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

Michael D. Catchen [1,2] Timothée Poisot [3,2] Laura Pollock [1,2] Andrew Gonzalez [1,2]

[1] McGill University [2] Québec Centre for Biodiversity Sciences [3] Université de Montréal

Correspondance to:

Michael D. Catchen — `michael.catchen@mail.mcgill.ca`

Last revision: *August 31, 2021*

Abstract: Species interactions and the networks that emerge from them structure ecosystem processes and enable biodiversity to persist through time. Still a robust understanding of interactions between species, how human activity is effecting these intearctions, and how these change will effect Earth's ecosystems in the future remains elusive. This knowledge-gap is largely driven by a shortfall of data—although species occurence data has rapidly increased in the last decade, species interaction data has lagged behind, largely due to the intrinsic difficulty of sampling interations. These sampling challenges bias data. Here, we demonstrate that the realized false-negative rate can be quite highly biased toward species with high relative abundance. We then simulate observation on both 243 empirical food webs and generated models to estimate the sampling effort required to reduce the false-negative rate to less than 10%. We then assess how false negatives effect measurements of network properties and models of network prediction. We conclude by discussing how understanding of false-negatives can inform how we design sampling of species interactions and the networks they form.

Introduction

Understanding which and how species interact is both a fundamental question of community ecology, but also an increasing imperative to mitigate the consequences of human activity on biodiversity (Makiola et al. 2020; Jordano 2016a) and to predict potential spillover of zoonotic disease (Becker et al. 2021). Over the past decade biodiversity data has become increasingly available. Modern remote-sensing has enabled collection of data on spatial scales and resolutions previously unimaginable, and improved in-situ sensing (Stephenson 2020) and adoption of open data practices (Kenall, Harold, and Foote 2014) have substantially amount of data available to ecologists. Still widespread data about species *interactions* remains elusive. Often observing an interaction between two species requires human sampling, because although remote methods can detect co-occurrence, this itself is not necessarily indicative of interaction (Blanchet, Cazelles, and Gravel 2020). This constraint induces biases on species interaction data subject to the spatial and temporal scales that humans can feasibly sample.

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

Because we cannot feasibly observe all (or even most) interactions that occur in nature, our samples end up capturing only a small fraction of those interactions. This means we can be reasonably confident two species actually interact if we have a record of it, but not at all confident that two species *do not* interact if we have no record of those species observed together. In other words, we can't distinguish true-negatives (two species *never* interact) from *false-negatives* (two species interact in some capacity, but we have not observed it). This is then amplified as the interaction data we have is geographically toward the usual suspects (Poisot, Bergeron, et al. 2021), This noise in data has practical consequences for answering questions about species interactions (de Aguiar et al. 2019)—these false-negatives could go on to effect the inferences we make about network properties and relations among species, and our predictions about how species will interact in the future.

This is compounded by semantic confusion about the definition of "interaction." Here distinguish between: a species *occurring*, a species being *observed occurring*, two species being observed *co-occurring*, and two species being observed *interacting* (fig. 1). In this manuscript, we refer to species either as "interacting"—two species co-occur (and, at least sometimes, interact)—or "not-interacting" (two species that, regardless of whether they co-occur, neither exhibits any meaningful effect on the biomass of the other). In fig. 1 we see that, under our definition, observing two species co-occurring is a prerequisite for observing an interaction between two species. But species are not observed with equal probability but instead in proportion to their relative biomass—you are much more likely to observe a species of high relative abundance than one of very low relative abundance (Poisot, Stouffer, and Gravel 2015). This assumes that there are no associations in species co-occurrence due to an interaction (perhaps because this interaction is "important" for both species) (Cazelles et al. 2016), but here we show increasing strength of associations leads to increasing probability of false-negatives in interaction data. Further observed co-occurrence is often equated with meaningful interaction strength, but this is not necessarily the case (Blanchet, Cazelles, and Gravel 2020; Strydom et al. 2021). Bears and salmon *interact*—a bear and the microbes in the soil of a dens interact, but less so.

[Figure 1 about here.]

Here, we show that the probability of observing a actual "non-interaction" between species depends on sampling effort, and suggest that surveys of species interactions can benefit from simulation modeling of detection probability (Jordano 2016b). We demonstrate that the realized false-negative rate of interactions is directly related the relative abundance of a particular species, relationship between total sampling effort (the total count of all individuals of all species seen) and false-negative rate. questions we pose and attempt to answer are: 1) How many times do you have to observe a non-interaction between two species to be confident in saying that is a true negative? 2) How "wrong" are the measurements of network structure as a function of false-negative probability? and lastly 3) How do false-negatives impact our ability to make reliable predictions about interactions? We show that positive associations in co-occurrence data can increase realized probability of false negatives, and demonstrate these positive associations are present in two spatially-replicated systems. We conclude by suggesting that simulation of sampling effort and species occurrence can and should be used to help design surveys of species diversity (Moore and McCarthy 2016), and by advocating use of null models like those presented here as a tool for guiding design of surveys of species interactions, and for modeling detection error in predictive ecological models.

