

METHOD

Novel analytic methods for predicting extinctions in ecological networks

Chris Jones^{1,2}  | Damaris Zurell³  | Karoline Wiesner² 

¹School of Mathematics, University of Bristol, Bristol, UK

²Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany

³Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

Correspondence

Karoline Wiesner

Email: karoline.wiesner@uni-potsdam.de

Funding information

Engineering and Physical Sciences Research Council, Grant/Award Number: EP/R513179/1; Research Focus Data-centric Sciences of the University of Potsdam

Handling Editor: Chuliang Song

Abstract

Ecological networks describe the interactions between different species, informing us how they rely on one another for food, pollination, and survival. If a species in an ecosystem is under threat of extinction, it can affect other species in the system and possibly result in their secondary extinction as well. Understanding how (primary) extinctions cause secondary extinctions on ecological networks has been considered previously using computational methods. However, these methods do not provide an explanation for the properties that make ecological networks robust, and they can be computationally expensive. We develop a new analytic model for predicting secondary extinctions that requires no stochastic simulation. Our model can predict secondary extinctions when primary extinctions occur at random or due to some targeting based on the number of links per species or risk of extinction, and can be applied to an ecological network of any number of layers. Using our model, we consider how false negatives and positives in network data affect predictions for network robustness. We have also extended the model to predict scenarios in which secondary extinctions occur once species lose a certain percentage of interaction strength, and to model the loss of interactions as opposed to just species extinction. From our model, it is possible to derive new analytic results such as how ecological networks are most robust when secondary species are of equal degree. Additionally, we show that both specialization and generalization in the distribution of interaction strength can be advantageous for network robustness, depending upon the extinction scenario being considered.

KEYWORDS

analytic prediction of extinctions, combinatorics, ecological extinctions, errors in interaction data, networks, robustness of mutualistic networks, species specialization

INTRODUCTION

No species exists in isolation, depending upon interactions with other species to feed, reproduce or maintain a stable

population (Cohen, 1978; Haeckel, 1866). Modeling the interactions between species is therefore of great importance in ecology, and one approach to this problem is to model interactions as an ecological network

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Ecological Monographs* published by Wiley Periodicals LLC on behalf of The Ecological Society of America.

(Bersier, 2007; Ings & Hawes, 2018). Ecosystems are increasingly threatened by the effects of climate change (Bellard et al., 2012; Dawson et al., 2011), which can cause sudden and widespread extinction events. Therefore, it is useful to model extinctions on ecological networks to understand the possible knock-on effects of species extinctions, as this may help to identify methods for conserving or reinforcing ecosystems in the future (Forup et al., 2008; Jongman, 1995; Tylianakis et al., 2010).

Species extinctions on ecological networks have been extensively studied in the past 20 years, with simplistic topological models providing predictions for the impact of extinctions under scenarios including extinctions that occur at random or with some ordering (Memmott et al., 2004), extinctions on networks made up of numerous trophic levels (Pocock et al., 2012), and extinctions that occur due to a loss of interaction strength over a certain threshold (Schleuning et al., 2016).

The models used are not the only possible approach to understanding the robustness of ecological networks. Other models consider the size of the largest component in the interaction network (Montoya et al., 2006; Solé & Montoya, 2001), or take a more dynamical approach as is the case with Bayesian network models (Aguilera et al., 2011; Ramazi et al., 2021). Here we restrict ourselves to what we refer to as simplistic topological network models, which originate from (Memmott et al., 2004). In these models, we are concerned with the point at which a given species goes extinct due to losing either a certain number of neighbors or a certain amount of interaction strength.

Previous work on simplistic topological network models has been largely computational, where extinctions are simulated to obtain predictions. Limited analytic work has been done to predict the robustness of ecological networks, which are either maximally or minimally nested (Burgos et al., 2007), but there is no existing analytic framework that can predict the robustness of any given simple ecological network. In the following, we develop such a model, which improves upon computational methods by providing an insight into the properties that make ecological networks robust, and by cutting computational costs.

We start by considering the same scenario put forward by Memmott et al. (2004), where a bipartite mutualistic network (such as a plant–pollinator network) undergoes extinctions on one trophic level, with species on the other trophic level experiencing secondary extinctions if they lose some or all of their neighbors. We show that, for random primary extinctions and a set quantity of interactions, ecological networks are maximally robust when secondary species have equally distributed

interactions. Secondary extinctions may also be predicted for targeted primary extinctions, and secondary extinctions are predictable on networks with more than two trophic layers (Pocock et al., 2012). Additionally, our model may also be used to predict the effects of errors in network data, where interactions are erroneously included or excluded.

The model is then developed further, taking into account the variable interaction strengths of neighboring species, where a species will go extinct if it loses a certain amount of interaction strength, as considered by Schleuning et al. (2016). For these scenarios, we show that if secondary species' interaction strength is maximally specialized then network robustness is constant regardless of network degree distribution or extinction sensitivity. If interaction strength is maximally generalized, networks with a high degree of secondary species have a robustness that is solely dependent upon extinction sensitivity. As a result, high specialization makes a network more robust if it is highly sensitive to interaction loss, and high generalization is better for robustness if interaction loss sensitivity is low. We also consider the scenario in which species go extinct gradually, modeled by the loss of interaction strength as opposed to entire species.

The wide range of applications and novel results presented here demonstrate the potential benefits of approaching the problem of predicting species extinctions on ecological networks using a combinatoric framework. As we argue in the following, this is an approach that can cut down on computational expense for predicting extinctions and provide new insights into the structures that make an ecological network robust, with several possible avenues for future development.

INTRODUCING THE ANALYTIC FRAMEWORK

In the model of Memmott et al. (2004), species in one trophic level (e.g., pollinators) undergo extinctions, and this impacts species in an adjacent, secondary trophic level (e.g., plants). If species in the secondary level lose all of their neighbors, they suffer a secondary extinction. Primary extinctions may occur at random or according to some ordering, such as highest to lowest degree, where the degree of a species is the number of interactions/links/edges it has. We can plot the proportion of secondary species that survive against the proportion of primary extinctions to visualize how robust a given ecological network is against extinction, and an example of such a “robustness curve” is shown in Figure 1.

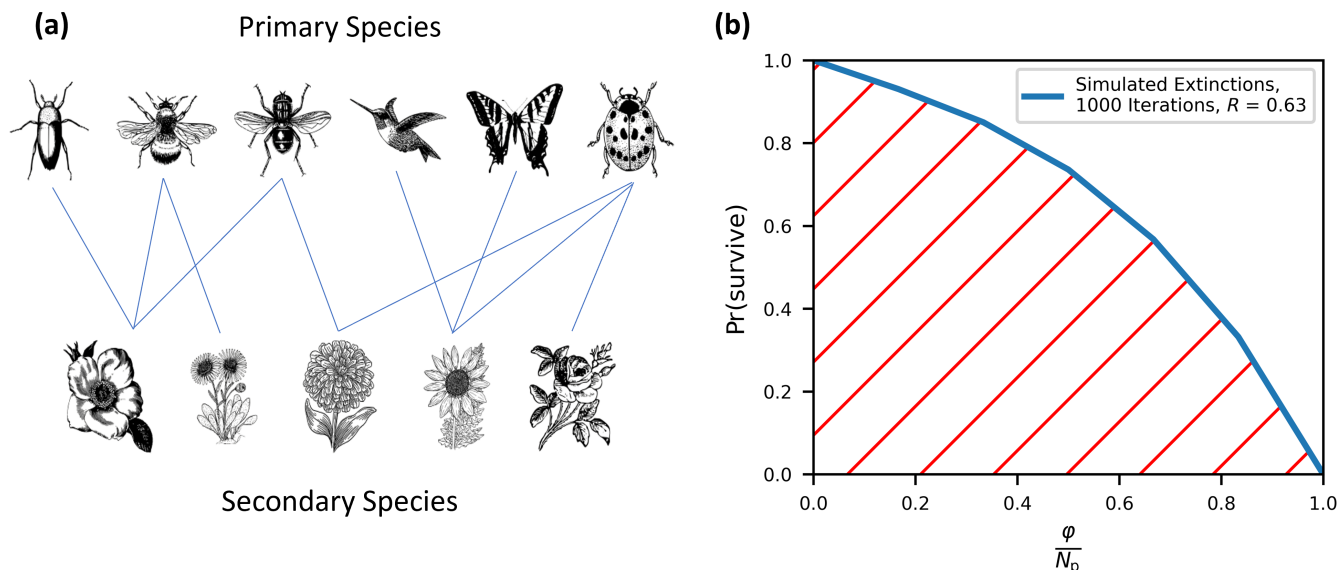


FIGURE 1 (a) An example plant–pollinator network and (b) its associated robustness curve. For this network, pollinators are treated as primary species and plants as secondary species. Random primary extinctions are simulated repeatedly and the proportion of surviving secondary species is recorded to generate the robustness curve. Illustrations used in (a) are all available in the public domain under a CC0 license.

