 1350 Ruiter, Peter C. de, Anje-Margriet Neutel, and John C. Moore. 1995. “Energetics, Patterns  
1351 of Interaction Strengths, and Stability in Real Ecosystems.” *Science* 269 (5228): 1257–60.  
1352 <https://doi.org/10.1126/science.269.5228.1257>.
- 1353 Saavedra, Serguei, Rudolf P. Rohr, Jordi Bascompte, Oscar Godoy, Nathan J. B. Kraft, and  
1354 Jonathan M. Levine. 2017. “A Structural Approach for Understanding Multispecies Coex-  
1355 istence.” *Ecological Monographs* 87 (3): 470–86. <https://doi.org/10.1002/ecm.1263>.
- 1356 Sala, Enric, and Michael H. Graham. 2002. “Community-Wide Distribution of Predatorprey  
1357 Interaction Strength in Kelp Forests.” *Proceedings of the National Academy of Sciences* 99  
1358 (6): 3678–83. <https://doi.org/10.1073/pnas.052028499>.
- 1359 Sander, Elizabeth L., J. Timothy Wootton, and Stefano Allesina. 2017. “Ecological Network  
1360 Inference from Long-Term Presence-Absence Data.” *Scientific Reports* 7 (1): 7154. <https://doi.org/10.1038/s41598-017-07009-x>.
- 1361
- 1362 Sasal, P., N. Niquil, and P. Bartoli. 1999. “Community Structure of Digenean Parasites of Sparid  
1363 and Labrid Fishes of the Mediterranean Sea: A New Approach.” *Parasitology* 119 ( Pt 6)  
1364 (December): 635–48. <https://doi.org/10.1017/s0031182099005077>.
- 1365 Sánchez, S., E. Serrano, M. S. Gómez, C. Feliu, and S. Morand. 2014. “Positive Co-Occurrence  
1366 of Flea Infestation at a Low Biological Cost in Two Rodent Hosts in the Canary Archipelago.”  
1367 *Parasitology* 141 (4): 511–21. <https://doi.org/10.1017/S0031182013001753>.
- 1368 Schmid-Araya, Jenny M., Peter E. Schmid, Steven P. Tod, and Genoveva F. Esteban. 2016.  
1369 “Trophic Positioning of Meiofauna Revealed by Stable Isotopes and Food Web Analyses.”  
1370 *Ecology* 97 (11): 3099–109. <https://doi.org/10.1002/ecy.1553>.
- 1371 Scholes, Robert J, Michele Walters, Eren Turak, Hannu Saarenmaa, Carlo HR Heip, Éamonn Ó  
1372 Tuama, Daniel P Faith, et al. 2012. “Building a Global Observing System for Biodiversity.”  
1373 *Current Opinion in Environmental Sustainability*, Open issue, 4 (1): 139–46. <https://doi.org/10.1016/j.cosust.2011.12.005>.
- 1374
- 1375 Segar, Simon T., Tom M. Fayle, Diane S. Srivastava, Thomas M. Lewinsohn, Owen T. Lewis,

