

# The missing link: discerning true from false negatives when sampling species interaction networks

[Michael D. Catchen](#) · [Timothée Poisot](#) · [Laura Pollock](#) · [Andrew Gonzalez](#) ·

<sup>1</sup> McGill University · <sup>2</sup> Québec Centre for Biodiversity Sciences · <sup>3</sup> Université de Montréal

## Correspondance to:

Michael D. Catchen — [michael.catchen@mail.mcgill.ca](mailto:michael.catchen@mail.mcgill.ca)

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



Last revision: *September 25, 2022*

# 1 Introduction

2 Understanding which and how species interact underlies many questions in evolutionary biology and  
3 ecology, but also is an increasing imperative to both mitigate the anthropogenic change on Earth's  
4 biodiversity (Jordano 2016a; Makiola *et al.* 2020) and to predict potential spillover of zoonotic disease  
5 (Becker *et al.* 2021). Over the past few decades biodiversity data has become increasingly available:  
6 remote-sensing has enabled collection of data on spatial scales and resolutions previously unimaginable,  
7 automation of in-situ sensing (Stephenson 2020), and adoption of open data practices (Kenall *et al.* 2014)  
8 have substantially amount of data available to ecologists. Still, widespread data about species *interactions*  
9 remains elusive (**TIM\_DATASET\_PAPER?**). Often, observing an interaction between two species  
10 requires human sampling, typically because remote sampling methods can only detect co-occurrence, and  
11 this itself is not necessarily indicative of interaction (Blanchet *et al.* 2020). This constraint induces biases  
12 on species interaction data subject to the spatial and temporal scales that humans can feasibly sample.

13 *Sampling effort* and its impact on the resulting data collected from ecosystems has encouraged a long  
14 history of discourse. The recorded number of species in a sample depends on the total number of  
15 observations (Walther *et al.* 1995; Willott 2001), as do estimates of population abundance (Griffiths 1998).  
16 This has motivated more quantitatively robust approaches to account for error in sampling data in many  
17 contexts: to determine if a given species is extinct (Boakes *et al.* 2015), to determine sampling design  
18 (Moore & McCarthy 2016), and to measure species richness across large scales (Carlson *et al.* 2020). In the  
19 context of interactions, the initial concern was the compounding effects of limited sampling effort  
20 combined with the amalgamation of data (across both study sites and across taxonomic scales) could lead  
21 any empirical set of observations to inadequately reflect the reality of how species interact (Paine 1988).  
22 Martinez *et al.* (1999) showed that in a plant-endophyte trophic network, network connectance is robust to  
23 sampling effort, but this done in the context of a system for which observation of 62,000 total interactions  
24 derived from 164,000 plant-stems was feasible. In some systems (e.g. megafauna food-webs) this many  
25 observations is either impractical or infeasible due to the absolute abundance of the species in question.

26 Because we cannot feasibly observe all (or even many) of the interactions that occur in nature, our  
27 samples end up capturing only a small fraction of those interactions. This means we can be reasonably  
28 confident two species actually interact if we have a record of it, but not at all confident that two species *do*  
29 *not* interact if we have no record of those species observed together. In other words, it is difficult to

30 distinguish true-negatives (two species *never* interact) from *false-negatives* (two species interact in some  
31 capacity, but we have not observed it). This is then amplified as the interaction data we have is  
32 geographically toward the usual suspects (Poisot *et al.* 2021a), This noise in data has practical  
33 consequences for answering questions about species interactions (de Aguiar *et al.* 2019)—these  
34 false-negatives could go on to effect the inferences we make about network properties and relations  
35 among species, and our predictions about how species will interact in the future.

36 This is compounded by semantic confusion about the definition of “interaction”. Here distinguish  
37 between: a species *occurring*, a species being *observed occurring*, two species being observed *co-occurring*,  
38 and two species being observed *interacting* (fig. ??). In this manuscript, we refer to species either as  
39 “interacting”—two species co-occur (and, at least sometimes, interact)—or “not-interacting” (two species  
40 that, regardless of whether they co-occur, neither exhibits any meaningful effect on the biomass of the  
41 other). In fig. ?? we see that, under our definition, observing two species co-occurring is a prerequisite for  
42 observing an interaction between two species.

43 But species are not observed with equal probability but instead in proportion to their relative  
44 biomass—you are much more likely to observe a species of high relative abundance than one of very low  
45 relative abundance (Poisot *et al.* 2015). This assumes that there are no associations in species  
46 co-occurrence due to an interaction (perhaps because this interaction is “important” for both species)  
47 (Cazelles *et al.* 2016), but here we show increasing strength of associations leads to increasing probability  
48 of false-negatives in interaction data. Further observed co-occurrence is often equated with meaningful  
49 interaction strength, but this is not necessarily the case (Blanchet *et al.* 2020; Strydom *et al.* 2021). Bears  
50 and salmon *interact*—a bear and the microbes in the soil of a dens interact, but less so.