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

To answer the titular question of this section, we present a naive model of interaction detection: we assume that every interacting pair of species is incorrectly observed as a not-interacting with an independent and fixed probability, which we denote p_{fn} and subsequently refer to as the False-Negative Rate (FNR). If we observe the same species not-interacting N times, then the probability of a true-negative (denoted p_{tn}) is given by $p_{tn} = 1 - (p_{fn})^N$. This relation (a special case of the negative-binomial distribution) is shown in fig. 2 for varying values of the false negative rate p_{fn}. This illustrates a fundamental link between our ability to reliably say an interaction doesn't exist—p_{tn}—and the number of times we have observed a given species. In addition, note that there also is no non-zero p_{fn} for which we can ever *prove* that an interaction does not exist—no matter how many observations of non-interaction N we have, $p_{tn} < 1$.

[Figure 2 about here.]

From fig. 2 (and general intuition) it is clear that the more times we see two species *occurring*, but *not* interacting, the more likely the interaction is a true negative. But how does one decide what this threshold of number of observations should be when planning to sample a given system? If false-negative rates presented in fig. 2 seem unrealistically high, consider that species are not observed independent of their relative abundance. In the next section we demonstrate that distribution of abundance in ecosystems can lead to realized values of p_{fn} similar to those in fig. 2 for species with low relative abundance, simply as a artifact of sampling effort.

False-negatives as a product of relative abundance

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

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

In fig. 3 (a) we see this model of observation applied to networks generated using the niche model (Richard J. Williams and Martinez 2000) across varying levels of species richness, and in (b) applied to 243 food-webs from the Mangal database. For all niche model simulations in this manuscript, for a given number of species S the number of interactions is drawn from the flexible-links model fit to Mangal data (MacDonald, Banville, and Poisot 2020), effectively drawing the number of interactions L for a random niche model food-web as $L \sim \text{BetaBinomial}(S^2 - S + 1, \mu\phi, (1-\mu)\phi)$, where the MAP estimate of (μ, ϕ) applied to Mangal data from MacDonald, Banville, and Poisot (2020) is $(\mu = 0.086, \phi = 24.3)$. All simulations were done with 500 independent replicates per unique number of observations n. All analyses presented here are done in Julia v1.6 (Bezanson et al. 2015) using both EcologicalNetworks.jl v0.5 and Mangal.jl v0.4 [Banville, Vissault, and Poisot (2021); ZENODO link TODO]. Note that the empirical data also is, due to the phenomena described here, very likely to *already* have many false negatives, which is why we are interested in prediction of networks in the first place—we'll revisit this in the final section.

[Figure 3 about here.]

In panel (c) of fig. 3, we show the expected number of total observations needed to obtain a "goal" number of observations (colors) of a particular "focal" species. As an example, if we hypothesize that A and B do not interact, and we want to see species A and B both co-occurring and not-interacting 10 times to be confident this is a negative (a la fig. 2), then we need an expected 10,000 observations of all species if the relative abundance of A is 0.00125.

Empirical data on interactions are subject to the practical limitations of funding and humanwork hours, and therefore existing data tend to fall on the order on 100s or 1000s observations of individuals per site (Resasco, Chacoff, and Vázquez 2021; Schwarz et al. 2020; Nielsen and Bascompte 2007). Clear aggregation of this data has proven difficult to find and a meta-analysis of network data and sampling effort seems both pertinent and necessary, in addition to the effects of aggregation of interactions across taxonomic scales (Giacomuzzo and Jordán 2021; Gauzens et al. 2013). Further, from fig. 3 it is evident that the number of species considered in a study is inseparable from the false-negative rate in that study, and this effect should be taken into account when designing samples of ecological networks in the future.

We conclude this section by advocating for the use of neutral models similar to above to generate expectations about the number of false-negatives in a data set of a given size. This could prove fruitful both for designing surveys of interactions (Canard et al. 2012), but also because we may want to incorporate models of observation error into predictive models (Joseph 2020). Additionally, one must consider the context for sampling—is the goal to detect a particular species A (as in fig. 3 (c)), or to get a representative sample of interactions across the species pool? This argument is well-considered when 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. 2 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, Stouffer, and Gravel 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 (R. M. Thompson and 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 4 about here.]

In fig. 4, 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 taxonmic identifiers, often at different resolutions. These particular datasets (Hadfield et al. 2014; R. M. Thompson and Townsend 2000) were usable because they already have been sorted to have a fixed taxonomic backbone (as part of EcologicalNetworks.jl (Banville, Vissault, and Poisot 2021)). Applying this in bulk to Mangal food-webs presents the difficulty of resolving different taxon identifiers across spatial samples of species with to different resolutions, which is why we can't simply apply this to the whole Mangal database—this highlights a general problem of resolving taxonomic indentifiers which use different names and different resolutions in different ecological datasets, which is a problem that needs to be addressed for computational approaches to scale up to the world of big-ecological-data we hope to build, although this is a task that may be aided via natural-language-processing methods.

The impact of false-negatives on network analysis and prediction

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

Effects of false-negatives on network properties

Here we simulate the process of observation with error to generate synthetic data with a known proportion of false negatives, and compare the computed network properties of the original "true" network to the computed properties of the "observed" network with added false-negatives. In fig. 5 we see the mean-squared error of connectance, mean degree-centrality, and spectral radius, computed across 2000, 2000, and 300 replicates respectively at each value of the false negative rate p_{fn}. All replicates use random food-webs simulated using the niche model (Richard J. Williams and Martinez 2000) with 100 species and connectance drawn from the flexible-links model (MacDonald, Banville, and Poisot 2020) as before.