- 1376 Vojtech Novotny, Roger L. Kitching, and Sarah C. Maunsell. 2020. “The Role of Evolution  
1377 in Shaping Ecological Networks.” *Trends in Ecology & Evolution* 35 (5): 454–66. <https://doi.org/10.1016/j.tree.2020.01.004>.
- 1378
- 1379 Seibold, Sebastian, Marc W. Cadotte, J. Scott MacIvor, Simon Thorn, and Jörg Müller. 2018.  
1380 “The Necessity of Multitrophic Approaches in Community Ecology.” *Trends in Ecology &*  
1381 *Evolution* 33 (10): 754–64. <https://doi.org/10.1016/j.tree.2018.07.001>.
- 1382 Simková, Andrea, Dusan Kadlec, Milan Gelnar, and Serge Morand. 2002. “Abundance-Prevalence  
1383 Relationship of Gill Congeneric Ectoparasites: Testing the Core Satellite Hypothesis and  
1384 Ecological Specialisation.” *Parasitology Research* 88 (7): 682–86. <https://doi.org/10.1007/s00436-002-0650-3>.
- 1385
- 1386 Skidmore, Andrew K., and Nathalie Pettorelli. 2015. “Agree on Biodiversity Metrics to Track  
1387 from Space: Ecologists and Space Agencies Must Forge a Global Monitoring Strategy.”  
1388 *Nature* 523 (7561): 403–6.
- 1389 Smith, Gabriel R., Roger D. Finlay, Jan Stenlid, Rimvydas Vasaitis, and Audrius Menkis. 2017.  
1390 “Growing Evidence for Facultative Biotrophy in Sapro trophic Fungi: Data from Microcosm  
1391 Tests with 201 Species of Wood-Decay Basidiomycetes.” *The New Phytologist* 215 (2):  
1392 747–55. <https://doi.org/10.1111/nph.14551>.
- 1393 Smith-Ramesh, Lauren M., Alexandria C. Moore, and Oswald J. Schmitz. 2016. “Global Syn-  
1394 thesis Suggests That Food Web Connectance Correlates to Invasion Resistance.” *Global  
1395 Change Biology*, August, n/a–. <https://doi.org/10.1111/gcb.13460>.
- 1396 Spiegelhalter, David J, Jonathan P Myles, David R Jones, and Keith R Abrams. 2000. “Bayesian  
1397 Methods in Health Technology Assessment: A Review.”
- 1398 Staniczenko, Phillip P. A., Prabu Sivasubramaniam, K. Blake Suttle, and Richard G. Pearson.  
1399 2017. “Linking Macroecology and Community Ecology: Refining Predictions of Species  
1400 Distributions Using Biotic Interaction Networks.” *Ecology Letters* 20 (6): 693–707. <https://doi.org/10.1111/ele.12770>.
- 1401
- 1402 Stephens, CR. 2009. “Using Biotic Interaction Networks for Prediction in Biodiversity and  
1403 Emerging Diseases.” *PLoS ONE* 4(5): <https://doi.org/doi:10.1371/journal.pone.0005137>.

- 1404 0005725.
- 1405 Stephenson, PJ. 2020. “Technological Advances in Biodiversity Monitoring: Applicability, Op-  
1406 portunities and Challenges.” *Current Opinion in Environmental Sustainability* 45 (August):  
1407 36–41. <https://doi.org/10.1016/j.cosust.2020.08.005>.
- 1408 Stock, Michiel, Timothée Poisot, Willem Waegeman, and Bernard De Baets. 2017. “Linear  
1409 Filtering Reveals False Negatives in Species Interaction Data.” *Scientific Reports* 7 (April):  
1410 45908. <https://doi.org/10.1038/srep45908>.
- 1411 Stouffer, Daniel B. 2019. “All Ecological Models Are Wrong, but Some Are Useful.” *Journal  
1412 of Animal Ecology* 88 (2): 192–95. <https://doi.org/10.1111/1365-2656.12949>.
- 1413 Stouffer, Daniel B., and Jordi Bascompte. 2010. “Understanding Food-Web Persistence from  
1414 Local to Global Scales.” *Ecology Letters* 13 (2): 154–61. [https://doi.org/10.1111/j.1461-0248.2009.01407.x](https://doi.org/10.1111/j.<br/>1415 1461-0248.2009.01407.x).
- 1416 Strydom, Tanya, Giulio V. Dalla Riva, and Timothée Poisot. 2021. “SVD Entropy Reveals the  
1417 High Complexity of Ecological Networks.” *Frontiers in Ecology and Evolution* 9. <https://doi.org/10.3389/fevo.2021.623141>.
- 1419 Syfert, Mindy M., Lucas Joppa, Matthew J. Smith, David A. Coomes, Steven P. Bachman, and  
1420 Neil A. Brummitt. 2014. “Using Species Distribution Models to Inform IUCN Red List  
1421 Assessments.” *Biological Conservation* 177 (September): 174–84. [https://doi.org/10.1016/j.biocon.2014.06.012](https://doi.org/10.<br/>1422 1016/j.biocon.2014.06.012).
- 1423 Terry, J. Christopher D., and Owen T. Lewis. 2020. “Finding Missing Links in Interaction  
1424 Networks.” *Ecology* 101 (7): e03047. <https://doi.org/10.1002/ecy.3047>.
- 1425 Thompson, Ross M., Ulrich Brose, Jennifer Dunne, Robert O. Hall, Sally Hladyz, Roger L.  
1426 Kitching, Neo D. Martinez, et al. 2012. “Food Webs: Reconciling the Structure and Func-  
1427 tion of Biodiversity.” *Trends in Ecology & Evolution* 27 (12): 689–97. [https://doi.org/10.1016/j.tree.2012.08.005](https://doi.org/<br/>1428 10.1016/j.tree.2012.08.005).
- 1429 Tinker, M. Tim, Paulo R. Guimarães, Mark Novak, Flavia Maria Darcie Marquitti, James L.  
1430 Bodkin, Michelle Staedler, Gena Bentall, and James A. Estes. 2012. “Structure and Mech-  
1431 anism of Diet Specialisation: Testing Models of Individual Variation in Resource Use with