51 Here, we show that the probability of observing a actual “non-interaction” between species depends on  
52 sampling effort, and suggest that surveys of species interactions can benefit from simulation modeling of  
53 detection probability (Jordano 2016b). We demonstrate that the realized false-negative rate of interactions  
54 is directly related the relative abundance of a particular species, relationship between total sampling effort  
55 (the total count of all individuals of all species seen) and false-negative rate. questions we pose and attempt  
56 to answer are: 1) How many times do you have to observe a non-interaction between two species to be  
57 confident in saying that is a true negative? 2) How “wrong” are the measurements of network structure as  
58 a function of false-negative probability? and lastly 3) How do false-negatives impact our ability to make  
59 reliable predictions about interactions? We show that positive associations in co-occurrence data can

60 increase realized probability of false negatives, and demonstrate these positive associations are present in  
61 two spatially-replicated systems. We conclude by suggesting that simulation of sampling effort and species  
62 occurrence can and should be used to help design surveys of species diversity (Moore & McCarthy 2016),  
63 and by advocating use of null models like those presented here as a tool for guiding design of surveys of  
64 species interactions, and for modeling detection error in predictive ecological models.

## 65 **How many observations of a non-interaction do we need to classify it as a** 66 **true negative?**

67 To answer the titular question of this section, we present a naive model of interaction detection: we  
68 assume that every interacting pair of species is incorrectly observed as a not-interacting with an  
69 independent and fixed probability, which we denote  $p_{fn}$  and subsequently refer to as the False-Negative  
70 Rate (FNR). If we observe the same species not-interacting  $N$  times, then the probability of a true-negative  
71 (denoted  $p_{tn}$ ) is given by  $p_{tn} = 1 - (p_{fn})^N$ . This relation (callend the geometric distribution, a special case  
72 of the negative-binomial distribution) is shown in fig. ?? for varying values of the false negative rate  $p_{fn}$ .  
73 This illustrates a fundamental link between our ability to reliably say an interaction doesn't  
74 exist— $p_{tn}$ —and the number of times we have observed a given species. In addition, note that there also is  
75 no non-zero  $p_{fn}$  for which we can ever *prove* that an interaction does not exist—no matter how many  
76 observations of non-interaction  $N$  we have,  $p_{tn} < 1$ .

77 From fig. 1 (A) (and general intuition) it is clear that the more times we see two species *occurring*, but *not*  
78 interacting, the more likely the interaction is a true negative. But how does one decide what this threshold  
79 of number of observations should be when planning to sample a given system? If false-negative rates  
80 presented in fig. 1 seem unrealistically high, consider that species are not observed independent of their  
81 relative abundance. In the next section we demonstrate that distribution of abundance in ecosystems can  
82 lead to realized values of  $p_{fn}$  similar to those in fig. ?? for species with low relative abundance, simply as a  
83 artifact of sampling effort.

## 84 False-negatives as a product of relative abundance

85 Here we show the realized false-negative rate of species interactions changes drastically with sampling  
86 effort, largely due to the intrinsic variation of abundances within a community. We do this by simulating  
87 the process of observation of species interactions, applied both to 243 empirical food webs from the  
88 Mangal database (Banville *et al.* 2021) as well as random food-webs generated using the niche model  
89 (Williams & Martinez 2000). Our neutral model of observation assumes each observed species is drawn  
90 from the distribution of those species' abundances at that place and time. Although there is no shortage of  
91 debate as to the processes the govern this distribution of abundances within a community, this abundance  
92 distribution can be reasonably-well described by a log-normal distribution (Volkov *et al.* 2003) (Note that  
93 in addition to the log-normal distribution, we also tested the case where the abundance distribution is  
94 derived from power-law scaling  $Z^{(T_i-1)}$  where  $T_i$  is the trophic level of species  $i$  and  $Z$  is a scaling  
95 coefficient. (Savage *et al.* 2004), which yields the same qualitative behavior, *supplement figure 1*). The  
96 practical consequence of abundance distributions spanning many order of magnitude is seeing two “rare”  
97 species interacting requires two low probability events: observing two rare species *at the same time*.