[Figure 5 about here.]

We consider three properties: connectance, mean-degree-centrality, and spectral radius, indicative of local, meso, and global structure. Connectance is effectively a node-level property, a proxy for the degree distribution. Degree-centrality captures a different aspect of network structure than connectance, more 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. 5 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 (R.

J. 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 (Poisot, Ouellet, et al. 2021; Stock et al. 2017). However, it is feasible this could induce too much noise for a interaction prediction model to detect the signal of interaction chance from to 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-occurence (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. 6, 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 6 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 dazta 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 models performance, even in the $p_{fn} = 0$ case.

We conclude be 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, Barberán, and Clauset 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 (R. J. 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 (P. L. Thompson and 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 samples 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. 2, 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 (Young, Valdovinos, and Newman 2021; Stock et al. 2017), 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 (Young, Valdovinos, and Newman 2021; Johnson and Larremore 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

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.

Banville, Francis, Steve Vissault, and Timothée Poisot. 2021. “Mangal.jl and EcologicalNetworks.jl: Two Complementary Packages for Analyzing Ecological Networks in Julia.” *Journal of Open Source Software* 6 (61): 2721. https://doi.org/10.21105/joss.02721.

Becker, Daniel J., Gregory F. Albery, Anna R. Sjodin, Timothée Poisot, Laura M. Bergner, Tad A. Dallas, Evan A. Eskew, et al. 2021. “Optimizing Predictive Models to Prioritize Viral Discovery in Zoonotic Reservoirs.” *bioRxiv*, August, 2020.05.22.111344. https://doi.org/10.1101/2020.05.22.111344.

Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B. Shah. 2015. “Julia: A Fresh Approach to Numerical Computing.” *arXiv:1411.1607 [cs]*, July. http://arxiv.org/abs/1411.1607.

Blanchet, F. Guillaume, Kevin Cazelles, and Dominique Gravel. 2020. “Co-Occurrence Is Not Evidence of Ecological Interactions.” *Ecology Letters* 23 (7): 1050–63. https://doi.org/10.1111/ele.13525.

Boakes, Elizabeth H., Tracy M. Rout, and Ben Collen. 2015. “Inferring Species Extinction: The Use of Sighting Records.” *Methods in Ecology and Evolution* 6 (6): 678–87. https://doi.org/10.1111/2041-210X.12365.

Canard, Elsa, Nicolas Mouquet, Lucile Marescot, Kevin J. Gaston, Dominique Gravel, and David Mouillot. 2012. “Emergence of Structural Patterns in Neutral Trophic Networks.” *PLOS ONE* 7 (8): e38295. https://doi.org/10.1371/journal.pone.0038295.

Carlson, Colin J., Tad A. Dallas, Laura W. Alexander, Alexandra L. Phelan, and Anna J. Phillips. 2020. “What Would It Take to Describe the Global Diversity of Parasites?” *Proceedings of the Royal Society B: Biological Sciences* 287 (1939): 20201841. https://doi.org/10.1098/rspb.2020.1841.

Cazelles, Kévin, Miguel B. Araújo, Nicolas Mouquet, and Dominique Gravel. 2016. “A Theory for Species Co-Occurrence in Interaction Networks.” *Theoretical Ecology* 9 (1): 39–48.

https://doi.org/10.1007/s12080-015-0281-9.

Connor, Nora, Albert Barberán, and Aaron Clauset. 2017. "Using Null Models to Infer Microbial Co-Occurrence Networks." *PLOS ONE* 12 (5): e0176751. https://doi.org/10.1371/journal.pone.0176751.

Gauzens, Benoît, Stéphane Legendre, Xavier Lazzaro, and Gérard Lacroix. 2013. "Food-Web Aggregation, Methodological and Functional Issues." *Oikos* 122 (11): 1606–15.

Giacomuzzo, Emanuele, and Ferenc Jordán. 2021. "Food Web Aggregation: Effects on Key Positions." Preprint. Ecology. https://doi.org/10.1101/2021.04.18.440319.

Griffiths, David. 1998. "Sampling Effort, Regression Method, and the Shape and Slope of Sizeabundance Relations." *Journal of Animal Ecology* 67 (5): 795–804. https://doi.org/10.1046/j.1365-2656.1998.00244.x.

Hadfield, Jarrod D., Boris R. Krasnov, Robert Poulin, and Shinichi Nakagawa. 2014. "A Tale of Two Phylogenies: Comparative Analyses of Ecological Interactions." *The American Naturalist* 183 (2): 174–87. https://doi.org/10.1086/674445.

Johnson, Erik K., and Daniel B. Larremore. 2021. "Bayesian Estimation of Population Size and Overlap from Random Subsamples." *bioRxiv*, July, 2021.07.06.451319. https://doi.org/10.1101/2021.07.06.451319.

Jordano, Pedro. 2016a. "Chasing Ecological Interactions." *PLOS Biology* 14 (9): e1002559. https://doi.org/10.1371/journal.pbio.1002559.