- 1432 Sea Otters.” *Ecology Letters* 15 (5): 475–83. <https://doi.org/10.1111/j.1461-0248.2012.01760.x>.
- 1433 Tipping, Michael E., and Christopher M. Bishop. 1999. “Probabilistic Principal Component Analysis.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61 (3): 611–22. <https://doi.org/10.1111/1467-9868.00196>.
- 1434 Trøjelsgaard, Kristian, and Jens M. Olesen. 2016. “Ecological Networks in Motion: Micro- and Macroscopic Variability Across Scales.” *Functional Ecology* 30 (12): 1926–35. <https://doi.org/10.1111/1365-2435.12710>.
- 1435 Urban, M. C., G. Bocedi, A. P. Hendry, J.- B. Mihoub, G. Peer, A. Singer, J. R. Bridle, et al. 2016. “Improving the Forecast for Biodiversity Under Climate Change.” *Science* 353 (6304): aad8466–66. <https://doi.org/10.1126/science.aad8466>.
- 1436 Vazquez, Diego P., Nico Blüthgen, Luciano Cagnolo, and Natacha P. Chacoff. 2009. “Uniting Pattern and Process in Plant-Animal Mutualistic Networks: A Review.” *Annals of Botany* 103 (9): 1445–57. <https://doi.org/10.1093/aob/mcp057>.
- 1437 Vázquez, Diego P., Carlos J. Melián, Neal M. Williams, Nico Blüthgen, Boris R. Krasnov, and Robert Poulin. 2007. “Species Abundance and Asymmetric Interaction Strength in Ecological Networks.” *Oikos* 116 (7): 1120–27. <https://doi.org/10.1111/j.0030-1299.2007.15828.x>.
- 1438 Vázquez, Diego P., William F. Morris, and Pedro Jordano. 2005. “Interaction Frequency as a Surrogate for the Total Effect of Animal Mutualists on Plants.” *Ecology Letters* 8 (10): 1088–94.
- 1439 Vieira, Marcos Costa, and Mário Almeida-Neto. 2015. “A Simple Stochastic Model for Complex Coextinctions in Mutualistic Networks: Robustness Decreases with Connectance.” *Ecology Letters* 18 (2): 144–52. <https://doi.org/10.1111/ele.12394>.
- 1440 Wardeh, Maya, Matthew Baylis, and Marcus S. C. Blagrove. 2021. “Predicting Mammalian Hosts in Which Novel Coronaviruses Can Be Generated.” *Nature Communications* 12 (1): 780. <https://doi.org/10.1038/s41467-021-21034-5>.

