

# The biological interpretation of probabilistic food webs

Francis Banville<sup>1,2,3,‡</sup> Tanya Strydom<sup>1,3,‡</sup> Timothée Poisot<sup>1,3</sup>

<sup>1</sup> Université de Montréal   <sup>2</sup> Université de Sherbrooke   <sup>3</sup> Quebec Centre for Biodiversity Science

‡ These authors contributed equally to the work

## Correspondance to:

Francis Banville — francis.banville@umontreal.ca

This work is released by its authors under a CC-BY 4.0 license



Last revision: *February 10, 2023*

Community ecologists are increasingly thinking probabilistically when it comes to food webs and other ecological networks. Assuredly, the benefits of representing ecological interactions as probabilistic events (e.g., how likely are species to interact?) instead of binary objects (e.g., do species interact?) are numerous, from a better assessment of the spatiotemporal variation of trophic interactions to an increase capacity to reconstruct networks from sparse data. However, probabilities need to be used with caution when working with species interactions. Indeed, depending on the system at hand and the method used to build probabilistic networks, probabilities can have different interpretations that imply different ways to manipulate them. At the core of these differences lie the distinction between assessing the likelihood that two groups of individuals *can* interact and the likelihood that they *will* interact. This impacts the spatial, temporal, and taxonomic scaling of interaction probabilities, thus enlightening the need to properly define them in their ecological context. Published data on probabilistic species interactions are poorly documented in the literature, which impedes our ability to use them appropriately. With these challenges in mind, we propose a general approach to thinking about probabilities in regards to ecological interactions, with a strong focus on food webs, and call for better definitions and conceptualizations of probabilistic ecological networks, both at the local and regional scales.

# 1 Introduction

2 Cataloging species interactions across space is a gargantuan task. At the core of this challenge lies the  
3 spatiotemporal variability of ecological networks (Poisot *et al.* 2012, 2015), which makes documenting the  
4 location and timing of interactions difficult. Indeed, it is not sufficient to know that two species have the  
5 biological capacity to interact to infer the realization of their interaction at a specific time and space  
6 (Dunne 2006). Taking food webs as an example, a predator species and its potential prey must first  
7 co-occur in order for a trophic interaction to take place (Blanchet *et al.* 2020). They must then encounter,  
8 which is conditional on their relative abundances in the ecosystem and the matching of their phenology  
9 (Poisot *et al.* 2015). Finally, the interaction occurs only if the predators have a desire to consume their prey  
10 and are able to capture and ingest them (Pulliam 1974). Environmental (e.g. temperature and presence of  
11 shelters) and biological (e.g. physiological state of both species and availability of other prey species)  
12 factors contribute to this variability by impacting species co-occurrence (Araujo *et al.* 2011) and the  
13 realization of their interactions (Poisot *et al.* 2015). In this context, the development of computational  
14 methods in ecology can help alleviate the colossal sampling efforts required to document species  
15 interactions across time and space (Strydom *et al.* 2021). Having a better portrait of species interactions  
16 and the emerging structure of their food webs is important since it lays the groundwork for understanding  
17 the functioning, dynamics, and resilience of ecosystems worldwide (e.g., Proulx *et al.* 2005; Pascual *et al.*  
18 2006; Delmas *et al.* 2019).

19 The recognition of the intrinsic variability of species interactions and the emergence of numerical  
20 methods have led ecologists to rethink their representation of ecological networks, slowly moving from a  
21 binary to a probabilistic view of species interactions (Poisot *et al.* 2016). This has several benefits. For  
22 example, probabilities represent the limit of our knowledge about species interactions and can inform us  
23 about the expected number of interactions and emerging network properties despite this limited  
24 knowledge (Poisot *et al.* 2016). They are also very helpful in predicting the spatial distribution of species  
25 within networks (Cazelles *et al.* 2016) and the temporal variability of interactions (Poisot *et al.* 2015),  
26 generating new ecological data (e.g., Strydom *et al.* 2022), and identifying priority sampling locations of  
27 species interactions (see Andrade-Pacheco *et al.* 2020 for an ecological example of a sampling optimization  
28 problem). Moreover, the high rate of false negatives in ecological network data, resulting from the  
29 difficulty of witnessing interactions between rare species, makes it hard to interpret non-observations of

species interactions ecologically (Catchen *et al.* 2023). Using probabilities instead of yes-no interactions accounts for these observation errors; in that case, only forbidden interactions (Jordano *et al.* 2003; Olesen *et al.* 2010) would have a probability value of zero (but see Gonzalez-Varo & Traveset 2016). Many measures have been developed to describe the structure (Poisot *et al.* 2016) and diversity (Ohlmann *et al.* 2019; Godsoe *et al.* 2022) of probabilistic interactions, which shows the potential of this framework in the study of a variety of ecological phenomena.