———. 2016b. "Sampling Networks of Ecological Interactions." *Functional Ecology*, September. https://doi.org/10.1111/1365-2435.12763.

Joseph, Maxwell B. 2020. "Neural Hierarchical Models of Ecological Populations." *Ecology Letters* 23 (4): 734–47. https://doi.org/10.1111/ele.13462.

Kenall, Amye, Simon Harold, and Christopher Foote. 2014. "An Open Future for Ecological and Evolutionary Data?" *BMC Evolutionary Biology* 14 (1): 66. https://doi.org/10.1186/1471-2148-14-66.

MacDonald, Arthur Andrew Meahan, Francis Banville, and Timothée Poisot. 2020. "Revisiting the Links-Species Scaling Relationship in Food Webs." *Patterns* 1 (7). https://doi.org/10.1016/j.patter.2020.100079.

Makiola, Andreas, Zacchaeus G. Compson, Donald J. Baird, Matthew A. Barnes, Sam P. Boerlijst, Agnès Bouchez, Georgina Brennan, et al. 2020. "Key Questions for Next-Generation Biomonitoring." *Frontiers in Environmental Science* 7. https://doi.org/10.3389/fenvs.2019.00197.

Martinez, Neo D., Bradford A. Hawkins, Hassan Ali Dawah, and Brian P. Feifarek. 1999. "Effects of Sampling Effort on Characterization of Food-Web Structure." *Ecology* 80 (3): 1044–55. https://doi.org/10.1890/0012-9658(1999)080%5B1044:EOSEOC%5D2.0.CO;2.

Moore, Alana L., and Michael A. McCarthy. 2016. "Optimizing Ecological Survey Effort over Space and Time." *Methods in Ecology and Evolution* 7 (8): 891–99. https://doi.org/10.1111/2041-210X.12564.

Nielsen, Anders, and Jordi Bascompte. 2007. "Ecological Networks, Nestedness and Sampling Effort." *Journal of Ecology* 95 (5): 1134–41. https://doi.org/10.1111/j.1365-2745.2007.01271.x.

Paine, R. T. 1988. "Road Maps of Interactions or Grist for Theoretical Development?" *Ecology* 69 (6): 1648–54. https://doi.org/10.2307/1941141.

Poisot, Timothée, Gabriel Bergeron, Kevin Cazelles, Tad Dallas, Dominique Gravel, Andrew MacDonald, Benjamin Mercier, Clément Violet, and Steve Vissault. 2021. "Global Knowledge Gaps in Species Interaction Networks Data." *Journal of Biogeography*, April, jbi.14127. https://doi.org/10.1111/jbi.14127.

Poisot, Timothée, Marie-Andrée Ouellet, Nardus Mollentze, Maxwell J. Farrell, Daniel J. Becker, Gregory F. Albery, Rory J. Gibb, Stephanie N. Seifert, and Colin J. Carlson. 2021. "Imputing the Mammalian Virome with Linear Filtering and Singular Value Decomposition." *arXiv:2105.14973 [q-Bio]*, May. http://arxiv.org/abs/2105.14973.

Poisot, Timothée, Daniel B. Stouffer, and Dominique Gravel. 2015. "Beyond Species: Why Ecological Interaction Networks Vary Through Space and Time." *Oikos* 124 (3): 243–51. https://doi.org/10.1111/oik.01719.

Resasco, Julian, Natacha P. Chacoff, and Diego P. Vázquez. 2021. "Plantpollinator Interactions Between Generalists Persist over Time and Space." *Ecology* 102 (6): e03359. https://doi.org/10.1002/ecy.3359.

Savage, Van M., James F. Gillooly, James H. Brown, Geoffrey B. West, and Eric L. Charnov. 2004. "Effects of Body Size and Temperature on Population Growth." *The American Naturalist* 163 (3): 429–41. https://doi.org/10.1086/381872.

Schwarz, Benjamin, Diego P. Vázquez, Paul J. CaraDonna, Tiffany M. Knight, Gita Benadi, Carsten F. Dormann, Benoit Gauzens, et al. 2020. "Temporal Scale-Dependence of Plant-pollinator Networks." *Oikos* 129 (9): 1289–1302. https://doi.org/10.1111/oik.07303.

Stephenson, PJ. 2020. "Technological Advances in Biodiversity Monitoring: Applicability, Opportunities and Challenges." *Current Opinion in Environmental Sustainability* 45 (August): 36–41. https://doi.org/10.1016/j.cosust.2020.08.005.

Stock, Michiel, Timothée Poisot, Willem Waegeman, and Bernard De Baets. 2017. "Linear Filtering Reveals False Negatives in Species Interaction Data." *Scientific Reports* 7 (1): 45908. https://doi.org/10.1038/srep45908.

Strydom, Tanya, Michael David Catchen, Francis Banville, Dominique Caron, Gabriel Dansereau, Philippe Desjardins-Proulx, Norma Forero, et al. 2021. "A Roadmap Toward Predicting Species Interaction Networks (Across Space and Time)." Preprint. EcoEvoRxiv. https://doi.org/10.32942/osf.io/eu7k3.