The area under the robustness curve in Figure 1b may be calculated to give a single metric for ecological network robustness, and this is given by (Burgos et al., 2007):

$$R = \frac{1}{N_p + 1} \sum_{\varphi=0}^{N_p} \text{Pr}(\text{survive}|\varphi), \quad (1)$$

where N_p is the total number of primary species, φ is the number of primary species that have become extinct at a given point, and $\text{Pr}(\text{survive}|\varphi)$ is the average probability of a randomly chosen secondary species surviving after some φ primary species have been removed. Note that we divide the sum of survival probabilities by $N_p + 1$ because we are averaging $\text{Pr}(\text{survive}|\varphi)$ over φ values from 0 to N_p , and so there are $N_p + 1$ different values of φ in total. Previously, calculations of the robustness curve and the robustness value R have been done computationally, with some analytic results being derived for extreme cases (Burgos et al., 2007). As we show in the following, it is, in fact, possible to analytically predict the robustness curve of any given simple ecological network for a variety of extinction scenarios.

The following model relies heavily upon binomial coefficients, which express the number of combinations of k elements that may be chosen from some population of size n . The binomial coefficient $\binom{n}{k}$ may be calculated as:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad (2)$$

under the condition that $0 \leq k \leq n$. Note that $!$ refers to the factorial function, where for some x , $x! = x \times (x - 1) \times (x - 2) \times \dots \times 2 \times 1$.

Let us consider some species A in the secondary trophic level, which initially has degree k_A and therefore k_A unique neighbors in the primary level. If species in the primary level go extinct at random, we want to know the probability that species A has degree $k_A - j$ (i.e., j extinct neighbors) after some φ number of primary species have gone extinct. If there are N_p primary species, then there are $\binom{N_p}{\varphi}$ different possible combinations of primary species extinctions. We then need to find how many of those combinations include j neighbors of A . There are $\binom{k_A}{j}$ possible combinations for removing j neighbors of A , and therefore there are $\binom{N_p - k_A}{\varphi - j}$ possible combinations for removing $\varphi - j$ species that are not neighbors of A , so long as $\varphi \geq j$. Multiplying $\binom{k_A}{j}$ by $\binom{N_p - k_A}{\varphi - j}$ gives us the total number of combinations of length φ , which include j neighbors of A , and so we may write the probability of A having degree $k_A - j$ once φ primary species are extinct as:

$$\Pr(k'_A = k_A - j | \varphi) = \begin{cases} \frac{\binom{k_A}{j} \binom{N_p - k_A}{\varphi - j}}{\binom{N_p}{\varphi}} & \text{if } \varphi \geq j, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where k'_A refers to A's actual degree value once φ primary species have gone extinct. This is the hypergeometric distribution, which describes a process of sampling without replacement where each sample may pass (a neighbor of A is removed) or fail (a nonneighboring primary species is removed). If we specify that species A goes extinct once its degree is $k'_A = k_A - i_k$ or below (i.e., it has lost at least i_k neighbors), then the disconnection probability for secondary species A once some φ primary species are extinct is:

$$\Pr(\text{A extinct} | \varphi) = \sum_{j=i_k}^{k_A} \Pr(k'_A = k_A - j | \varphi), \quad (4)$$

Because extinction probability is only dependent upon the initial degree of a given secondary species, the total number of primary species in the network, and the number of primary species removed, we may extend this to all secondary species of initial degree k . Consequently, the average secondary extinction probability over the entire network is:

$$\Pr(\text{extinct} | \varphi) = \sum_{k=0} p(k) \sum_{j=i_k}^{k_A} \Pr(k' = k - j | \varphi), \quad (5)$$

where $p(k)$ is the probability of some randomly chosen secondary species having an initial degree of k . Given that $\Pr(\text{survive} | \varphi)$ is simply $1 - \Pr(\text{extinct} | \varphi)$, we can rewrite the expression for robustness R from Equation (1) as:

$$R = 1 - \frac{1}{N_p + 1} \sum_{\varphi=0}^{N_p} \Pr(\text{extinct} | \varphi). \quad (6)$$

Here we note that this analytic model is considerably computationally cheaper than brute force simulation. With an efficient implementation, calculating $\Pr(\text{extinct} | \varphi)$ analytically takes $O(p)$ time, where p is the number of unique nonzero entries in the secondary species degree distribution. By contrast, estimating $\Pr(\text{extinct} | \varphi)$ computationally once takes $O(N_s)$ time, where N_s is the number of secondary species and $N_s \geq p$. In practice, it is often necessary to run several thousand simulations to produce an accurate estimate of $\Pr(\text{extinct} | \varphi)$, and so our analytic approach is substantially computationally cheaper than the brute force method.

In Figure 2a, we demonstrate the results of our model by comparing the analytically predicted robustness curve for an ecological network against the average curve obtained computationally when all neighbors must be removed for extinction to occur (i.e., $i_k = k$). The ecological data used are from a study of

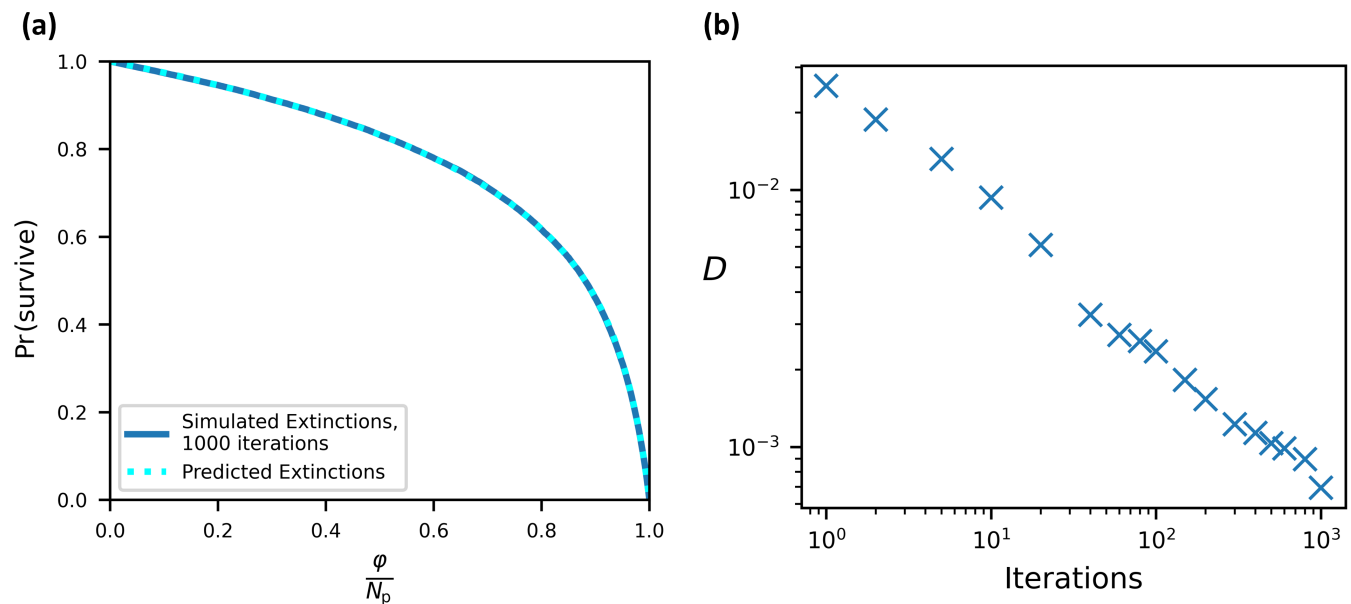


FIGURE 2 (a) Analytically predicted and computationally simulated robustness curves of a real-world network from a study by Kato (2000) and (b) the absolute curve divergence between the analytic and simulated curve as computational simulations increase, plotted on a log-log scale.

plant–pollinator networks in Japan by Kato (2000). We can see that the computationally obtained curve converges to our predicted curve as the number of simulations increases, indicating that our method accurately predicts the average robustness curve.

We also compare the absolute curve divergence between predicted and simulated curves for an increasing number of simulations in Figure 2b. The absolute curve divergence D is given by:

$$D = \frac{\sum_{\varphi=0}^{N_p} |\Pr(\text{survive}|\varphi)_{\text{predict}} - \Pr(\text{survive}|\varphi)_{\text{sim}}|}{N_p}, \quad (7)$$

where $\Pr(\text{survive}|\varphi)_{\text{predict}}$ and $\Pr(\text{survive}|\varphi)_{\text{sim}}$ are the predicted and average simulated secondary species survival probabilities, respectively. In Figure 2 we can see that the computationally generated result converges toward our prediction in the limit of a large number of simulations, indicating that our analytic model is accurate.