However, representing species interactions probabilistically can also be challenging. Beyond methodological difficulties in estimating these numbers, there are important conceptual challenges in defining what we mean by “probability of interactions.” To the best of our knowledge, because the building blocks of this mathematical representation of food webs are still being laid, there is no clear definition found in the literature. This is worrisome, since working with probabilistic species interactions without clear guidelines could be misleading as much for field ecologists as for computational ecologists who use and generate these data. In this contribution, we outline different ways to define and interpret interactions probabilities in network ecology and propose an approach to thinking about them. These definitions mostly depend on the study system (e.g. local network or metaweb) and on the method used to generate them. We show that different definitions can have different ecological implications, especially regarding spatial, temporal, and taxonomic scaling. Although we will focus on food webs, our observations and advice can be applied to all types of ecological networks, from plant-pollinator to host-parasite networks. Specifically, we argue that probabilities should be better documented, defined mathematically, and used with caution when describing species interactions.

## Definitions and interpretations

### Food-web representations

The basic unit of food webs and other ecological networks are individuals that interact with each others (e.g., by predation; Elton 2001), forming individual-based networks. The aggregation of these individuals into more or less homogeneous groups (e.g., populations, species, trophic species, families) allows us to represent networks at broader scales, which impacts the properties and behaviour of these systems (Guimarães 2020). A network’s nodes can thus designate distinct levels of organization, whereas the edges linking these nodes can describe a variety of interaction measures. When using a Boolean (yes-no)

58 representation of biotic interactions, the observation that one individual from group (or node)  $i$  interacts  
 59 with another individual from group  $j$  is enough to set the interaction  $A_{i,j}$  to 1. This simplified  
 60 representation of food webs is a highly valuable source of ecological information (Pascual *et al.* 2006) even  
 61 though it overlooks important factors regarding interaction strengths. These, in turn, can be represented  
 62 using weighted interactions, which better describe the energy flows, demographic impacts or frequencies  
 63 of interactions between nodes (Berlow *et al.* 2004; Borrett & Scharler 2019), with  $A_{i,j} \in \mathbb{N}$  or  $\mathbb{R}$  depending  
 64 on the measure. For example, they can be used to estimate the average number of prey individuals  
 65 consumed by the predators in a given time period. This extra amount of ecological information typically  
 66 comes at a cost of greater sampling effort or data requirement in predictive models (Strydom *et al.* 2021),  
 67 which can lead to high uncertainties when building these types of networks. Therefore, important  
 68 methodological and conceptual decisions must be made when sampling and building food webs.

69 The uncertainty and spatiotemporal variability of both types of trophic interactions (Boolean and  
 70 weighted) can be represented probabilistically. On one hand, Boolean interactions follow a Bernoulli  
 71 distribution  $A_{i,j} \sim \text{Bernoulli}(p)$ , with  $p$  being the probability of interactions. The only two possible  
 72 outcomes are the presence ( $A_{i,j} = 1$ ) or absence ( $A_{i,j} = 0$ ) of an interaction between the two nodes.  
 73 Weighted interactions, on the other hand, can follow various probability distributions depending on the  
 74 measure used. In this case, the event's outcome is the value of interaction strength. For instance, weights  
 75 can follow a Poisson distribution  $A_{i,j} \sim \text{Poisson}(\lambda)$  when predicting frequencies of interactions between  
 76 pairs of nodes, with  $\lambda$  being the expected rate of interaction. The definition and interpretation of  
 77 parameters like  $p$  and  $\lambda$  are inextricably linked to environmental and biological factors such as species  
 78 relative abundance, traits, area, and time, depending on the type of interaction. Because Boolean species  
 79 interactions are much more documented in the literature, our primary focus in this contribution will be on  
 80 addressing the challenges in defining and interpreting  $p$  for pairwise species interactions.