Thompson, Patrick L., and Andrew Gonzalez. 2017. "Dispersal Governs the Reorganization of Ecological Networks Under Environmental Change." *Nature Ecology & Evolution* 1 (6): 1–8. https://doi.org/10.1038/s41559-017-0162.

Thompson, Ross M., and Colin R. Townsend. 2000. "Is Resolution the Solution?: The Effect of Taxonomic Resolution on the Calculated Properties of Three Stream Food Webs." *Freshwater Biology* 44 (3): 413–22. https://doi.org/10.1046/j.1365-2427.2000.00579.x.

Volkov, Igor, Jayanth R. Banavar, Stephen P. Hubbell, and Amos Maritan. 2003. "Neutral Theory and Relative Species Abundance in Ecology." *Nature* 424 (6952): 1035–37. https://doi.org/10.1038/nature01883.

Walther, B. A., P. Cotgreave, R. D. Price, R. D. Gregory, and D. H. Clayton. 1995. “Sampling Effort and Parasite Species Richness.” *Parasitology Today* 11 (8): 306–10. https://doi.org/10.1016/0169-4758(95)80047-6.

Williams, R. J., E. L. Berlow, J. A. Dunne, A.-L. Barabasi, and N. D. Martinez. 2002. “Two Degrees of Separation in Complex Food Webs.” *Proceedings of the National Academy of Sciences* 99 (20): 12913–16. https://doi.org/10.1073/pnas.192448799.

Williams, Richard J., and Neo D. Martinez. 2000. “Simple Rules Yield Complex Food Webs.” *Nature* 404 (6774): 180–83. https://doi.org/10.1038/35004572.

Willott, S. J. 2001. “Species Accumulation Curves and the Measure of Sampling Effort.” *Journal of Applied Ecology* 38 (2): 484–86. https://doi.org/10.1046/j.1365-2664.2001.00589.x.

Young, Jean-Gabriel, Fernanda S. Valdovinos, and M. E. J. Newman. 2021. “Reconstruction of Plantpollinator Networks from Observational Data.” *Nature Communications* 12 (1): 3911. https://doi.org/10.1038/s41467-021-24149-x.

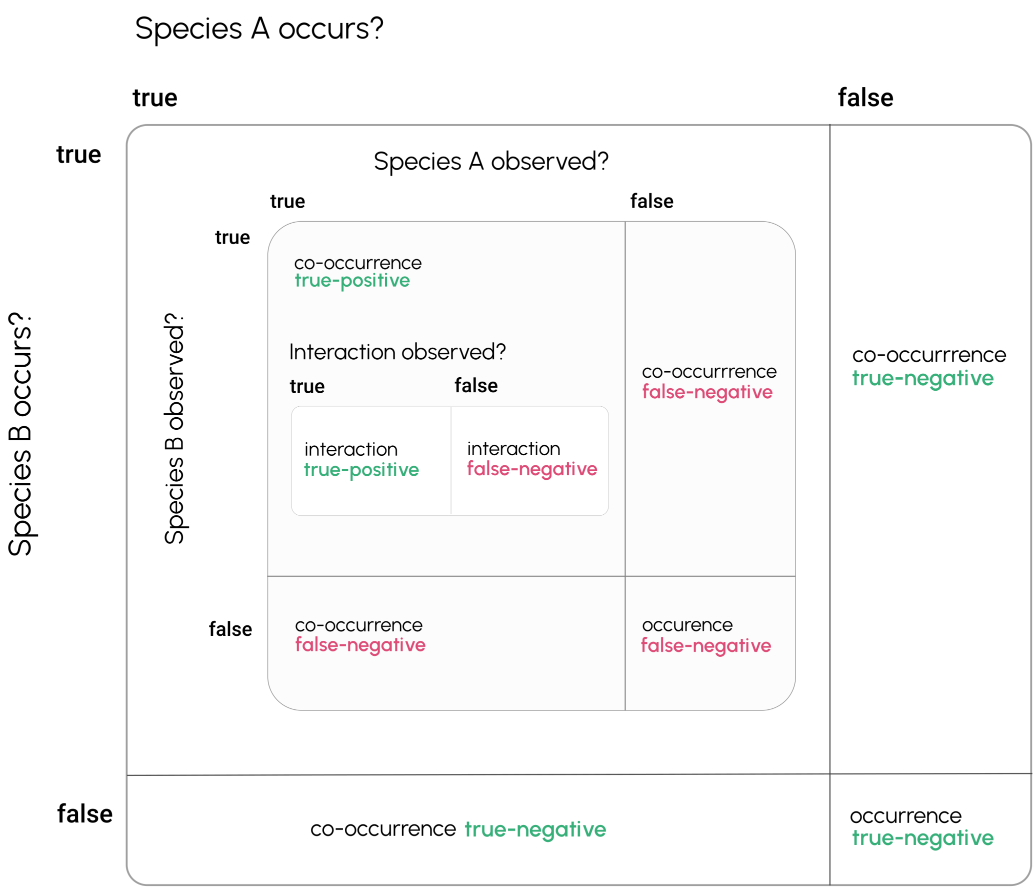

Figure 1: Taxonomy of false-negatives in data for two hypothetical species A and B, where in reality A and B do interact in some capacity.