We implemented our model in Python 3.9, with our code archived in Jones (2023). Our code may be used to replicate our results throughout this paper, and can form a basis for implementing our model and its extensions in other contexts. The predictions presented throughout this paper rely heavily upon the hypergeometric distribution, which may be easily calculated in Python using the “`scipy.stats.hypergeom`” function. For example, to calculate $\Pr(\text{A extinct}|\varphi)$ from Equation (4), we can run:

```
probAExtinct = 1 - scipy.stats.hypergeom.cdf
                (threshold-1, noPrimary, degreeA,
                 removedSpecies)
```

where `threshold` corresponds to i_A , `noPrimary` corresponds to N_p , `degreeA` corresponds to k_A and `removedSpecies` corresponds to φ . We used `scipy.stats.hypergeom.cdf` to calculate the cumulative distribution of the hypergeometric distribution, which is the same as taking the sum of $\Pr(k'_A = k_A - j|\varphi)$ over values of j . An equivalent calculation may be made in R using the “`phyper`” function to give

```
probAExtinct = 1 - phyper(threshold-1,
                          removedSpecies, noPrimary -
                          removedSpecies, degreeA)
```

DEGREE EQUALITY MAXIMIZES ROBUSTNESS

Using the analytic model presented in the preceding section, we can prove that, when a secondary species must

lose all of its neighbors to go extinct, an ecological network of fixed average degree is most robust when all secondary species are as close together in degree values as possible. When the average degree $\langle k \rangle$ is an integer, this entails a degree distribution of:

$$p(k) = \begin{cases} 1 & \text{if } k = \langle k \rangle, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

and if $\langle k \rangle$ is not an integer, then the robustness maximizing distribution is:

$$p(k) = \begin{cases} 1 + \langle k \rangle - \lceil \langle k \rangle \rceil & \text{if } k = \lceil \langle k \rangle \rceil, \\ 1 - \langle k \rangle + \lfloor \langle k \rangle \rfloor & \text{if } k = \lfloor \langle k \rangle \rfloor, \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

where $\lceil \langle k \rangle \rceil$ and $\lfloor \langle k \rangle \rfloor$ are the average degree rounded up and rounded down, respectively, to the nearest integer.

To prove this, we may first show how robustness R is solely dependent upon the degree values of secondary species when secondary species only go extinct upon the loss of all of their neighbors. For a given secondary species A of degree k_A , its average extinction probability throughout the process of random primary species extinctions is given by:

$$\begin{aligned} \Pr(\text{A extinct}|\varphi) &= \frac{1}{N_p + 1} \sum_{\varphi=0}^{N_p} \begin{cases} \frac{\binom{k_A}{k_A} \binom{N_p - k_A}{\varphi - k_A}}{\binom{N_p}{\varphi}} & \text{if } \varphi \geq k_A, \\ 0 & \text{otherwise,} \end{cases} \\ &= \frac{1}{N_p + 1} \sum_{\varphi=k_A}^{N_p} \frac{\binom{N_p - k_A}{\varphi - k_A}}{\binom{N_p}{\varphi}}, \\ &= \frac{1}{N_p + 1} \sum_{\varphi=k_A}^{N_p} \frac{\varphi! (N_p - k_A)!}{N_p! (\varphi - k_A)!}. \end{aligned} \quad (10)$$

Multiplying Equation (10) by $\frac{k_A!}{k_A!}$ gives:

$$\begin{aligned} \Pr(\text{A extinct}|\varphi) &= \frac{k_A! (N_p - k_A)!}{N_p! (N_p + 1)} \sum_{\varphi=k_A}^{N_p} \frac{\varphi!}{k_A! (\varphi - k_A)!}, \\ &= \frac{1}{N_p + 1} \frac{1}{\binom{N_p}{k_A}} \sum_{\varphi=k_A}^{N_p} \binom{\varphi}{k_A}, \\ &= \frac{\binom{N_p + 1}{k_A + 1}}{(N_p + 1) \binom{N_p}{k_A}}, \\ &= \frac{1}{k_A + 1}. \end{aligned} \quad (11)$$

Scaling this result up across the entire network, we can calculate robustness R as:

$$R = 1 - \sum_{k=0} p(k) \frac{1}{k+1}, \quad (12)$$

$$= \sum_{k=0} p(k) \frac{k}{k+1}.$$

Therefore, the robustness of an ecological network is solely dependent upon the degree values of secondary species when secondary species only go extinct upon the loss of all of their neighbors.

To prove that degree equality among secondary species maximizes robustness, let us consider two secondary species A and B that we rewire by removing an edge from B and adding one to A. How would rewiring an edge like this affect the network's robustness? The change in robustness ΔR is given by:

$$\Delta R = \frac{1}{N_s} \left(\frac{k_A + 1}{k_A + 2} + \frac{k_B - 1}{k_B} - \frac{k_A}{k_A + 1} - \frac{k_B}{k_B + 1} \right),$$

$$= \frac{1}{N_s} \left(\frac{(k_A + 1)^2 - k_A(k_A + 2)}{(k_A + 1)(k_A + 2)} + \frac{(k_B - 1)(k_B + 1) - k_B^2}{k_B(k_B + 1)} \right),$$

$$= \frac{1}{N_s} \left(\frac{1}{(k_A + 1)(k_A + 2)} - \frac{1}{k_B(k_B + 1)} \right), \quad (13)$$

where N_s is the number of secondary species in the network. Therefore, $\Delta R > 0$ when $k_A + 1 < k_B$, $\Delta R = 0$ when $k_A + 1 = k_B$ and $\Delta R < 0$ when $k_A + 1 > k_B$. This entails the fact that robustness always increases or stays the same if species A is of a lower initial degree than species B and an edge is rewired from B to A. Applying this process repeatedly to a network until robustness cannot be improved further gives either the degree distribution described in Equation (8) for integer $\langle k \rangle$, or the distribution of Equation (9) when $\langle k \rangle$ is not an integer. In other words, the network is most robust when secondary species are as equal in degree values as possible.

This result tells us the structural properties that make secondary species in ecological networks robust against random primary species extinctions, namely equally distributed interactions among secondary species. Previous research has indicated this before (Burgos et al., 2007). However, this was not conclusively proven, and in the work of Burgos et al. (2007) robustness was related to nestedness as opposed to secondary species degree distribution. As previously noted, network nestedness is dependent upon how primary species interact with secondary species, but our results demonstrate the fact that robustness against random extinctions is determined by

the degree distribution of secondary species alone. Therefore, for random primary extinctions, nestedness is not necessarily an indicator of secondary species robustness.

To illustrate how robustness increases as secondary species become equal in degree when edges are rewired from high- to low-degree secondary species, we provide a series of example networks in Figure 3, each with equal numbers of interactions and primary and secondary species, but different robustness values.

We can see that the most "unequal" network (i) has the lowest robustness, and the network where all secondary species are of equal degree (iv) has the highest robustness. Additionally, networks (ii) and (iii) have the same secondary species degree distributions and robustness values as one another, as the only difference between them is the degree distribution of primary species. The difference in primary species degree distribution means that they are not considered to have the same nestedness as one another, but they are equally robust, demonstrating the fact that nestedness and robustness are not necessarily related in the context of random extinctions.

TARGETED SPECIES EXTINCTIONS

The model demonstrated in the preceding sections only predicts robustness when primary species are removed at random, but it is also possible to adjust the model to predict robustness when primary species are removed in descending or ascending degree order. This means that species with many links (descending order) or few links (ascending order) go extinct first. In these scenarios, primary species are effectively sorted into some n different groups based on degree value. All species within a group are removed in a random order before moving on to the next group, which is higher or lower in degree value, depending on the scenario. If we consider some secondary species A, it will have some $k_{A,l}$ neighbors in a given primary species group where all primary species have degree l . As before, we set some threshold number i_k of neighboring species that must be lost before a given secondary species of degree k goes extinct. We can then say that the secondary species will go extinct as we remove primary species from some group d if it satisfies the conditions $\sum_{l=d}^n k_{A,l} \geq i_k$ and $\sum_{l=d+1}^n k_{A,l} < i_k$ for descending degree order removal, and $\sum_{l=0}^d k_{A,l} \geq i_k$ and $\sum_{l=0}^{d-1} k_{A,l} < i_k$ for ascending degree order removal. We can therefore write the probability of some secondary species A going extinct when primary species are removed in descending degree order, given that species are being removed from group d and some ϕ_d species have been removed from group d , as:

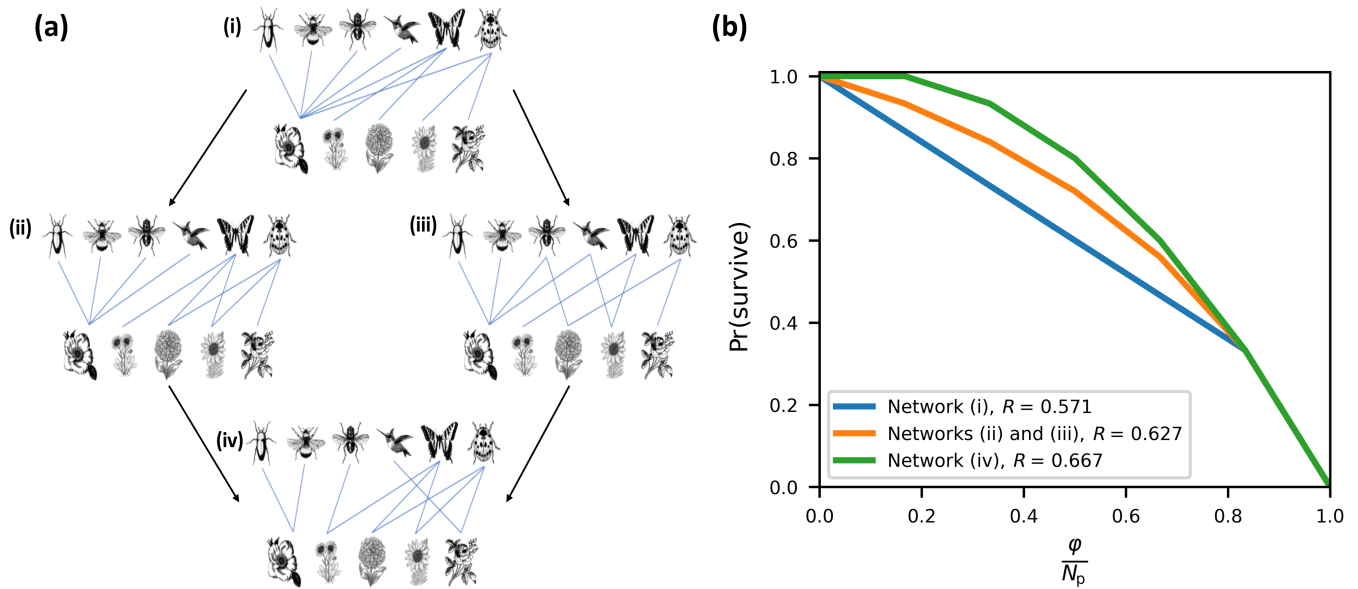


FIGURE 3 (a) Example networks with different secondary species degree distributions and (b) network robustness curves for the four example networks. Network (1) exhibits the lowest robustness, networks (2) and (3) have the same robustness even though they have different primary species degree distributions, and network (4) has the highest robustness, where all secondary species have the same degree. Illustrations used in (a) are all available in the public domain under a CC0 license.