98 To simulate the process of observation, for an ecological network  $A$  with  $S$  species, we sample abundances  
99 for each species from a standard-log-normal distribution. For each true interaction in  $A$  (i.e.  $A_{ij} = 1$ ) we  
100 estimate the probability of observing both species  $i$  and  $j$  at given place and time by simulating  $n$   
101 observations of individuals, where the species of the individual observed at the 1, 2, ...,  $n$ -th observation is  
102 drawn from the generated log-normal distribution of abundances. For each pair of species  $(i, j)$ , if both  $i$   
103 and  $j$  are observed within the  $n$  observations, the interaction is tallied as a true positive if  $A_{ij} = 1$  and a  
104 false positive otherwise. Similarly, if only one of  $i$  and  $j$  are observed—but *not both*—in these  $n$   
105 observations, but  $A_{ij} = 1$ , this is counted as a false-negative, and a true-negative otherwise.

106 In fig. 1 (C) and (D) we see this model of observation applied to networks generated using the niche model  
107 (Williams & Martinez 2000) across varying levels of species richness, and in (b) applied to 243 food-webs  
108 from the Mangal database. For all niche model simulations in this manuscript, for a given number of  
109 species  $S$  the number of interactions is drawn from the flexible-links model fit to Mangal data (MacDonald  
110 *et al.* 2020), effectively drawing the number of interactions  $L$  for a random niche model food-web as  
111  $L \sim \text{BetaBinomial}(S^2 - S + 1, \mu\phi, (1 - \mu)\phi)$ , where the MAP estimate of  $(\mu, \phi)$  applied to Mangal data  
112 from MacDonald *et al.* (2020) is  $(\mu = 0.086, \phi = 24.3)$ . All simulations were done with 500 independent

113 replicates per unique number of observations  $n$ . All analyses presented here are done in Julia v1.6  
114 (Bezanson *et al.* 2015) using both EcologicalNetworks.jl v0.5 and Mangal.jl v0.4 [Banville *et al.* (2021);  
115 ZENODO link TODO]. Note that the empirical data also is, due to the phenomena described here, very  
116 likely to *already* have many false negatives, which is why we are interested in prediction of networks in the  
117 first place—we’ll revisit this in the final section.

118 [Figure 1 about here.]

119 In panel (c) of fig. ??, we show the expected number of total observations needed to obtain a “goal”  
120 number of observations (colors) of a particular “focal” species. As an example, if we hypothesize that  $A$   
121 and  $B$  do not interact, and we want to see species  $A$  and  $B$  both co-occurring and not-interacting 10 times  
122 to be confident this is a negative (a la fig. ??), then we need an expected 10,000 observations of all species if  
123 the relative abundance of  $A$  is 0.00125.

124 Empirical data on interactions are subject to the practical limitations of funding and human-work hours,  
125 and therefore existing data tend to fall on the order on 100s or 1000s observations of individuals per site  
126 (Nielsen & Bascompte 2007; Schwarz *et al.* 2020; Resasco *et al.* 2021). Clear aggregation of this data has  
127 proven difficult to find and a meta-analysis of network data and sampling effort seems both pertinent and  
128 necessary, in addition to the effects of aggregation of interactions across taxonomic scales (Gauzens *et al.*  
129 2013; Giacomuzzo & Jordán 2021). Further, from fig. ?? it is evident that the number of species considered  
130 in a study is inseparable from the false-negative rate in that study, and this effect should be taken into  
131 account when designing samples of ecological networks in the future.

132 We conclude this section by advocating for the use of neutral models similar to above to generate  
133 expectations about the number of false-negatives in a data set of a given size. This could prove fruitful  
134 both for designing surveys of interactions (Canard *et al.* 2012), but also because we may want to  
135 incorporate models of observation error into predictive models (Joseph 2020). Additionally, one must  
136 consider the context for sampling—is the goal to detect a particular species  $A$  (as in fig. ?? (c)), or to get a  
137 representative sample of interactions across the species pool? This argument is well-considered when  
138 sampling species (Willott 2001), but has not yet been internalized for designing samples of communities.

## Positive associations can increase the probability of false-negatives

This model above doesn't consider the possibility that there are positive or negative associations which shift the probability of observing two species together due to their interaction (Cazelles *et al.* 2016).

However, here we demonstrate that the probability of observing a false negative can be *higher* if there is some positive association between occurrence of species  $A$  and  $B$ .

If we denote the probability that we observe an interaction we know exists between  $A$  and  $B$  as  $P(AB)$ , and if there is *no* association between the marginal probabilities of observing  $A$  and observing  $B$ , denoted  $P(A)$  and  $P(B)$  respectively, then the probability of observing the interaction  $P(AB) = P(A)P(B)$ . In the other case where there is some positive strength of association between observing both  $A$  and  $B$  because this interaction is "important" for each species, then the probability of observation both  $A$  and  $B$ ,  $P(AB)$ , is greater than  $P(A)P(B)$  as  $P(A)$  and  $P(B)$  are not independent and instead are positively correlated, *i.e.*  $P(AB) > P(A)P(B)$ . In this case, the probability of observing a false negative in our naive model from fig. ?? is  $p_{fn} = 1 - P(AB)$  which due to the above inequality implies  $p_{fn} \geq 1 - P(A)P(B)$  which indicates increasingly greater probability of a false negative as  $P(AB) \rightarrow P(AB) \gg P(A)P(B)$ .