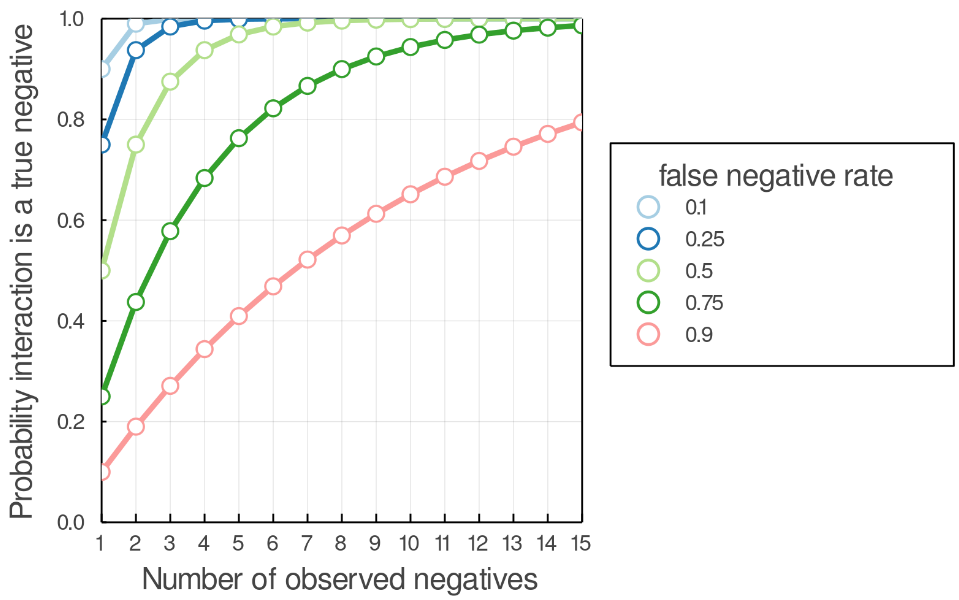

Figure 2: The probability an observed interaction is a "true negative" (y-axis) given how many times it has been sampled as a non-interaction (x-axis). Each color reflects a different value of p_{fn}, the false-negative rate (FNR). This is effectively the cdf of the negative-binomial distribution with $r = 1$. It's the birthday paradox, but backwards.