$$\Pr(A \text{ extinct}, d | \varphi_d) = \begin{cases} 0 & \text{if } \sum_{l=d}^n k_{A,l} < i_k \\ & \text{or } \varphi_d < j_{A,d}, \\ \frac{\binom{k_{A,d}}{j_{A,d}} \binom{N_d - k_{A,d}}{\varphi_d - j_{A,d}}}{\binom{N_d}{\varphi_d}} & \text{if } \sum_{l=d}^n k_{A,l} \geq i_k \\ & \text{and } \sum_{l=d+1}^n k_{A,l} < i_k \\ & \text{and } \varphi_d \geq j_{A,d}, \\ 1 & \text{if } \sum_{l=d+1}^n k_{A,l} \geq i_k, \end{cases} \quad (14)$$

where N_d is the number of primary species in group d and $j_{A,d} = i_k - \sum_{l=d+1}^n k_{A,l}$. Similarly, for primary species removal in ascending degree order we have:

$$\Pr(A \text{ extinct}, d | \varphi_d) = \begin{cases} 0 & \text{if } \sum_{l=0}^d k_{A,l} < i_k \\ & \text{or } \varphi_d < j_{A,d}, \\ \frac{\binom{k_{A,d}}{j_{A,d}} \binom{N_d - k_{A,d}}{\varphi_d - j_{A,d}}}{\binom{N_d}{\varphi_d}} & \text{if } \sum_{l=0}^d k_{A,l} \geq i_k \\ & \text{and } \sum_{l=0}^{d-1} k_{A,l} < i_k \\ & \text{and } \varphi_d \geq j_{A,d}, \\ 1 & \text{if } \sum_{l=0}^{d-1} k_{A,l} \geq i_k, \end{cases} \quad (15)$$

where $j_{A,d} = i_k - \sum_{l=0}^{d-1} k_{A,l}$. The extinction probabilities $\Pr(A \text{ extinct} | d, \varphi_d)$ from Equations (14) and (15) can be averaged over all secondary species to give:

$$\Pr(\text{extinct} | \varphi) = \frac{1}{N_s} \sum_A \sum_{d=0}^n \Pr(A \text{ extinct} | d, \varphi_d), \quad (16)$$

which is the average extinction probability for any number of removed primary species. Therefore, we can predict the robustness curves for descending and ascending degree removal of primary species, and an example is given in Figure 4.

As before, our analytic model can successfully predict the average extinction probabilities for secondary species as primary species are removed. The scenarios in which primary species are removed in descending and ascending degree order have been referred to as the “worst” and “best” case scenarios, respectively. However, our analytic model suggests that from the perspective of robustness, this may not exactly be the case. Under descending degree order removal, a secondary species’ extinction probability only depends upon how many lowest-degree neighbors it has, and for ascending degree order removal it depends upon the number of highest-degree neighbors. In Figure 5, we provide an example ecosystem for which descending degree order removal gives higher network robustness than ascending degree removal.

While this is a specifically constructed example, it demonstrates that finding the true worst- or best-case scenario for secondary extinctions is not necessarily a case of removing primary species in descending or ascending degree order. As such, a possible future line of inquiry is to try to establish the true worst- and best-case scenarios for species extinction on any given network.

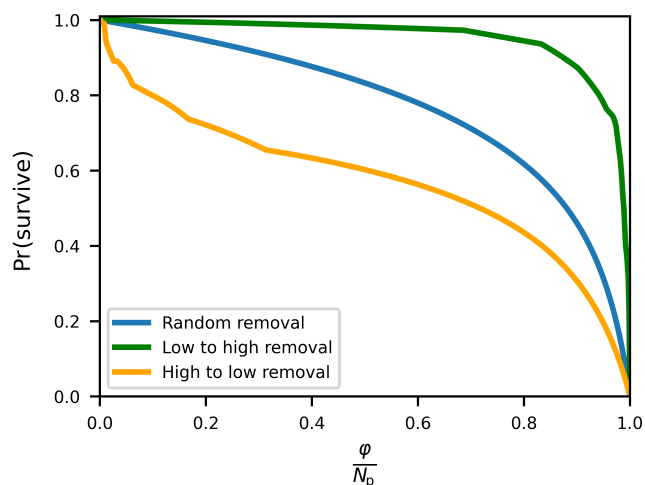


FIGURE 4 Analytically predicted robustness curves for targeted primary species removal. Predictions are made on the plant–pollinator network from a study by Kato (2000). The blue curve is for random removal, included as a point of comparison. The green curve is for removal of primary species in ascending degree order (lowest degree first) and the orange curve is for removal of primary species in descending degree order (highest degree first).

It should also be noted that according to our model, the probability of a secondary species going extinct is only dependent upon the number of primary species it interacts within one degree group. For descending extinctions, a secondary species' extinction is only dependent upon the number of primary species in its lowest-degree group of neighbors, and for ascending extinctions it solely depends on the number of neighbors in its highest-degree group. The number of distinct groups does not necessarily alter a network's robustness, although having more distinct groups of primary species results in more predictable extinctions, with the most extreme case occurring when each primary species is in a group by itself, making extinctions perfectly predictable.

Thus far in this section we have considered species extinctions that are targeted based on degree value, but this is only one possible extinction ordering. Recent research has examined extinction scenarios in which species are lost according to their extinction risk as assessed by the International Union for Conservation of Nature (IUCN) Red List (Lamperty & Brosi, 2022). Species are ranked in the Red List from Critically Endangered to Least Concern, and in work by Lamperty and Brosi (2022), frugivore species in a seed-dispersal network are removed from highest to lowest extinction risk. Because species only belong to one of these extinction risk categories, it is necessary to simulate extinctions from each risk category in descending risk order, with the order of extinctions within each group randomized. This framework fits well with our targeted species extinction model.

Using data from (Bello et al., 2017), we can replicate the results of (Lamperty & Brosi, 2022), predicting the survival of plant species in a seed-dispersal network as frugivore species are lost in descending extinction risk order. We calculate extinction probability using:

$$\Pr(A \text{ extinct}|\varphi) = \begin{cases} 0 & \text{if } \sum_{l=d}^n k_{A,l} < i_k \\ & \text{or } \varphi_d < j_{A,d}, \\ \frac{\binom{k_{A,d}}{j_{A,d}} \binom{N_d - k_{A,d}}{\varphi_d - j_{A,d}}}{\binom{N_d}{\varphi_d}} & \text{if } \sum_{l=d}^n k_{A,l} \geq i_k \\ & \text{and } \sum_{l=d+1}^n k_{A,l} < i_k \\ & \text{and } \varphi_d \geq j_{A,d}, \\ 1 & \text{if } \sum_{l=d+1}^n k_{A,l} \geq i_k, \end{cases} \quad (17)$$

which is the same as Equation (14), except instead of our primary species groups being organized by degree value, they are now organized by extinction risk. In Figure 6, we show predicted plant species survival probabilities as frugivore species are removed in descending extinction risk order, with plant species going extinct once they have lost all of their frugivore neighbors.

This demonstrates the fact that our framework for targeted species extinctions may be extended to any ordering of primary species loss where primary species are sorted into groups that become extinct in some order, but extinction within groups occurs at random.

MULTILAYER ECOSYSTEMS

So far, we have only demonstrated our model for bipartite systems, that is, those that include only two groups that interact with one another. However, real-world ecosystems can exist on several distinct layers, for example, predators may feed on pollinators, which in turn pollinate plants. Another extension for our model is to predict species extinction in a group of species not directly adjacent to the group undergoing primary extinction. This is predictable analytically, but only for the scenario in which a species must lose all of its neighbors to go extinct.