However this does not consider variation in species abundance in space and time, (Poisot *et al.* 2015). If positive or negative associations between species structure variation in the distribution of  $P(AB)$  across space/time, then the spatial/temporal biases induced by data collection would further impact the realized false negative rate, as the probability of false negative would not be constant for each pair of species across sites. To test for this association empirical data, we use two datasets: a set of host-parasite interactions sampled across 51 sites with 327 total taxa (Hadfield *et al.* 2014) and a set of 18 New Zealand freshwater stream food webs with 566 total taxa (Thompson & Townsend 2000). We simply compute the empirical marginal distribution of species occurrence, and compare the product of the marginals,  $P(A)P(B)$ , to the empirical joint distribution  $P(AB)$ .

[Figure 2 about here.]

In fig. 2, both host-parasite system (top) and food-web (bottom) exhibit these positive associations. There is no reason to expect the strength of this association to be the same in different systems. At the moment, computing this metric for all of the networks in the Mangal database proves challenging as most data sets use different taxonomic identifiers, often at different resolutions. These particular datasets (Thompson &

167 Townsend 2000; Hadfield *et al.* 2014) were usable because they already have been sorted to have a fixed  
168 taxonomic backbone (as part of EcologicalNetworks.jl (Banville *et al.* 2021)). Applying this in bulk to  
169 Mangal food-webs presents the difficulty of resolving different taxon identifiers across spatial samples of  
170 species with to different resolutions, which is why we can't simply apply this to the whole Mangal  
171 database—this highlights a general problem of resolving taxonomic identifiers which use different  
172 names and different resolutions in different ecological datasets, which is a problem that needs to be  
173 addressed for computational approaches to scale up to the world of big-ecological-data we hope to build,  
174 although this is a task that may be aided via natural-language-processing methods.

## 175 **The impact of false-negatives on network analysis and prediction**

176 We now transition toward assessing the effects of false negatives in data on the properties of the networks  
177 which we derive from this interaction data, and their effect on models for predicting interactions in the  
178 future.

### 179 **Effects of false-negatives on network properties**

180 Here we simulate the process of observation with error to generate synthetic data with a known  
181 proportion of false negatives, and compare the computed network properties of the original “true”  
182 network to the computed properties of the “observed” network with added false-negatives. In fig. 3 we see  
183 the mean-squared error of connectance, mean degree-centrality, and spectral radius, computed across  
184 2000, 2000, and 300 replicates respectively at each value of the false negative rate  $p_{fn}$ . All replicates use  
185 random food-webs simulated using the niche model (Williams & Martinez 2000) with 100 species and  
186 connectance drawn from the flexible-links model (MacDonald *et al.* 2020) as before.

187 [Figure 3 about here.]

188 We consider three properties: connectance, mean-degree-centrality, and spectral radius, indicative of  
189 local, meso, and global structure. Connectance is effectively a node-level property, a proxy for the degree  
190 distribution. Degree-centrality captures a different aspect of network structure than connectance, more  
191 indicative of meso-level properties that describe local ‘regions’ of nodes interact. Spectral radius



(equivalent to the magnitude of the largest eigenvalue of  $A$ ) is a measure of global structure, and demonstrates the most variability in response to false-negatives. For example, if a false-negative splits a metaweb into two components, spectral-radius becomes the largest eigenvalue of each of those two components. Also note that the form of this error function varies little as species richness changes (*supplemental figure 2*). Practically, *fig. 3* shows us that different scales of measuring network structure vary in their response to false negatives—connectance responds roughly linearly to false negatives, whereas mean-degree-centrality decisively does not. This implies that false-negatives adversely could effect indirect interactions (Williams *et al.* 2002).

## Effects of false negatives on ability to make predictions

Here, we assess the effect of false negatives in data on our ability to make predictions about interactions. The prevalence of false-negatives in data is the catalyst for interaction prediction in the first place, and as a result methods have been proposed to counteract this bias (Stock *et al.* 2017; Poisot *et al.* 2021b). However, it is feasible this could induce too much noise for a interaction prediction model to detect the signal of interaction chance from the latent properties of each species derived from the empirical network if the number of false-negatives in a dataset becomes too overwhelming.