81 The first aspect to take into consideration when estimating or interpreting probabilities of interactions is  
 82 knowing if they describe the likelihood of potential or realized interactions. A potential interaction is  
 83 defined as the biological capacity of two species to interact (i.e., the probability that they *can* interact)  
 84 whereas a realized interaction refers to the materialization or observation of this interaction in a delineated  
 85 space and time period (i.e., the probability that they interact). Here, we will use the terms *metaweb* to  
 86 designate networks of potential interactions and *local networks* for those of realized interactions. Frequent  
 87 confusion arises among ecologists over the use of these two terms, especially in a probabilistic context.

88 Indeed, in many studies of probabilistic ecological networks, it remains unclear when authors describe  
89 potential or realized interactions, or when so-called probabilities are actually *interaction scores*. Likewise,  
90 probabilistic potential interactions are often used as realized interactions (and conversely), even when the  
91 type of interaction is clearly indicated. We believe that a better understanding of these differences and  
92 concepts would alleviate interpretation errors and help ecologists use these numbers more appropriately.

### 93 **Probabilistic metawebs**

94 Metawebs are networks of potential interactions, representing the probability that two taxa can interact  
95 regardless of biological plasticity, environmental variability or co-occurrence. They are the network  
96 analogue to the species pool, where local networks originate from a subset of both species (nodes) and  
97 interactions (edges). For this reason, they have been initially designed for broad spatial, temporal, and  
98 taxonomic scales (e.g, species food webs at the continental scale). However, in the next section, we argue  
99 that this concept can also be used at smaller scales, with similar ecological meaning.

100 We can express the probability that two taxa  $i$  and  $j$  can interact in a metaweb  $M$  as

$$P_M(i \rightarrow j), \quad (1)$$

101 which is context independent. In other words, the probability that two species can interact is not  
102 contingent on location, time, and environmental factors. Nevertheless, one aspect of a metaweb that could  
103 be conditional on these factors is the list of species (or taxa) it is built from when assembled for a specific  
104 region.

105 Starting from a selected set of species, metawebs can be built using different data sources, including  
106 literature review, field work, and predictive models (e.g., the metaweb of Canadian mammals inferred by  
107 Strydom *et al.* 2022). Every pair of species that has been observed to interact at least once can be given a  
108 probability of interaction of 1; we know that they *can* interact. This means that rare interactions can  
109 technically be given high probabilities in the metaweb. Unobserved interactions, on the other hand, are  
110 given lower probabilities, going as low as 0 for forbidden links. Two important nuances must however be  
111 made here. Because of observation errors due to taxonomic misidentifications and ecological  
112 misinterpretations, many observations of interactions are actually false positives. Similarly, forbidden

interactions can be false negatives if e.g. they have been assessed for specific phenotypes, locations or time. Implementing a Bayesian framework, which updates prior probabilities of interactions with empirical data, could lessen these errors.

## Probabilistic local networks

As opposed to metawebs, probabilistic local food webs represent the likelihood that two species will interact at a specific location and within a given time period; in other words, they are context dependant. To be specific, space is defined here as the geographic coordinates ( $x, y$ ) of the spatial boundaries delineating the system, where  $t$  is the time interval ( $t$ ) during which interactions were sampled or for which they were predicted. We put these variables in bold to emphasize that they are not single values, but rather continued dimensions that could be outlined differently depending on the study system. Regardless of how they were defined, they always delineate a specific area  $A$  and duration  $D$ . These could refer to the sampled area and duration or to the targeted location and time period.

Many factors could be taken into consideration when estimating the probability that a predator species  $i$  interacts with a given prey species  $j$  locally. One of the most important is species co-occurrence, which describes if both species can be found at location and time ( $x, y, t$ ). Surely, the probability that the interaction is realized must be 0 when species do not co-occur. Similarly, species relative abundance is another important predictor of the probability of interaction, because it impacts the probability that species will randomly encounter (Canard *et al.* 2012; Canard *et al.* 2014; Poisot *et al.* 2015). Interaction probabilities can also be conditional on other biological and environmental variables, such as temperature, precipitation, presence of shelters, phenotypic plasticity, phenology, and presence of other interacting species in the network. These conditions can affect species traits, which greatly impact the likelihood of interactions (Poisot *et al.* 2015). When probabilities of interactions are estimated using specific values of these variables, ecologists should make them explicit in their metadata, preferably using mathematical equations to avoid any ambiguity. In the next section, we will see examples of why this matters when it comes to spatial, temporal, and taxonomic scaling of biotic interactions. Here, we will use the variable  $\Omega$  as a substitute for the biological and ecological context in which interaction probabilities were estimated. This gives us the following equation for the probability that species (or taxa)  $i$  and  $j$  interact in a local network  $N$ :

$$P_N(i \rightarrow j | \mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{\Omega}). \quad (2)$$