Scenarios such as this have previously been considered by (Pocock et al., 2012), using computational methods. Here we construct an example network to demonstrate robustness predictions on multilayer networks. Let us consider a system of plants, pollinators and predators, where predators feed on pollinators, who in turn feed on plants. We want to know the probability of a predator going extinct after a certain number of plant extinctions. For some predator species A, the species will go extinct if all of the pollinator species it is connected to become extinct, which only occurs once all of their plant

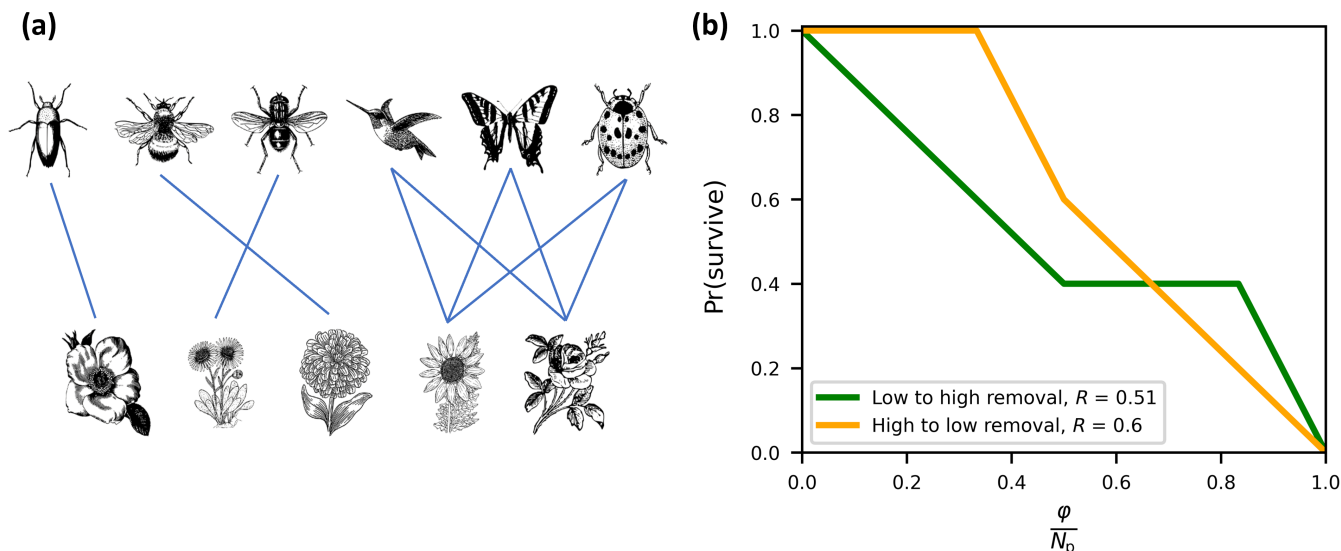


FIGURE 5 (a) Example ecological network and (b) targeted removal robustness curves. Removing low-degree primary species (pollinators) first gives a lower robustness than removing high-degree primary species. Illustrations used in (a) are all available in the public domain under a CC0 license.

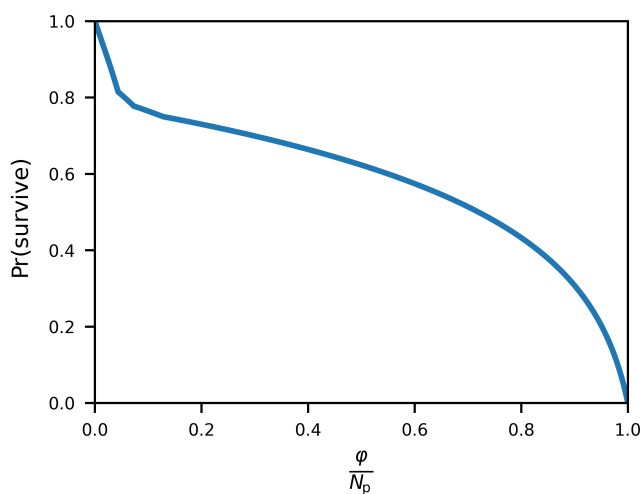


FIGURE 6 Analytically predicted robustness curve of a real-world seed-dispersal network, where primary frugivore species go extinct according to their IUCN extinction risk. This replicates results from Lamperty & Brosi (2022) using data from Bello et al. (2017).

species neighbors become extinct. Therefore, the extinction probability of predator species A is simply dependent upon the u_A initial number of unique plant species to which it is connected via its pollinator species neighbors, before any extinctions occur. Therefore, we can treat the predator species in this system as our secondary species and the plant species as our primary species.

Similar to Equation (3), the extinction probability for a secondary species A once some φ number of primary species have been removed is:

$$\Pr(u'_A = 0 | \varphi) = \begin{cases} \frac{\binom{N_p - u_A}{\varphi - u_A}}{\binom{N_p}{\varphi}} & \text{if } \varphi \geq u_A, \\ 0 & \text{otherwise,} \end{cases} \quad (18)$$

As before, we can average Equation (18) over $p(u)$, the distribution of secondary species connected to u unique primary species to predict secondary extinctions as primary species are removed to give:

$$\Pr(\text{extinct} | \varphi) = \sum_{u=0} p(u) \Pr(u' = 0 | \varphi). \quad (19)$$

This extinction probability may then be used to generate analytically predicted robustness curves such as the one given in Figure 7.

While this expands the scope of our analytic model beyond simply two-layer ecosystems, it is important to note that thus far we can only model extinctions on multilayer systems if species go extinct after losing all of their neighbors. Therefore, we cannot consider as many different extinction scenarios on multilayer networks as we can for bipartite networks.

In this scenario, the survival of a secondary species is only dependent upon the number of unique primary species to which it is connected via intermediary species. Theoretically, this may be scaled up to a network of arbitrarily many layers, and so secondary species will be separated from primary species by an arbitrary number of intermediary species. In this case, it may be useful to determine how the inclusion of additional layers affects

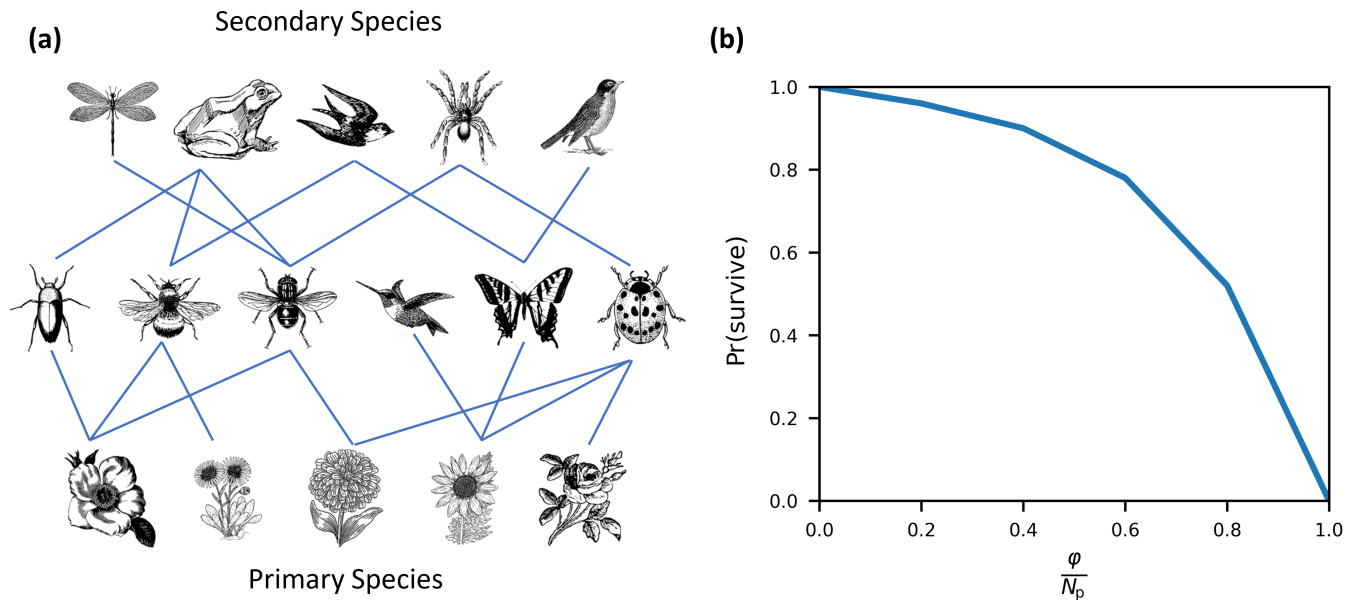


FIGURE 7 (a) Example of a three-layer network of plants (primary species), pollinators, and predators (secondary species), and (b) the associated analytically predicted robustness curve. Illustrations used in (a) are all available in the public domain under a CC0 license.

robustness, especially if some layers have few interactions and act as bottlenecks in the network.

FALSE NEGATIVES AND POSITIVES IN NETWORK DATA

Beyond simply predicting robustness, we may also be interested in how predictions of robustness are affected by errors in network data, as ecological data can be error prone (de Aguiar et al., 2019; Kangas et al., 2018). For example, networks may vary across environmental gradients (Pellissier et al., 2018) or may constitute metawebs inferred from proxies (Maiorano et al., 2020; Morales-Castilla et al., 2015). One may be interested in how the robustness of networks changes as true edges are removed (false negatives) or false edges are added in (false positives). In the simplest case, let us consider the random removal and addition of edges. Because robustness against random primary species removal only depends upon the degree distribution of secondary species, we can analytically predict how robustness will change as edges are randomly removed or added by modeling the changes to the secondary degree distribution.