To test this, we use the same predictive model and dataset as in Strydom *et al.* (2021) to predict a metaweb from various empirical slices of the species pool observed across space. This dataset from Hadfield *et al.* (2014) describes host-parasite interaction networks sampled across 51 sites. We partition the data into 80-20 training-test split, and then seed the training data with false negatives varying rates, but crucially do nothing to the test data. We use the same model, a neural-network with 3 feed-forward layers to predict outputs based on features extracted from co-occurrence (see Strydom *et al.* (2021) for more details). The single modification we make to the model is not enforcing a number of positives in the training data as this constraint is eventually impossible for increasing FNR. In *fig. 4*, we show receiving-operating-characteristic (ROC) and precision-recall (PR) curves for the model with varying levels of synthetic false-negatives added to the data.

[Figure 4 about here.]

Interestingly, the performance of the model from Strydom *et al.* (2021) changes little with many added false-negatives, which is good evidence in favor neural-networks as a class of model for interaction

detection. Again, similar to our caveat in the previous section, this data is *already* likely to have many false-negatives, so the effects of adding more as we do in this illustration might be mitigated because there are already non-simulated false-negatives in the original data which impact the model's performance, even in the  $p_{fn} = 0$  case.

We conclude by proposing that simulating the effects of false negatives in this way can serve as an additional validation tool when aiming to detect structural properties of networks using generative null models (Connor *et al.* 2017), or when evaluating the robustness of a predictive model.

## Discussion

Here, we have demonstrated that we expect false-negatives in species interaction datasets purely due to the distribution of abundances within a community. Positive associations between species occurrence (Cazelles *et al.* 2016) can increase the realized false-negative rate if the sampling effort is limited, and we have presented evidence of this non-random structure of co-occurrence in two sets of spatially-replicated ecological network samples. We have also shown that false-negatives can cause varying responses in our measurements of network properties and further could impact our ability to reliably predict interactions, which highlights the need for further research into methods for correcting this bias in existing data (Stock *et al.* 2017). A brief caveat here is that we do not consider the rate of false-positives—in large part false-positives can be explained by misidentification of species, although this could be a relevant consideration in some cases.

What does the future hold for this research? A better understanding of how false-negatives impact our analyses and prediction of ecological networks is a practical necessity. False-negatives could pose a problem for many forms of inference in network ecology. For example, if we aim to measure structural or dynamic stability of a network, or to infer indirect interactions (Williams *et al.* 2002), these estimates could be prone to error if the observed network is not sampled “enough”. What exactly “enough” means is then specific to the application, and should be assessed via methods like those here when designing samples. Further, predictions about network rewiring (Thompson & Gonzalez 2017) due to a changing climate could be error-prone without accounting for interactions that have not been observed but that still may become climatically infeasible.

This highlights the need for a quantitatively robust approach to sample design: for interactions (Jordano

2016b) and otherwise (Carlson *et al.* 2020). The primary takeaway is that when planning the sampling effort across sites, it is necessary to take both the size of the species pool into account. Further, simulating the process of observation could be a powerful tool for planing study design which takes relative abundance into account, and provide a null baseline for detection of interaction strength. A model similar to that here can and should be used to provide a neutral expectation of true-negative probability given a number of observations of individuals at a given place and time.

As we derive from fig. ??, we can never guarantee there are no false-negatives in data. In recent years, there has been interest toward explicitly accounting for false-negatives in models (Stock *et al.* 2017; Young *et al.* 2021), and toward a predictive approach toward interactions —rather than expect that our samples can fully capture all interactions, we know that some interactions between species will not be observed due to finite sampling capacity, and instead we must impute the true metaweb of interactions given a set of samples (Strydom *et al.* 2021). As a result, better predictive approaches are needed for interaction networks (Strydom *et al.* 2021), and building models that explicitly account for observation error is a necessary step forward for predictive ecological models (Johnson & Larremore 2021; Young *et al.* 2021). Neural networks, like the one used to predict interactions in the above section, have been used to reflect hidden states which account for detection error in occupancy modeling (Joseph 2020), and could be integrated in the predictive models of the future.

A better conceptual framework for designing surveys and monitoring networks, and incorporating sequential observations over time is clearly needed (Carlson *et al.* 2020), combined with a meta-analysis of sampling effort and taxonomic resolution in existing data. Incorporating a better understanding of sampling effects and bias on both the future design of biodiversity monitoring systems, and the predictive models we wish to apply to this data, is imperative in making actionable predictions about the future of ecological interactions on our planet.

## References

- preBanville, F., Vissault, S. & Poisot, T. (2021). [Mangal.jl and EcologicalNetworks.jl: Two complementary packages for analyzing ecological networks in Julia](#). *Journal of Open Source Software*, 6, 2721.
- Becker, D.J., Albery, G.F., Sjodin, A.R., Poisot, T., Bergner, L.M., Dallas, T.A., *et al.* (2021). [Optimizing predictive models to prioritize viral discovery in zoonotic reservoirs](#). *bioRxiv*, 2020.05.22.111344.