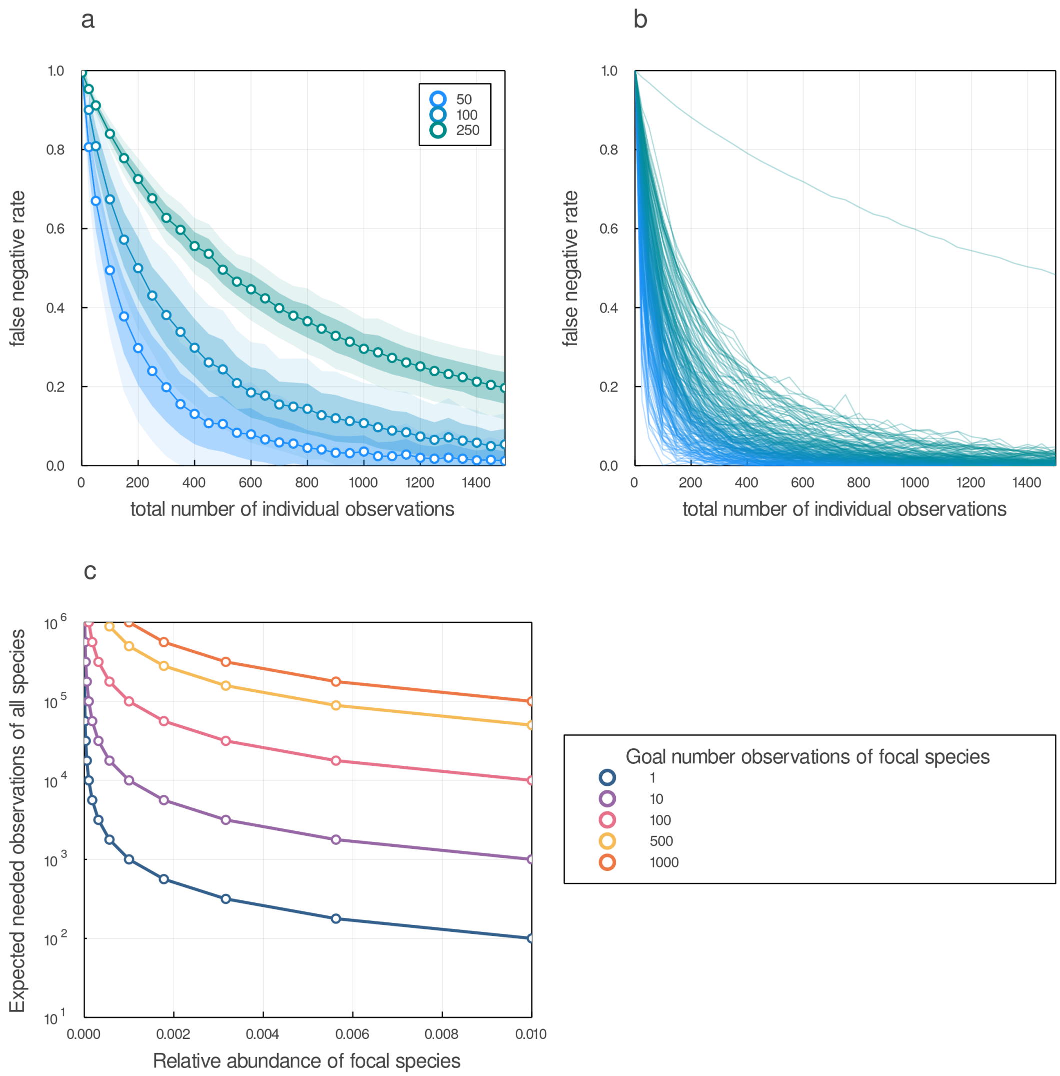

Figure 3: A and B: False negative rate (y-axis) as a function of total sampling effort (x-axis) and network size, computed using the method described above. For 500 independent draws from the niche model (Richard J. Williams and Martinez 2000) at varying levels of species richness (colors) with connectance drawn according to the flexible-links model (MacDonald, Banville, and Poisot 2020) 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 1σ in the first shade and 2σ 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. C) 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)

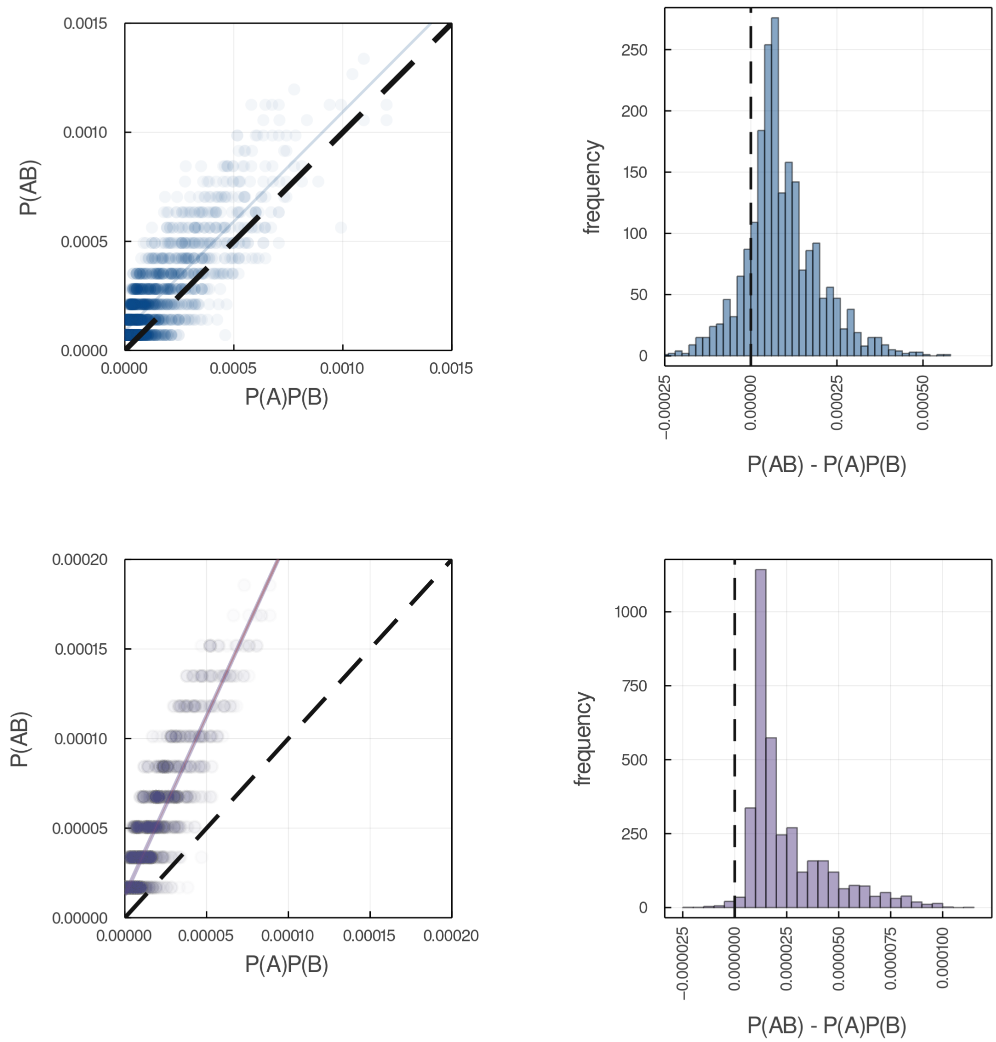

Figure 4: Top: Hadfield, Bottom: NZ Stream Foodwebs. Effectively a version of Cazelles et al. (2016) figure 1 panel A. Both distributions have $\mu \neq 0$ with $p < 10^{-50}$