For random edge removal and addition, we can predict changes to the secondary species degree distribution using the hypergeometric distribution. For modeling false negatives, let us consider some species A with initial degree k_A before any edges are removed. To determine

the probability that A has degree value $k'_A = k_A - j$ once some ϕ_n edges have been removed from the network, we use a hypergeometric distribution of a similar form to that shown in Equation (3), except we are now choosing from the total number of edges E that are initially in the network. The probability that A has degree $k'_A = k_A - j$ is given by

$$\Pr(k'_A = k_A - j | \phi_n) = \begin{cases} \frac{\binom{k_A}{j} \binom{E - k_A}{\phi_n - j}}{\binom{E}{\phi_n}} & \text{if } \phi_n \geq j, \\ 0 & \text{otherwise.} \end{cases} \quad (20)$$

Equation (20) may then be scaled up across the entire network to update the secondary degree distribution. Values in the updated degree distribution are given by:

$$p(k')_n = \frac{1}{N_s} \sum_A \Pr(k'_A = k_A - j | \phi_n). \quad (21)$$

For false positives, we can again use the hypergeometric distribution to predict changes to the network's secondary degree distribution, but we now consider the probability of adding edges to some species A. For species A with initial degree k_A before any edges are added, the number of potential edges that may be added is $N_p - k_A$, assuming that there are only single interactions

between primary and secondary species. Additionally, there are $N_s N_p - E$ possible edges that do not exist in the initial network and could potentially be added. Therefore, the probability that species A has degree value $k'_A = k_A + j$ after some φ_p false-positive edges have been added is:

$$\Pr(k'_A = k_A + j | \varphi_p) = \begin{cases} \frac{\binom{N_p - k_A}{j} \binom{N_s N_p - E - N_p + k_A}{\varphi_p - j}}{\binom{N_s N_p - E}{\varphi_n}} & \text{if } \varphi_p \geq j, \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

Similar to before, we can average Equation (22) across all species to update the degree distribution to give

$$p(k')_p = \frac{1}{N_s} \sum_A \Pr(k'_A = k_A + j | \varphi_p). \quad (23)$$

While Equations (21) and (23) give us updated degree distributions for the inclusion of false negatives and positives, respectively, they only tell us how false negatives and positives affect a network separately. What if we wish to model a network that includes both false negatives and positives?

To model false negatives and positives occurring on the same network, we can consider the convolution of Equations (20) and (22). This can ultimately give a degree distribution for a network that includes both false negatives and positives, although it is important to note that this method assumes false negatives and positives are independent. In other words, it is assumed that false negatives do not remove false-positive edges, and false-positive edges do not re-introduce true edges, which were removed as false negatives.

To predict the probability that some species A has degree k'_A after φ_n false negatives have been removed from and φ_p false positives have been added to the network, we can calculate:

$$\Pr(k'_A | \varphi_p \text{ and } \varphi_n) = \sum_{m=0} \Pr(m | \varphi_p) \Pr(k'_A - m | \varphi_n), \quad (24)$$

which is the convolution of Equations (20) and (22). As before, we can average over all species to get the updated degree distribution

$$p(k')_{n,p} = \frac{1}{N_s} \sum_A \Pr(k'_A | \varphi_p \text{ and } \varphi_n). \quad (25)$$

Using these equations, we can predict changes to a network's degree distribution when false negatives and positives are included, both separately and together. In the following, we use these updated degree distributions to predict how network robustness changes with the inclusion of false negatives and positives.

In Figure 8, we show the robustness curve for a plant-pollinator network undergoing random extinctions where $i_k = k$, comparing the originally predicted curve against the predicted curves for including false negatives alone, false positives alone, and false negatives and positives together.

In terms of robustness, false positives alone have a more significant impact than false negatives alone in small quantities. This is evident in Figure 8c, where we examine the total difference in robustness between networks with only false negatives or positives and the original network. We can see that false positives increase robustness more than false negatives decrease robustness up to and including a change in edges $\Delta\%E$ of approximately 40%, which for the network in question is 450 edges (out of 1125 edges in the original network).

The fact that false negatives have a greater impact on robustness than false positives in large quantities is perhaps not surprising, given that large quantities of false negatives entail the total disconnection of multiple species from the network. Large quantities of false positives appear to result in diminishing returns for increases to robustness, likely explained by the fact that excessive false positives tend to reinforce species that are already very well connected, and so are not made substantially more resistant to extinction with the addition of extra interactions.

We can also see from Figure 8b that if equal quantities of false negatives and positives are included simultaneously, robustness increases. This indicates that when false negatives and positives both occur in comparable quantities in network data, false positives will have a greater impact on robustness predictions than false negatives.

Additionally, this scenario is comparable to randomly rewiring the interactions in the network, as the total number of edges is conserved with edges randomly removed from some species and added to others. However, for random rewiring it may be permissible to rewire edges to reproduce edges that exist in the original network, which we disallow in our assumptions for adding false positives, so they are not necessarily identical scenarios. Nevertheless, based on these results

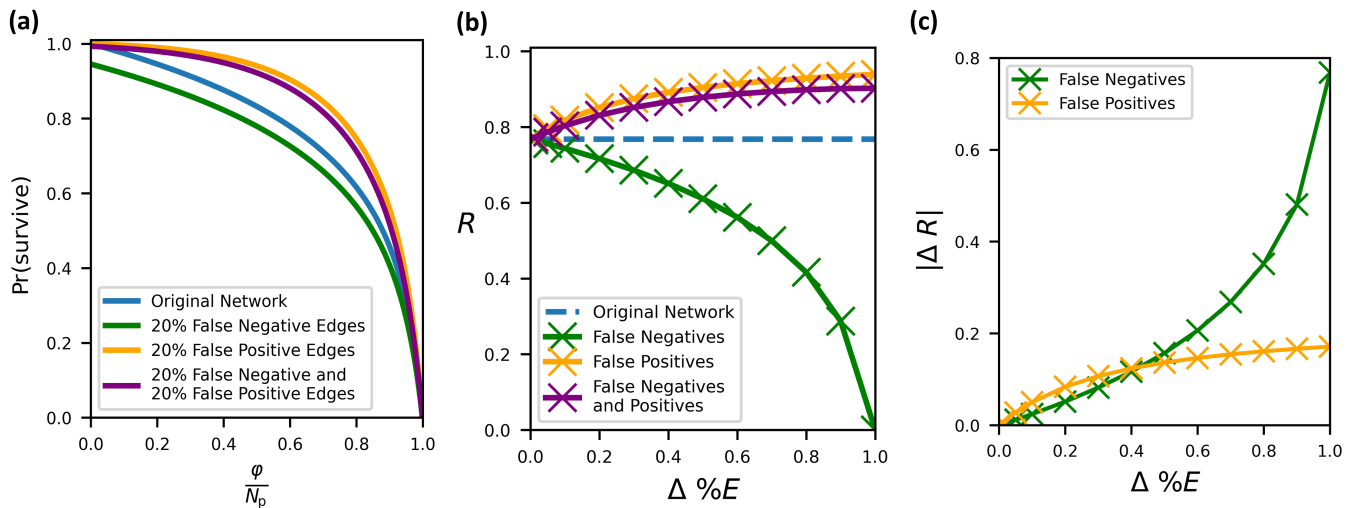


FIGURE 8 (a) Analytically predicted robustness curves for a real-world ecological network from a study by (Kato, 2000) for the original network, the network with false-negative edges randomly removed, the network with false-positive edges randomly added and the network with both false negatives and false positives simultaneously. Note that false edges totaling 20% of the original number of edges in the network are added or removed. (b) Shows how network robustness changes as edges are added or removed, and (c) shows the difference between the robustness of networks with only false negatives or false positives and the original network.

we might expect that the random rewiring of ecological networks improves their robustness, and this may be worthy of future consideration.

Here we have only considered a single ecological network; in Appendix S1: Section S1 we verify similar results on a dataset of 18 real-world plant–pollinator networks. For this dataset, we find that if we measure robustness where 20% of the original number of edges are added/removed for false positives/negatives alone, then the net change in robustness is typically positive, that is, false positives increase robustness more than false negatives decrease it for most networks. Additionally, we find that robustness increases on all networks we consider when 20% of the original number of edges are added and removed for false negatives and positives simultaneously. This supports our conclusion that false positives have a greater impact on our assessment of network robustness than false negatives.

If we assume that ecological network data gathering in the real world is reasonably accurate, that is, unlikely to over/under record interactions by more than 20%, then we would expect false positives to introduce more error into calculations of robustness than false negatives. Particular care is needed for robustness analyses based on metawebs of potential trophic interactions for which the false-positive and false-negative rates are difficult to ascertain (Maiorano et al., 2020; Morales-Castilla et al., 2015).

While this result indicates that false positives have more impact than false negatives, it only provides one perspective for how these errors may be introduced into

network data. One future avenue of inquiry is to establish the likely sources of errors and model those, as opposed to modeling errors randomly.

EXTINCTIONS DUE TO LOSS OF INTERACTION STRENGTH

In previous sections, we have only considered networks in which interactions are weighted equally; that is, each one of a secondary species' interactions is as important for its survival. However, in real ecological networks, a secondary species may interact more with one primary species than another, and this has an impact on a secondary species survivability (Berlow et al., 1999). We can specify a certain percentage of total interaction strength T that a species must lose before it goes extinct, an approach used before by (Schleuning et al., 2016). For some species A , there will be a threshold value of interaction strength i_A that must be lost for A to go extinct, and we may calculate i_A as $i_A = \lceil Tk_A \rceil$.