## From potential to realized interactions

Even though the spatiotemporal variation of interactions is not considered in metawebs, they can still be useful to reconstruct local networks of realized interactions. Indeed, local networks are formed from subsets of their metaweb (called subnetworks), which are obtained by selecting a subset of both species and interactions (Dunne 2006). Because a community's composition is arguably easier to sample (or predict) than its interactions, the biggest challenge is to sample links from the metaweb. This becomes a conceptual issue when we consider how potential and realized interactions differ. Despite these concerns, metawebs remain an important source of ecological information that can be leveraged for realistically predicting spatially explicit networks. First, metawebs set the upper limit of species interactions (McLeod *et al.* 2021), i.e. the probability that two species interact at a specific location is always lower or equal to the probability of their potential interaction:

$$P_N(i \rightarrow j | \mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{\Omega}) \leq P_M(i \rightarrow j). \quad (3)$$

Therefore, building local networks from their metaweb keeping the same values of interaction probability would generate systematic biases in the prediction. Second, the structure of local networks is constrained by the one of their metaweb (Saravia *et al.* 2022). This suggests that a metaweb not only constraints the pairwise interactions of its corresponding local networks, but also their emerging properties. Inferring the structure of local networks from the metaweb could thus help estimate more realistically the likelihood that potential interactions are realized and observed locally (Strydom *et al.* 2021).

## Scaling

- We should expect a certain number of interactions to be realized depending on the probability value. This number depends on the number of trials, which also depends on the ecological context (e.g., environmental conditions, scale) in which probabilities were estimated. This is in contrast with probabilities in metawebs.



## 164 **Spatial and temporal scales**

165 How do interaction probabilities scale spatially and temporally?

- 166 • Why do probabilistic local food webs scale with area and time but not probabilistic metawebs?
- 167 • In metawebs, interaction probabilities do not scale with space and time because they depend solely
- 168 on the biological capacity of two species to interact.
- 169 • In local food webs, interaction probabilities scale with space and time because there are more
- 170 opportunities of interactions (e.g., more environmental conditions) in a larger area and longer time
- 171 period.
- 172 • What are some network area relationships in probabilistic local food webs?
- 173 • We know that local networks can inform regional networks. However, can regional networks
- 174 inform local networks?

175 **Figure:** Empirical example of the association between the number of interactions in realized local food  
176 webs and the number of interactions in the corresponding species subnetworks of regional networks. We  
177 should expect the interaction to be linear below the 1:1 line.

178 **Papers:** there might be something in these McLeod *et al.* (2020); McLeod *et al.* (2021); Wood *et al.* (2015)

## 179 **Taxonomic scale**

180 How do interaction probabilities scale taxonomically?

- 181 • There are different biological interpretations of probabilities in food webs at the individual level and
- 182 at higher taxonomic levels.
- 183 • How does the scaling up of the nodes from an individual to population to any higher taxonomic
- 184 group change our interpretation of interaction probabilities? How does the aggregation change our
- 185 interpretation?
- 186 • How is it similar and different to spatial and temporal scaling? Basically, all kinds of scaling are just
- 187 different ways to aggregate individuals or nodes.

188 **Figure:** Conceptual figure of how a scale up of the nodes from an individual to a population to any higher  
189 taxonomic group change our interpretation of the probability of interaction.

## 190 **Concluding remarks**

191 Here we present some advice moving forward.

- 192 • What can we learn from other systems/fields (e.g., social networks, probabilistic graph theory)?
- 193 • What even are probabilities? What is the probability that we will ever know the answer to that?
- 194 • Be careful of how we define probabilities. Be sure to be explicit about these things. Be sure to specify
- 195 the type of interaction, the spatial, temporal, and taxonomic scale when presenting new data on
- 196 interaction probabilities. We need better metadata documentation.
- 197 • Be careful to use and manipulate interaction probabilities properly depending on how they were
- 198 defined and obtained. Different interpretations imply different scaling, and thus different ways to
- 199 manipulate these numbers.
- 200 • Maybe mention thinking about a workflow to predict probabilistic local food webs from probabilistic
- 201 metawebs.