- 1459 Weinstein, Ben G., and Catherine H. Graham. 2017. “On Comparing Traits and Abundance for  
1460 Predicting Species Interactions with Imperfect Detection.” *Food Webs* 11: 17–25.
- 1461 Wells, Konstans, and Robert B. O’Hara. 2013. “Species Interactions: Estimating Per-Individual  
1462 Interaction Strength and Covariates Before Simplifying Data into Per-Species Ecological  
1463 Networks.” *Methods in Ecology and Evolution* 4 (1): 1–8. <https://doi.org/10.1111/j.2041-210x.2012.00249.x>.
- 1465 Wiens, J. A. 1989. “Spatial Scaling in Ecology.” *Functional Ecology* 3 (4): 385–97. <https://doi.org/10.2307/2389612>.
- 1467 Wisz, Mary Susanne, Julien Pottier, W. Daniel Kissling, Loïc Pellissier, Jonathan Lenoir, Christian F. Damgaard, Carsten F. Dormann, et al. 2013. “The Role of Biotic Interactions in  
1468 Shaping Distributions and Realised Assemblages of Species: Implications for Species Dis-  
1469 tribution Modelling.” *Biological Reviews* 88 (1): 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>.
- 1472 Wootton, J. Timothy. 1997. “Estimates and Tests of Per Capita Interaction Strength: Diet,  
1473 Abundance, and Impact of Intertidally Foraging Birds.” *Ecological Monographs* 67 (1): 45–  
1474 64. [https://doi.org/10.1890/0012-9615\(1997\)067%5B0045:EATOPC%5D2.0.CO;2](https://doi.org/10.1890/0012-9615(1997)067%5B0045:EATOPC%5D2.0.CO;2).
- 1475 Wootton, J. Timothy, and Mark Emmerson. 2005. “Measurement of Interaction Strength in  
1476 Nature.” *Annual Review of Ecology, Evolution, and Systematics* 36 (1): 419–44. <https://doi.org/10.1146/annurev.ecolsys.36.091704.175535>.
- 1478 Yasami, Yasser, and Farshad Safaei. 2018. “A Novel Multilayer Model for Missing Link Pre-  
1479 diction and Future Link Forecasting in Dynamic Complex Networks.” *Physica A: Statistical  
1480 Mechanics and Its Applications* 492 (February): 2166–97. <https://doi.org/10.1016/j.physa.2017.11.134>.
- 1482 Yen, Tzu-Chi, and Daniel B. Larremore. 2020. “Community Detection in Bipartite Networks  
1483 with Stochastic Blockmodels.” *Physical Review E* 102 (3): 032309. <https://doi.org/10.1103/PhysRevE.102.032309>.
- 1485 Yodzis, P., and S. Innes. 1992. “Body Size and Consumer-Resource Dynamics.” *The American  
1486 Naturalist* 139 (6): 1151–75. <https://doi.org/10.1086/285380>.

- 1487 Zhang, Minhua, and Fangliang He. 2021. “Plant Breeding Systems Influence the Seasonal Dy-  
1488 namics of Plant-Pollinator Networks in a Subtropical Forest.” *Oecologia*, February. <https://doi.org/10.1007/s00442-021-04863-5>.
- 1489
- 1490 Zhao, Naizhuo, Katia Charland, Mabel Carabali, Elaine O. Nsoesie, Mathieu Maheu-Giroux,  
1491 Erin Rees, Mengru Yuan, Cesar Garcia Balaguera, Gloria Jaramillo Ramirez, and Kate  
1492 Zinszer. 2020. “Machine Learning and Dengue Forecasting: Comparing Random Forests  
1493 and Artificial Neural Networks for Predicting Dengue Burden at National and Sub-National  
1494 Scales in Colombia.” Edited by Marc Choisy. *PLOS Neglected Tropical Diseases* 14 (9):  
1495 e0008056. <https://doi.org/10.1371/journal.pntd.0008056>.
- 1496
- 1497 Zurell, Damaris, Niklaus E. Zimmermann, Helge Gross, Andri Baltensweiler, Thomas Sattler,  
1498 and Rafael O. Wüest. 2020. “Testing Species Assemblage Predictions from Stacked and  
Joint Species Distribution Models.” *Journal of Biogeography* 47 (1): 101–13. <https://doi.org/10.1111/jbi.13608>.
- 1499