To model the loss of interaction strength mathematically, let us define two vectors associated with some species A . The first is the vector of interaction strength weights, $\mathbf{w}_A = [w_{A,1}, w_{A,2}, \dots, w_{A,k_A}]$, where \mathbf{w}_A has individual elements $w_{A,z}$ that correspond to the interaction strength shared with the z th neighbor of A .

Once some number of neighbors of A has gone extinct, we need to determine the probability that some combination of weights $w_{A,z}$ has been removed such that their sum meets or exceeds A 's extinction threshold i_A .

We may do this by specifying some random vector $\mathbf{Y}_A = [Y_{A,1}, Y_{A,2}, \dots, Y_{A,k_A}]$, where \mathbf{Y}_A has individual elements $Y_{A,z}$ corresponding to whether or not A has lost its z th neighbor. Each element $Y_{A,z}$ is a random variable that takes on values of 0 or 1. $Y_{A,z} = 0$ if the z th neighbor of A is still present in the network, and $Y_{A,z} = 1$ if the z th neighbor of A has gone extinct.

Therefore, if species A has lost some j neighbors, we may express the probability that A has gone extinct as

$$\begin{aligned} \Pr(\text{A extinct}|j) &= \Pr\left(\mathbf{w}_A \times \mathbf{Y}_A \geq i_A \mid \sum_{z=1}^{k_A} Y_{A,z} = j\right) \\ &= \Pr\left(\sum_{z=1}^{k_A} w_{A,z} Y_{A,z} \geq i_A \mid \sum_{z=1}^{k_A} Y_{A,z} = j\right), \end{aligned} \tag{26}$$

where $\mathbf{w}_A \times \mathbf{Y}_A = \sum_{z=1}^{k_A} w_{A,z} Y_{A,z}$ is the dot product of \mathbf{w}_A and \mathbf{Y}_A . Using Equation (26) we can update our extinction probability for some secondary species A given the extinction of some φ primary species to:

$$\begin{aligned} \Pr(\text{A extinct}|\varphi) &= \sum_{j=0}^{k_A} \Pr(k'_A = k_A - j|\varphi) \\ &\Pr\left(\sum_{z=1}^{k_A} w_{A,z} Y_{A,z} \geq i_A \mid \sum_{z=1}^{k_A} Y_{A,z} = j\right). \end{aligned} \tag{27}$$

Therefore, the robustness of an ecological network for this scenario may be calculated as:

$$R = 1 - \frac{1}{N_p + 1} \sum_{\varphi=0}^{N_p} \sum_A \Pr(\text{A extinct}|\varphi). \tag{28}$$

However, it is challenging to predict robustness analytically in this scenario because the extinction probability $\Pr(\text{A extinct}|j)$ depends on the weight vector \mathbf{w}_A , the elements of which do not necessarily follow any particular distribution. It is instead necessary to estimate $\Pr(\text{A extinct}|j)$ in some way. The brute force method is to repeatedly randomly sample the vector \mathbf{Y}_A under the condition $\sum_{z=1}^{k_A} Y_{A,z} = j$ and count the number of samples for which $\sum_{z=1}^{k_A} w_{A,z} Y_{A,z} \geq i_A$ to estimate $\Pr(\text{A extinct}|j)$. However, this Monte Carlo method approach of repeated random sampling must be repeated many times to give an accurate estimate, and is subject to statistical fluctuations.

Instead, we have developed a deterministic sampling method, where for some species A the weight vector \mathbf{w}_A has its elements $w_{A,z}$ arranged in descending size order, such that $w_{A,1} \geq w_{A,2} \geq \dots \geq w_{A,k_A}$. We then estimate

$\Pr\left(\sum_{z=1}^{k_A} w_{A,z} Y_{A,z} \geq i_A \mid \sum_{z=1}^{k_A} Y_{A,z} = j\right)$ by considering the sum of the first δ elements of $\mathbf{w}_A \cdot \mathbf{Y}_A$, that is, we calculate $\Pr\left(\sum_{z=1}^{\delta} w_{A,z} Y_{A,z} \geq i_A \mid \sum_{z=1}^{\delta} Y_{A,z} \leq j\right)$. We refer to the maximum number of elements δ as the “depth” of our sampling, with higher δ values giving more accurate estimations with the trade-off of greater computational cost.

Due to the ordering of the weight vector \mathbf{w}_A , we can ensure that our sampling method is deterministic, while also improving upon the Monte Carlo method in terms of estimation quality and speed by using a method that allows us to calculate $\Pr\left(\sum_{z=1}^{\delta} w_{A,z} Y_{A,z} \geq i_A \mid \sum_{z=1}^{\delta} Y_{A,z} \leq j\right)$ without having to explicitly calculate all possible $\sum_{z=1}^{\delta} w_{A,z} Y_{A,z}$ values. Further details of how our sampling method works are provided in Appendix S1: Section S2.

Using our deterministic sampling method, we can provide quasi-analytic predictions for secondary species survival on networks undergoing random primary extinctions, where interaction strength is weighted unevenly and secondary extinctions occur after the loss of a certain percentage of interaction strength. Example predictions are given in Figure 9, alongside results showing how our deterministic estimate becomes increasingly accurate with greater depth, and a comparison between the time taken to estimate $\Pr(\text{A extinct}|j)$ and prediction accuracy for our deterministic sampling method and for a brute force Monte Carlo method.

From this, we can see that as the depth of the deterministic estimation increases, we get diminishing returns in terms of prediction accuracy, and that it is more computationally efficient to use the deterministic estimation method as opposed to Monte Carlo simulation to get the same level of prediction accuracy.

SPECIES SPECIALIZATION AND GENERALIZATION

Having developed an analytic framework for secondary species extinctions when interaction strength is weighted unevenly, we can examine some extreme scenarios of interaction strength weighting. One property of interest in ecological networks is specialization (Blüthgen et al., 2006), where specialist species tend to interact very strongly with a small number of species, and generalist species tend to interact with many species evenly. Given an ecological network with a set number of primary and secondary species, and a set distribution of interactions, we can examine the most specialist interaction weighting and the most generalist interaction weighting.

In the most specialist case, each secondary species weights one of its interactions at close to 100% of its interaction strength, and all others close to 0%.

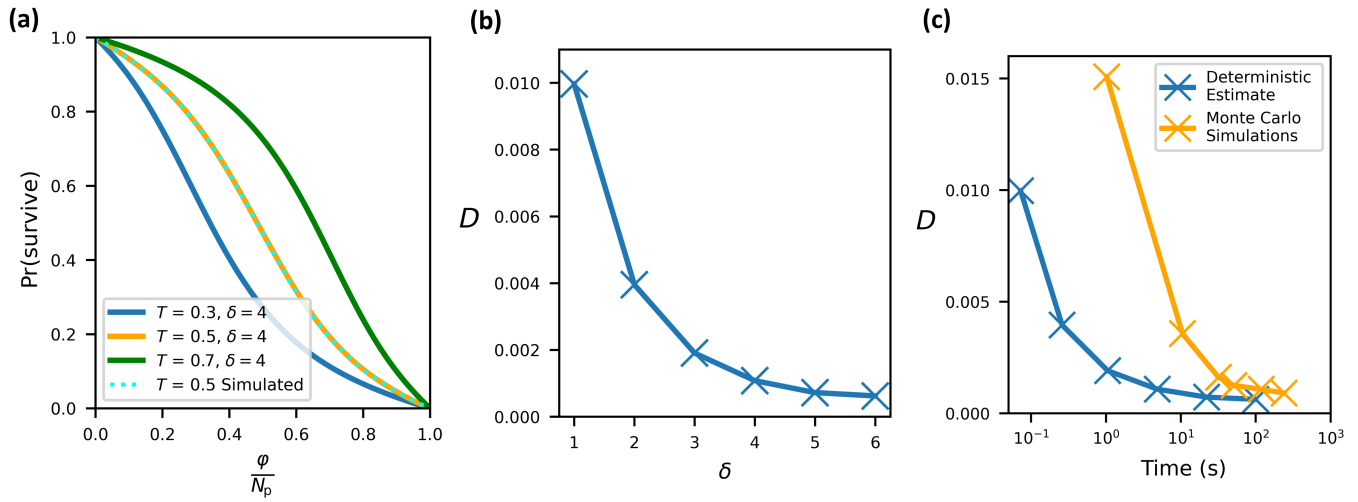


FIGURE 9 (a) Analytically predicted and computationally simulated robustness curves for a real-world network from a study by (Kato, 2000) where unevenly weighted interaction strength is taken into account and extinctions occur over a specified threshold T of interaction strength loss. Analytic predictions are given for threshold values of 70%, 50%, and 30%, and computationally simulated curve averaged over 5000 iterations is given for 50%. (b) Comparison between depth of the estimation for $\text{Pr}(A \text{ extinct}|j)$ and divergence (as defined in Equation (7)) between analytically predicted robustness curve and the simulated curve averaged over 5000 iterations. (c) The divergence between the predicted robustness curve and the 5000 iteration simulation curve compared against the time taken, with data for both the deterministic estimation method and the Monte Carlo method.