276 Bezanson, J., Edelman, A., Karpinski, S. & Shah, V.B. (2015). [Julia: A Fresh Approach to Numerical](#)  
277 [Computing](#). *arXiv:1411.1607 [cs]*.

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

280 Boakes, E.H., Rout, T.M. & Collen, B. (2015). [Inferring species extinction: The use of sighting records](#).  
281 *Methods in Ecology and Evolution*, 6, 678–687.

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

284 Carlson, C.J., Dallas, T.A., Alexander, L.W., Phelan, A.L. & Phillips, A.J. (2020). [What would it take to](#)  
285 [describe the global diversity of parasites?](#) *Proceedings of the Royal Society B: Biological Sciences*, 287,  
286 20201841.

287 Cazelles, K., Araújo, M.B., Mouquet, N. & Gravel, D. (2016). [A theory for species co-occurrence in](#)  
288 [interaction networks](#). *Theoretical Ecology*, 9, 39–48.

289 Connor, N., Barberán, A. & Clauset, A. (2017). [Using null models to infer microbial co-occurrence](#)  
290 [networks](#). *PLOS ONE*, 12, e0176751.

291 de Aguiar, M.A.M., Newman, E.A., Pires, M.M., Yeakel, J.D., Boettiger, C., Burkle, L.A., *et al.* (2019).  
292 [Revealing biases in the sampling of ecological interaction networks](#). *PeerJ*, 7, e7566.

293 Gauzens, B., Legendre, S., Lazzaro, X. & Lacroix, G. (2013). Food-web aggregation, methodological and  
294 functional issues. *Oikos*, 122, 1606–1615.

295 Giacomuzzo, E. & Jordán, F. (2021). [Food web aggregation: Effects on key positions](#) (Preprint). *Ecology*.

296 Griffiths, D. (1998). [Sampling effort, regression method, and the shape and slope of sizeabundance](#)  
297 [relations](#). *Journal of Animal Ecology*, 67, 795–804.

298 Hadfield, J.D., Krasnov, B.R., Poulin, R. & Nakagawa, S. (2014). [A Tale of Two Phylogenies: Comparative](#)  
299 [Analyses of Ecological Interactions](#). *The American Naturalist*, 183, 174–187.

300 Johnson, E.K. & Larremore, D.B. (2021). [Bayesian estimation of population size and overlap from random](#)  
301 [subsamples](#). *bioRxiv*, 2021.07.06.451319.

302 Jordano, P. (2016a). [Chasing Ecological Interactions](#). *PLOS Biology*, 14, e1002559.

303 Jordano, P. (2016b). [Sampling networks of ecological interactions](#). *Functional Ecology*.

304 Joseph, M.B. (2020). [Neural hierarchical models of ecological populations](#). *Ecology Letters*, 23, 734–747.

305 Kenall, A., Harold, S. & Foote, C. (2014). [An open future for ecological and evolutionary data?](#) *BMC*  
306 *Evolutionary Biology*, 14, 66.

307 MacDonald, A.A.M., Banville, F. & Poisot, T. (2020). [Revisiting the Links-Species Scaling Relationship in](#)  
308 [Food Webs](#). *Patterns*, 1.

309 Makiola, A., Compson, Z.G., Baird, D.J., Barnes, M.A., Boerlijst, S.P., Bouchez, A., *et al.* (2020). [Key](#)  
310 [Questions for Next-Generation Biomonitoring](#). *Frontiers in Environmental Science*, 7.

311 Martinez, N.D., Hawkins, B.A., Dawah, H.A. & Feifarek, B.P. (1999). [Effects of Sampling Effort on](#)  
312 [Characterization of Food-Web Structure](#). *Ecology*, 80, 1044–1055.

313 Moore, A.L. & McCarthy, M.A. (2016). [Optimizing ecological survey effort over space and time](#). *Methods*  
314 *in Ecology and Evolution*, 7, 891–899.

315 Nielsen, A. & Bascompte, J. (2007). [Ecological networks, nestedness and sampling effort](#). *Journal of*  
316 *Ecology*, 95, 1134–1141.

317 Paine, R.T. (1988). [Road Maps of Interactions or Grist for Theoretical Development?](#) *Ecology*, 69,  
318 1648–1654.

319 Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., *et al.* (2021a). [Global](#)  
320 [knowledge gaps in species interaction networks data](#). *Journal of Biogeography*, jbi.14127.

321 Poisot, T., Ouellet, M.-A., Mollentze, N., Farrell, M.J., Becker, D.J., Albery, G.F., *et al.* (2021b). [Imputing the](#)  
322 [mammalian virome with linear filtering and singular value decomposition](#). *arXiv:2105.14973 [q-bio]*.

