

The biological interpretation of probabilistic food webs

Francis Banville^{1,2,3,‡} Tanya Strydom^{1,3,‡} Timothée Poisot^{1,3}

¹ Université de Montréal ² Université de Sherbrooke ³ 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 8, 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 on the same territory in order for a trophic interaction to take place (Blanchet *et al.* 2020). They
8 must then encounter, which is conditional on their relative abundances in the ecosystem and the
9 matching of their phenology (Poisot *et al.* 2015). Finally, the interaction occurs only if the predators have a
10 desire to consume their prey and are able to capture and ingest them (Pulliam 1974). Environmental (e.g.
11 temperature and presence of shelters) and biological (e.g. physiological state of both species and
12 availability of other prey species) factors contribute to this variability by impacting species co-occurrence
13 (Araujo *et al.* 2011) and the realization of their interactions (Poisot *et al.* 2015). In this context, the
14 development of computational methods in ecology can help alleviate the colossal sampling efforts
15 required to document species interactions across time and space (Strydom *et al.* 2021). Having a better
16 portrait of species interactions and the emerging structure of their food webs is important since it lays the
17 groundwork for understanding the functioning, dynamics, and resilience of ecosystems worldwide (e.g.,
18 Proulx *et al.* 2005; Pascual *et al.* 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 λ being the expected rate of interaction. The definition and interpretation of
 77 parameters like p and λ 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 using 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
 85 delineated space and time period (i.e., the probability that they *interact*). Here, we will use the terms
 86 *metaweb* to designate networks of potential interactions and *local networks* for those of realized
 87 interactions. Frequent confusion arises among ecologists over the use of these two terms, especially in a

88 probabilistic context. Indeed, in many studies of probabilistic ecological networks, it remains unclear
89 when authors describe potential or realized interactions, or when so-called probabilities are actually
90 *interaction scores*. A better understanding of these differences would alleviate interpretation errors and
91 help ecologists use these numbers more appropriately.

92 **Probabilistic metawebs**

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

99 We can express the probability that two taxa i and j can interact as

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

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

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

113 Bayesian framework, which updates prior probabilities of interactions with empirical data, could lessen
114 these errors.

115 One potential application of metawebs is the prediction of networks of realized interactions. Even though
116 the spatiotemporal variation of interactions is not considered in metawebs, they can nevertheless be a
117 useful tool to reconstruct local networks. First, metawebs set the upper limit of species interactions,
118 i.e. the probability that two species interact at a specific location is always lower than the probability of
119 their potential interaction ($P(i \rightarrow j) < P(i \rightarrow j|x_1, x_2, \dots)$). This implies that predicting local networks
120 from metawebs realistically might need be done by filtering out interactions from its subnetworks.
121 Second, the structure of local networks is constrained by the one of their metaweb. Inferring their
122 structure from the metaweb first could help predicting which potential interactions are not realized.
123 Ultimately, we believe that predicting probabilistic local networks from metawebs must be done with
124 caution, as long as uncertainty is adequately propagated.

125 **Probabilistic local networks**

$$P(i \rightarrow j|C, A, N, E, t)$$

126 i = predator j = prey C = co-occurrence (and relative abundance?) A = area N = relative abundance (or
127 network?) E = environment (including network?) t = time

128 What does a probability in the context of a local network mean? A cautionary tale of how we define
129 probabilities.

- 130 • It means the probability that two taxa will interact at a given location.
- 131 • What do we mean by saying that two taxa will interact? We usually mean that at least one individual
132 from one group will interact with (e.g., predate) at least one other individual from the other group.
- 133 • The probability is conditional on the environmental and local abundance contexts.
- 134 • We should expect a certain number of interactions to be realized depending on the probability value.
135 This number depends on the number of trials, which also depends on the ecological context (e.g.,
136 environmental conditions, scale) in which probabilities were estimated. This is in contrast with
137 probabilities in metawebs.

138 **Scaling**

139 **Spatial and temporal scales**

140 How do interaction probabilities scale spatially and temporally?

- 141 • Why do probabilistic local food webs scale with area and time but not probabilistic metawebs?
- 142 • In metawebs, interaction probabilities do not scale with space and time because they depend solely
143 on the biological capacity of two species to interact.
- 144 • In local food webs, interaction probabilities scale with space and time because there are more
145 opportunities of interactions (e.g., more environmental conditions) in a larger area and longer time
146 period.
- 147 • What are some network area relationships in probabilistic local food webs?
- 148 • We know that local networks can inform regional networks. However, can regional networks
149 inform local networks?

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

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

154 **Taxonomic scale**

155 How do interaction probabilities scale taxonomically?

- 156 • There are different biological interpretations of probabilities in food webs at the individual level and
157 at higher taxonomic levels.
- 158 • How does the scaling up of the nodes from an individual to population to any higher taxonomic
159 group change our interpretation of interaction probabilities? How does the aggregation change our
160 interpretation?
- 161 • How is it similar and different to spatial and temporal scaling? Basically, all kinds of scaling are just
162 different ways to aggregate individuals or nodes.

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

165 **Concluding remarks**

166 Here we present some advice moving forward.

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

177 **Acknowledgement**

178 We acknowledge that this study was conducted on land within the traditional unceded territory of the
179 Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwininiwak nations.
180 This work was supported by the Institute for Data Valorisation (IVADO) and the Natural Sciences and
181 Engineering Research Council of Canada (NSERC) Collaborative Research and Training Experience
182 (CREATE) program, through the Computational Biodiversity Science and Services (BIOS²) program. A
183 special thank to all members of the Black Holes and Revelations working group (organized by BIOS²) for
184 their insightful discussions and valuable feedbacks on this manuscript.

References

- Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meité, A., Juziwelo, L., *et al.* (2020). Finding hotspots: Development of an adaptive spatial sampling approach. *Scientific Reports*, 10, 10939.
- Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, 34, 897–908.
- Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004). Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.
- Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23, 1050–1063.
- Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.
- Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false negatives when sampling species interaction networks.
- Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in interaction networks. *Theoretical Ecology*, 9, 39–48.
- Delmas, E., Besson, M., Brice, M.-H., Burkle, L.A., Dalla Riva, G.V., Fortin, M.-J., *et al.* (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94, 16–36.
- Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.
- Godsoe, W., Murray, R. & Iritani, R. (2022). Species interactions and diversity: A unified framework using Hill numbers. *Oikos*, n/a, e09282.
- Gonzalez-Varo, J.P. & Traveset, A. (2016). The Labile Limits of Forbidden Interactions. *Trends in Ecology & Evolution*, 31, 700–710.
- Guimarães, P.R. (2020). The Structure of Ecological Networks Across Levels of Organization. *Annual Review of Ecology, Evolution, and Systematics*, 51, 433–460.

211 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). Invariant properties in coevolutionary networks of
 212 plantanimal interactions. *Ecology Letters*, 6, 69–81.

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

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

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

219 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). Missing and
 220 forbidden links in mutualistic networks. *Proceedings of the Royal Society B: Biological Sciences*, 278,
 221 725–732.

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

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

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

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

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

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

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

236 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). A
 237 roadmap towards predicting species interaction networks (across space and time). *Philosophical*
 238 *Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.

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