## 202 **Acknowledgement**

203 We acknowledge that this study was conducted on land within the traditional unceded territory of the  
204 Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwininiwak nations.  
205 This work was supported by the Institute for Data Valorisation (IVADO) and the Natural Sciences and  
206 Engineering Research Council of Canada (NSERC) Collaborative Research and Training Experience  
207 (CREATE) program, through the Computational Biodiversity Science and Services (BIOS<sup>2</sup>) program. A  
208 special thank to all members of the Black Holes and Revelations working group (organized by BIOS<sup>2</sup>) for  
209 their insightful discussions and valuable feedbacks on this manuscript.

## 210 **References**

- 211 Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meïté, A., Juziwele, L., *et al.* (2020). Finding  
212 hotspots: Development of an adaptive spatial sampling approach. *Scientific Reports*, 10, 10939.
- 213 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to  
214 assess the impacts of climate change. *Ecography*, 34, 897–908.

215 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004).  
 216 Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.

217 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions.  
 218 *Ecology Letters*, 23, 1050–1063.

219 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and  
 220 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.

221 Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014). Empirical  
 222 Evaluation of Neutral Interactions in Host-Parasite Networks. *The American Naturalist*, 183, 468–479.

223 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of  
 224 Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7, e38295.

225 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false  
 226 negatives when sampling species interaction networks.

227 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in  
 228 interaction networks. *Theoretical Ecology*, 9, 39–48.

229 Delmas, E., Besson, M., Brice, M.-H., Burkle, L.A., Dalla Riva, G.V., Fortin, M.-J., *et al.* (2019). Analysing  
 230 ecological networks of species interactions. *Biological Reviews*, 94, 16–36.

231 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and*  
 232 *dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.

233 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.

234 Godsoe, W., Murray, R. & Iritani, R. (2022). Species interactions and diversity: A unified framework using  
 235 Hill numbers. *Oikos*, n/a, e09282.

236 Gonzalez-Varo, J.P. & Traveset, A. (2016). The Labile Limits of Forbidden Interactions. *Trends in Ecology &*  
 237 *Evolution*, 31, 700–710.

238 Guimarães, P.R. (2020). The Structure of Ecological Networks Across Levels of Organization. *Annual*  
 239 *Review of Ecology, Evolution, and Systematics*, 51, 433–460.

240 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). Invariant properties in coevolutionary networks of  
 241 plant-animal interactions. *Ecology Letters*, 6, 69–81.

242 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021). Sampling and  
 243 asymptotic network properties of spatial multi-trophic networks. *Oikos*, n/a.

244 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). Effects of species traits, motif profiles, and environment on  
 245 spatial variation in multi-trophic antagonistic networks. *Ecosphere*, 11, e03018.

246 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). Diversity indices  
 247 for ecological networks: A unifying framework using Hill numbers. *Ecology Letters*, 22, 737–747.

248 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). Missing and  
 249 forbidden links in mutualistic networks. *Proceedings of the Royal Society B: Biological Sciences*, 278,  
 250 725–732.

251 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food*  
 252 *Webs*. Oxford University Press, USA.

253 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). The dissimilarity of species  
 254 interaction networks. *Ecology Letters*, 15, 1353–1361.

255 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). The structure of  
 256 probabilistic networks. *Methods in Ecology and Evolution*, 7, 303–312.

257 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary  
 258 through space and time. *Oikos*, 124, 243–251.

259 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). Network thinking in ecology and evolution. *Trends*  
 260 *in Ecology & Evolution*, SPECIAL issue: BUMPER book REVIEW, 20, 345–353.

261 Pulliam, H.R. (1974). On the Theory of Optimal Diets. *The American Naturalist*, 108, 59–74.

262 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). Ecological network  
 263 assembly: How the regional metaweb influences local food webs. *Journal of Animal Ecology*, 91,  
 264 630–642.

265 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). Food web  
 266 reconstruction through phylogenetic transfer of low-rank network representation. *Methods in Ecology*  
 267 *and Evolution*, 13.

268 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). A  
 269 roadmap towards predicting species interaction networks (across space and time). *Philosophical*

270      *Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.

271      Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of  
272      sampling on food web structure. *Ecology and Evolution*, 5, 3769–3782.