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

325 Resasco, J., Chacoff, N.P. & Vázquez, D.P. (2021). [Plantpollinator interactions between generalists persist](#)  
326 [over time and space](#). *Ecology*, 102, e03359.

327 Savage, V.M., Gillooly, J.F., Brown, J.H., West, G.B. & Charnov, E.L. (2004). [Effects of Body Size and](#)  
328 [Temperature on Population Growth](#). *The American Naturalist*, 163, 429–441.

329 Schwarz, B., Vázquez, D.P., CaraDonna, P.J., Knight, T.M., Benadi, G., Dormann, C.F., *et al.* (2020).  
 330 [Temporal scale-dependence of plantpollinator networks](#). *Oikos*, 129, 1289–1302.

331 Stephenson, P. (2020). [Technological advances in biodiversity monitoring: Applicability, opportunities](#)  
 332 [and challenges](#). *Current Opinion in Environmental Sustainability*, 45, 36–41.

333 Stock, M., Poisot, T., Waegeman, W. & De Baets, B. (2017). [Linear filtering reveals false negatives in](#)  
 334 [species interaction data](#). *Scientific Reports*, 7, 45908.

335 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). [A](#)  
 336 [Roadmap Toward Predicting Species Interaction Networks \(Across Space and Time\)](#) (Preprint).  
 337 EcoEvoRxiv.

338 Thompson, P.L. & Gonzalez, A. (2017). [Dispersal governs the reorganization of ecological networks under](#)  
 339 [environmental change](#). *Nature Ecology & Evolution*, 1, 1–8.

340 Thompson, R.M. & Townsend, C.R. (2000). [Is resolution the solution?: The effect of taxonomic resolution](#)  
 341 [on the calculated properties of three stream food webs](#). *Freshwater Biology*, 44, 413–422.

342 Volkov, I., Banavar, J.R., Hubbell, S.P. & Maritan, A. (2003). [Neutral theory and relative species abundance](#)  
 343 [in ecology](#). *Nature*, 424, 1035–1037.

344 Walther, B.A., Cotgreave, P., Price, R.D., Gregory, R.D. & Clayton, D.H. (1995). [Sampling Effort and](#)  
 345 [Parasite Species Richness](#). *Parasitology Today*, 11, 306–310.

346 Williams, R.J., Berlow, E.L., Dunne, J.A., Barabasi, A.-L. & Martinez, N.D. (2002). [Two degrees of](#)  
 347 [separation in complex food webs](#). *Proceedings of the National Academy of Sciences*, 99, 12913–12916.

348 Williams, R.J. & Martinez, N.D. (2000). [Simple rules yield complex food webs](#). *Nature*, 404, 180–183.

349 Willott, S.J. (2001). [Species accumulation curves and the measure of sampling effort](#). *Journal of Applied*  
 350 *Ecology*, 38, 484–486.

351 Young, J.-G., Valdovinos, F.S. & Newman, M.E.J. (2021). [Reconstruction of plantpollinator networks from](#)  
 352 [observational data](#). *Nature Communications*, 12, 3911.