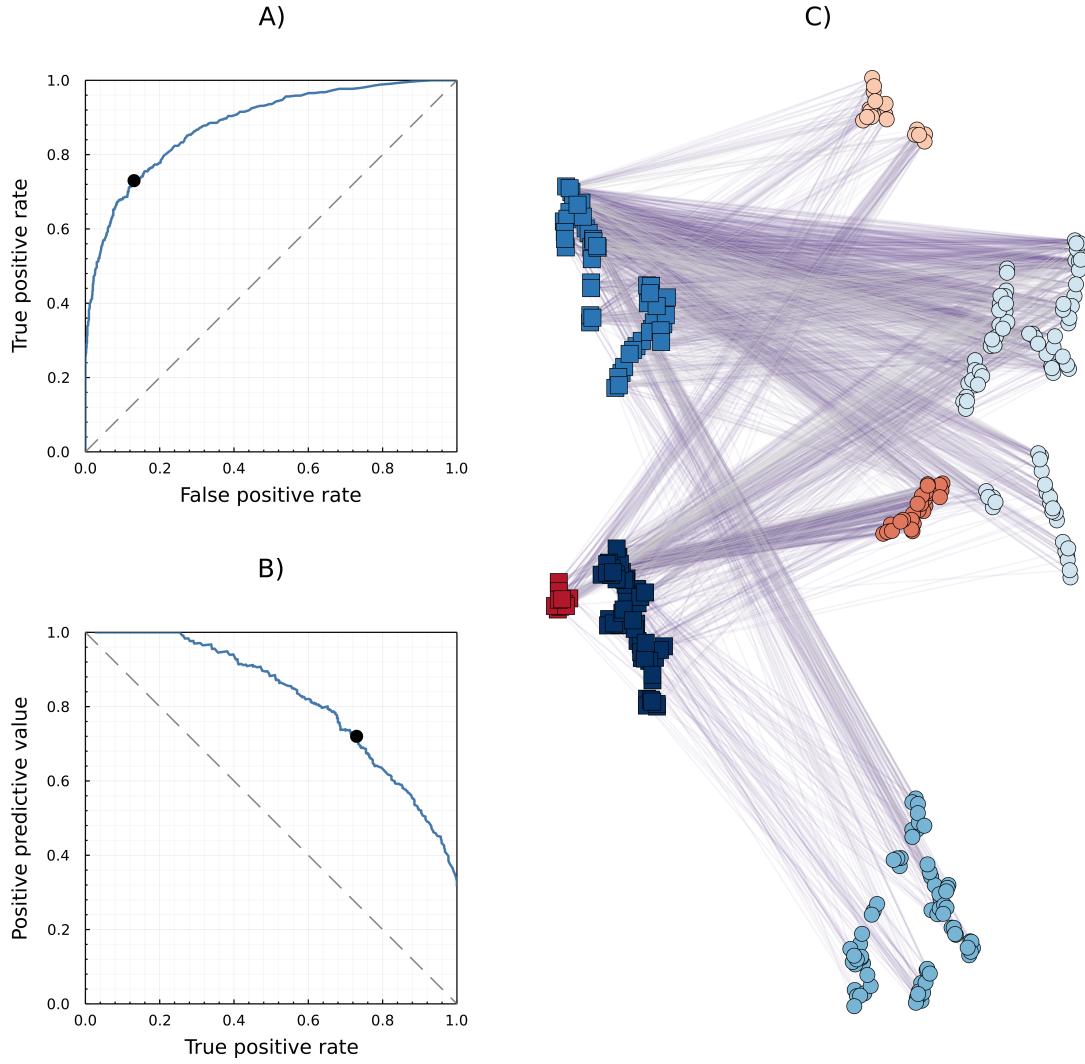


Figure 1: Proof-of-Concept: An empirical metaweb (from Hadfield et al. 2014), i.e. a list of known possible interactions within a species pool, is converted into latent features using probabilistic PCA, then used to train a deep neural network to predict species interactions. Panels A and B represent, respectively, the ROC curve and the precision-recall curve, with the best classifier (according to Youden's J) represented by a black dot. The expected performance of a neutral “random-guessing” classifier is shown with a dashed line. Panel C shows the imputed using t-distributed stochastic neighbour embedding (tSNE), and the colours of nodes are the cluster to which they are assigned based on a  $k$ -means clustering of the tSNE output. Empirical interactions are shown in purple, and imputed interactions in grey.