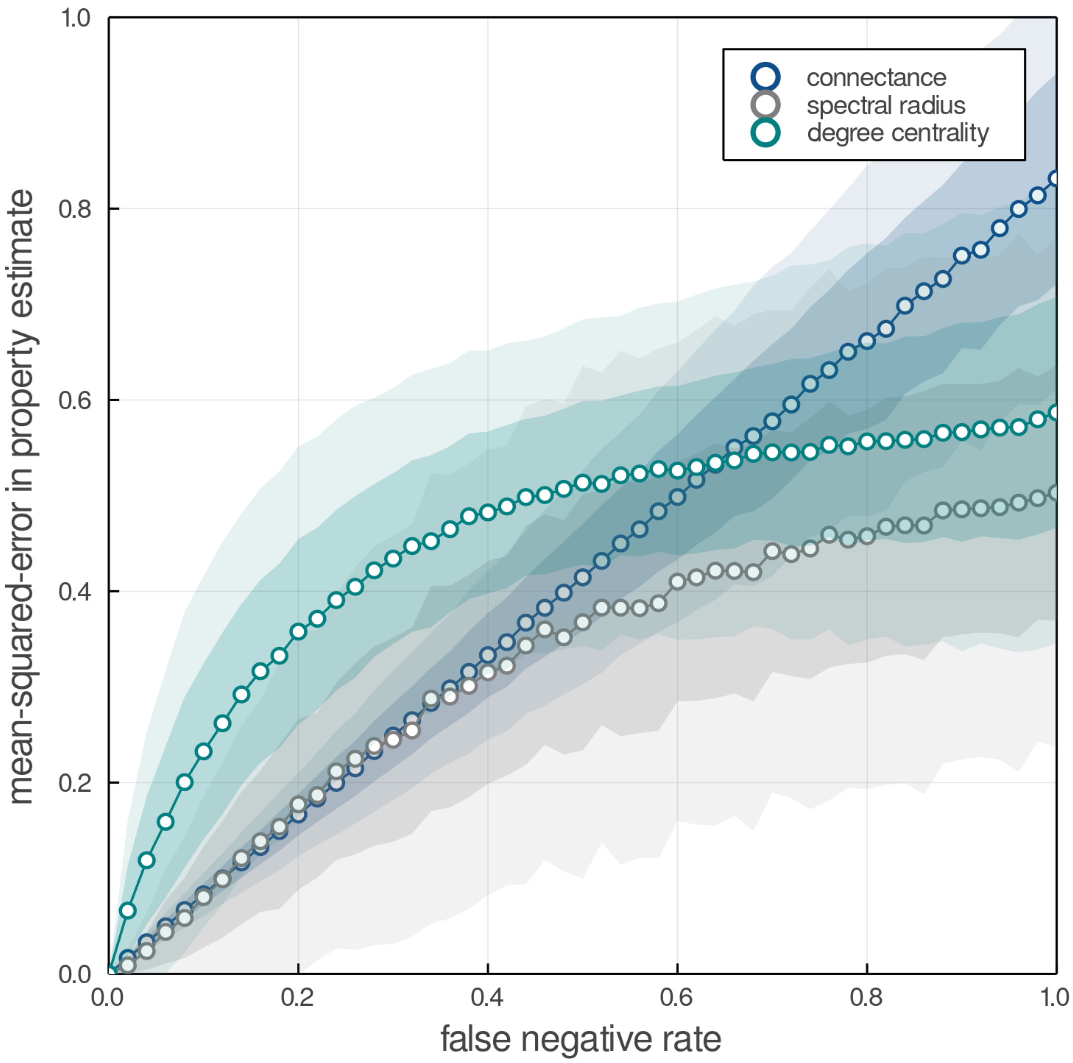

Figure 5: The mean-squared error (y-axis) of various network properties (different colors) across various simulated false-negative rates (x-axis). Means denoted with points, with 1σ in the first shade and 2σ in the second.

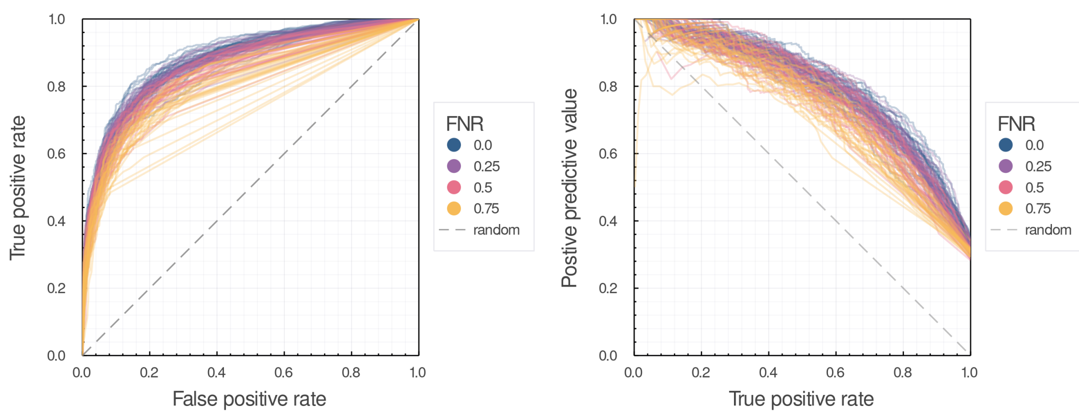

Figure 6: Receiver-operating-characteristic (left) and precision-recall (right) curves for the model on varying levels of false-negatives in the data (colors). For each value of FNR, we run 30 random training/test splits on 80/20 percent of the data. Replica of figure 1 in Strydom et al. (2021)