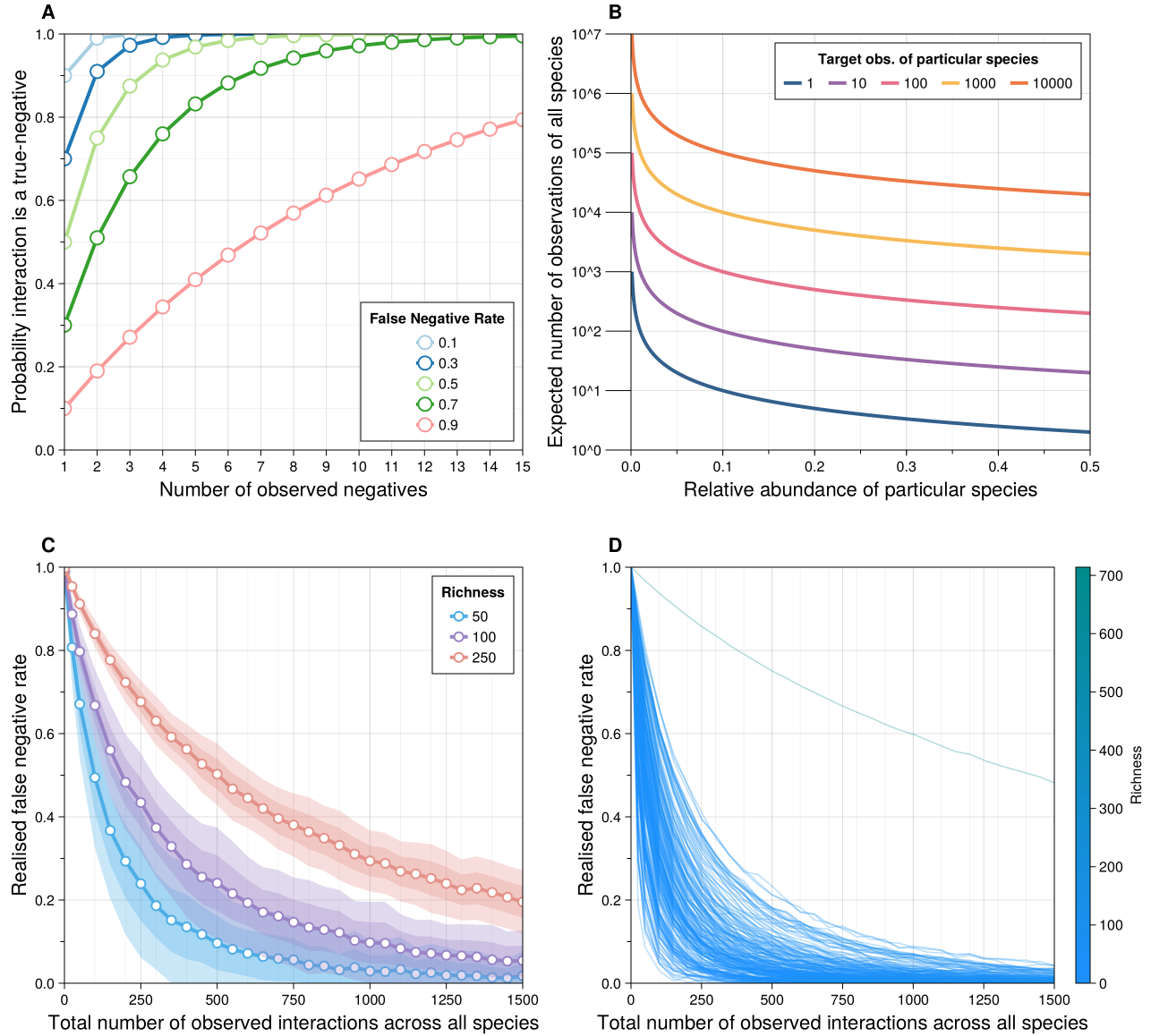


Figure 1: A) The probability an observed interaction is a true negative (y-axis) given how many times it has been sampled as a non-interaction (x-axis). Each color reflects a different value of  $p_{fn}$ , the false-negative rate (FNR). this is effectively the cdf of the negative-binomial distribution with  $r = 1$ . (B) The expected needed observations of all individuals of all species (y-axis) required to obtain a goal number of observations (colors) of a particular species, and a function of the relative abundance of that focal species (x-axis). (C) and (D): False negative rate (y-axis) as a function of total sampling effort (x-axis) and network size, computed using the method described above. For 500 independent draws from the niche model Williams2000SimRul at varying levels of species richness (colors) with connectance drawn according to the flexible-links model MacDonald2020RevLin as described in the main text. For each draw from the niche model, 200 sets of 1500 observations are simulated, for which each the mean false negative rate at each observation-step is computed. Means denoted with points, with  $1\sigma$  in the first shade and  $2\sigma$  in the second. B: empirical food webs from Mangal database in teal, applied to the same process as the A. The outlier on panel B is a 714 species food-web

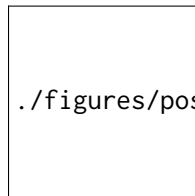


Figure 2: Top: Hadfield, Bottom: NZ Stream Foodwebs. Effectively a version of Cazelles *et al.* (2016) figure 1 panel A. Both distributions have  $\mu \neq 0$  with  $p < 10^{-50}$



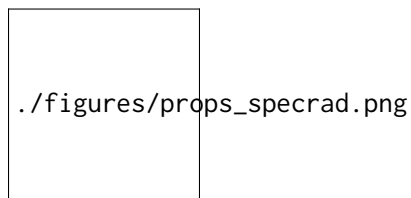


Figure 3: The mean-squared error (y-axis) of various network properties (different colors) across various simulated false-negative rates (x-axis). Means denoted with points, with  $1\sigma$  in the first shade and  $2\sigma$  in the second.

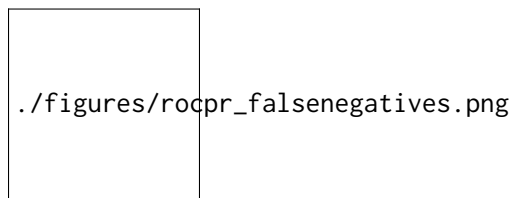


Figure 4: Receiver-operating-characteristic (left) and precision-recall (right) curves for the model on varying levels of false-negatives in the data (colors). For each value of FNR, we run 30 random training/test splits on 80/20 percent of the data. Replica of figure 1 in Strydom *et al.* (2021)