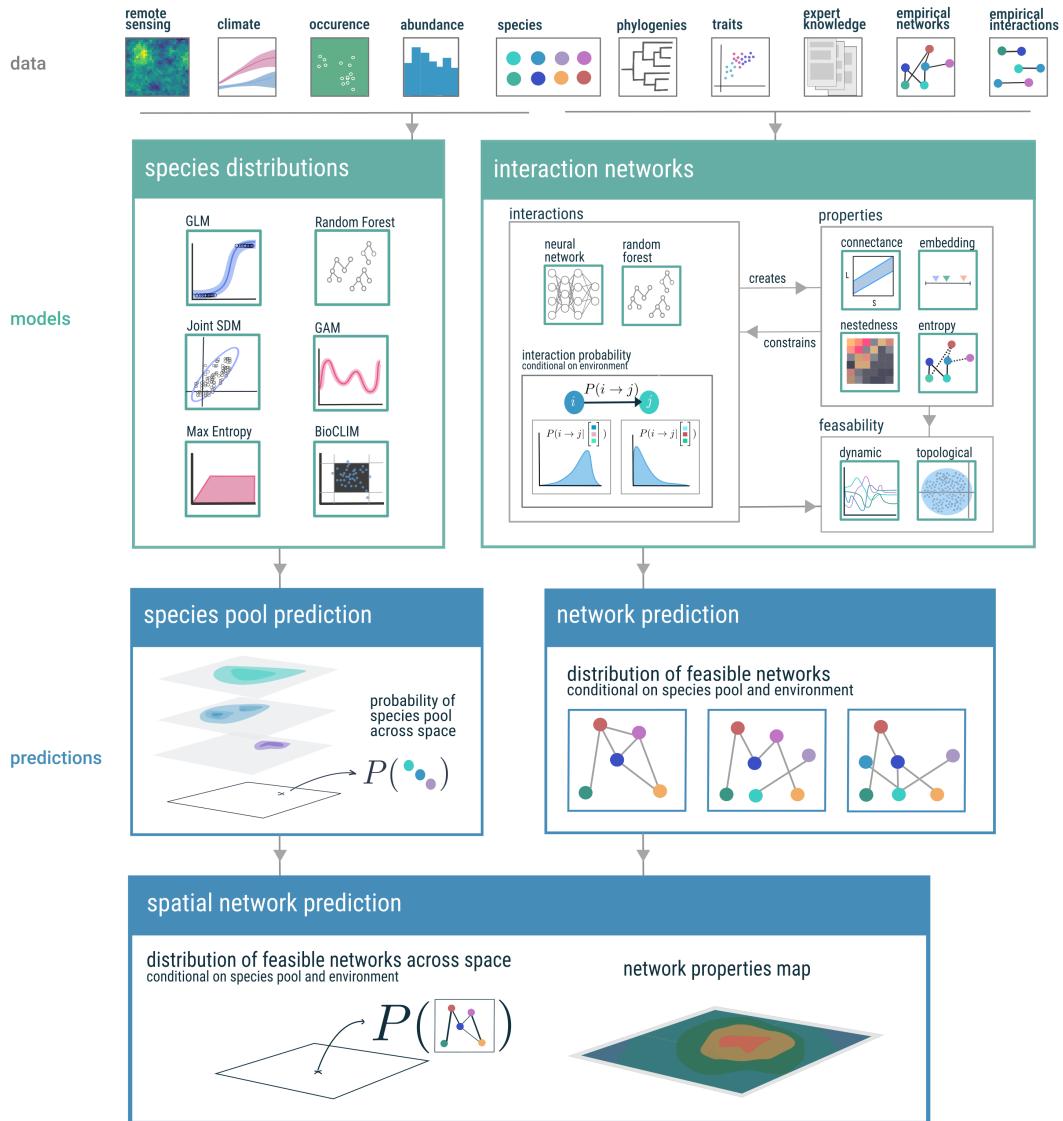


Figure 2: A conceptual roadmap highlighting key areas for the prediction of ecological networks. Starting with the input of data from multiple sources, followed by a modelling framework for ecological networks and the landscape, which are then ultimately combined to allow for the prediction of spatially explicit networks.

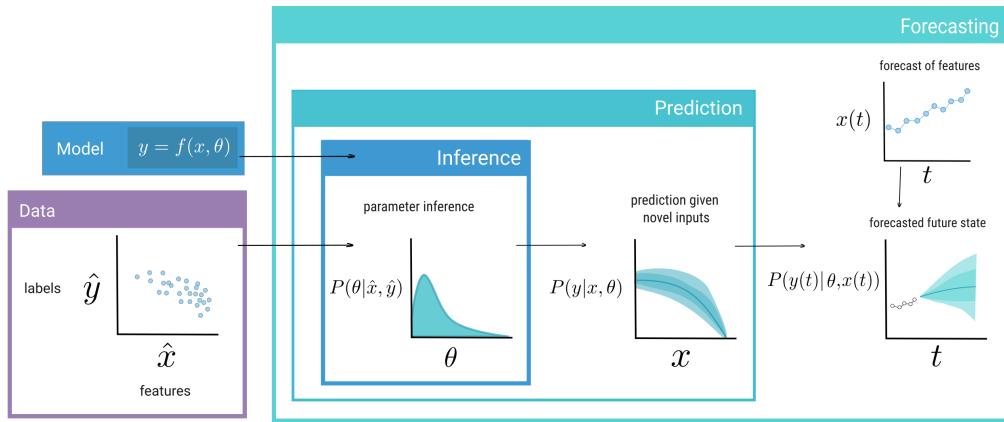


Figure 3: The nested nature of developing predictive and forecasting models, showcases the *forward problem* and how this relies on a hierarchical structure of the modelling process. The choice of a specific modelling technique and framework, as well as the data retained to be part of this model, proceeds directly from our assumptions about which ecological mechanisms are important in shaping both extant and future data.