Therefore, a given secondary species A only goes extinct when it loses the neighbor with which it shares almost all interaction strength. If a neighbor of A goes extinct, the probability of losing the heavily weighted neighbor is simply $\frac{1}{k_A}$, so $\text{Pr}(A \text{ extinct}|j) = \frac{j}{k_A}$. This gives an extinction probability for some species A of:

$$\begin{aligned} \text{Pr}(A \text{ extinct}|\varphi) &= \sum_{j=0}^{k_A} \text{Pr}(k'_A = k_A - j|\varphi) \frac{j}{k_A}, \\ &= \frac{\mathbb{E}(k'_A)}{k_A}, \\ &= \frac{\varphi}{N_p}, \end{aligned} \quad (29)$$

which we derive from the fact that $\mathbb{E}(k'_A) = k_A \frac{\varphi}{N_p}$ because $\text{Pr}(k'_A = k_A - j|\varphi)$ describes the hypergeometric distribution. This results in a robustness value of $R = 0.5$, regardless of the secondary degree distribution, number of primary species or threshold.

For the most generalist case, each secondary species weights all of its interactions evenly, which means $\text{Pr}(A \text{ extinct}|j) = 0$ when $j < i_k$, and $\text{Pr}(A \text{ extinct}|j) = 1$ when $j \geq i_k$. Therefore, $\text{Pr}(A \text{ extinct}|\varphi) = \sum_{j=i_k}^{k_A} \text{Pr}(k'_A = k_A - j|\varphi)$, the same as Equation (4).

Given these results, when is it more advantageous for a network to be highly specialist or highly generalist in terms of robustness? Let us consider some secondary

species A that is connected to all primary species in its network, that is, $k_A = N_p$. Every primary species that becomes extinct will reduce the degree of A , with A going extinct once the number of primary species removed φ is equal to or exceeds the extinction threshold i_k . Therefore, the extinction probability for A is given by:

$$\text{Pr}(A \text{ extinct}|\varphi) = \begin{cases} 1 & \text{if } \varphi \geq i_k, \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

If all secondary species in a network have $k = N_p$, then when they are maximally generalist, the network robustness is $R = \frac{i_k}{N_p}$. Therefore, such a network is more robust when secondary species are maximally generalist if more than 50% of interaction strength must be lost before a secondary species goes extinct, that is, when the sensitivity threshold $T > 0.5$. Conversely, the network is more robust when secondary species are maximally specialist if less than 50% of interaction strength must be lost to make secondary species go extinct, so $T < 0.5$. To illustrate this, we provide robustness curves in Figure 10a a single secondary species with $k_A = N_p$, and of Figure 10b an entire real-world network.

From these results, we know that either extreme of species specialization can be advantageous from the perspective of maximizing network robustness, depending upon the sensitivity of the network, that is, the proportion of interaction strength that must be lost for

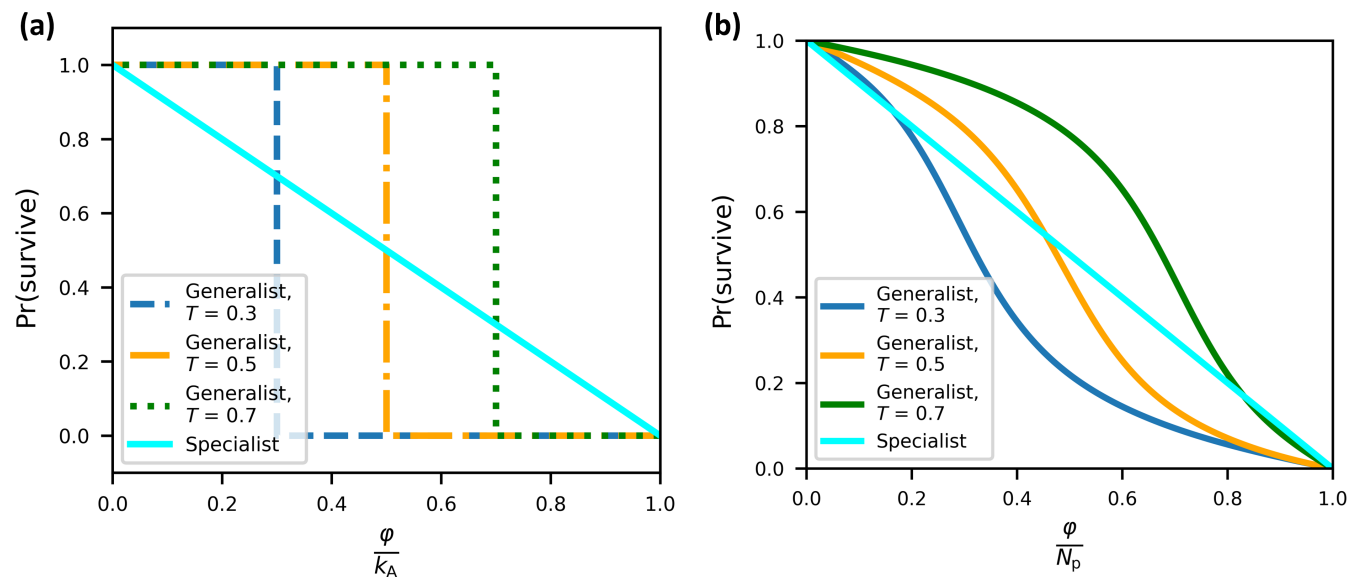


FIGURE 10 (a) Shows species survival for maximal generalists and maximal specialists at various sensitivity thresholds T when $k_A = N_p$. (b) Gives robustness curves for maximal generalization and maximal specialization on a real-world network from a study by Kato (2000) at various sensitivity thresholds T .

secondary species to go extinct. However, we see in Figure 10b that it is not strictly the case on real networks that maximum generalization is always better for robustness than maximum specialization when $T > 0.5$, as the maximum generalist curve when $T = 0.5$ gives $R = 0.471$, whereas the maximum specialist curve has $R = 0.5$. This is because secondary species typically have $k < N_p$ on real networks. Additionally, we note that the robustness values from the maximum generalist and maximum specialist interaction weightings do not necessarily give the maximum and minimum robustness values for a given threshold. Nevertheless, these results are still indicative of the fact that species generalization and specialization can both improve network robustness in different contexts, and so we might expect that in the real world, a network that has developed to be highly generalist is less sensitive to interaction loss than a network that has developed to be highly specialist.

INTERACTION LOSS

The models we have considered are only concerned with the loss of primary species as a whole, where a primary species is removed at each “step” in the extinction process. Considering the loss of entire species at a time can skew our understanding of network robustness, for example a plant–animal network with more animals than plants will appear more robust against primary extinctions of animals than against primary extinctions of plants (Schleuning et al., 2016). However, is this a

realistic understanding of how species go extinct? There may be more animal species than plant species, but what if there is a very large population of each plant species and a small population of each animal species? Extinctions may be experienced more gradually, where a species’ population dies off over time rather than all at once (Valiente-Banuet et al., 2015). This process can be modeled by examining the loss of interactions as opposed to the loss of species, which in network terms entails considering edge removal as opposed to node removal.

If the interaction strength between species is represented as integer values, then we can treat each unit of interaction strength as an edge, so secondary species have degree values equal to the sum of their interaction strength with other species. We then have E “edges” (i.e., total interaction strength on the network), and we remove some φ units of interaction strength. For a given secondary species A , after removing some φ interaction strength the probability that it has lost some interaction strength j is:

$$\Pr(k'_A = k_A - j | \varphi) = \begin{cases} \frac{\binom{k_A}{j} \binom{E - k_A}{\varphi - j}}{\binom{E}{\varphi}} & \text{if } \varphi \geq j, \\ 0 & \text{otherwise,} \end{cases} \quad (31)$$

which is the same as Equation (20) for predicting the change in secondary species degree due to false negatives.

However, it should be noted that for this scenario, k_A is the initial total interaction strength of A, and k'_A is the total interaction strength of A after removing φ interaction strength, because we are treating each unit of interaction strength as an edge. From this, it is straightforward to predict secondary species' survival probability and network robustness using a similar logic as in the section *Introducing the analytic framework*. Predictions for interaction loss on an ecological network are given in Figure 11.

These predictions of interaction loss give different robustness values than predictions of species extinctions. For example, for species extinctions on a real network (Kato, 2000) when true interaction strength values are used (as shown in Figure 9a) and $T=0.7$, we have $R_{\text{species}}=0.619$. By contrast, for interaction loss on the same network when $T=0.7$, we have $R_{\text{inter}}=0.653$. Therefore, modeling secondary species extinctions as an outcome of interaction loss as opposed to primary species extinctions gives a different perspective on network robustness, allowing one to identify networks that are fragile against primary species loss but robust against interaction loss, and vice versa.

A similar logic to that presented in the section *Introducing the analytic framework* may be followed to show that a network with a set number of primary species, secondary species and total interaction strength is maximally robust against random interaction loss when all secondary species have equal interaction strength, and a proof of this is provided in Appendix S1: Section S3.

While we have discussed extinctions in this section as being “secondary” extinctions for terminological consistency to denote the group we are interested in measuring, in this scenario there are no de facto primary and secondary

groups of species. Therefore, a possible future extension is to consider the vulnerability of all species in the network to gradual interaction loss, alongside the structures that make a network most robust in such a scenario.

DISCUSSION

In conclusion, we have successfully extended the robustness framework of (Memmott et al., 2004) such that we may make predictions of ecological network robustness analytically. For random extinctions, we have shown that degree equality among secondary species entails maximum robustness for a set number of interactions. We are also able to predict secondary extinctions as primary species go extinct according to some degree or extinction risk-based targeting, and we can predict secondary extinctions on ecological networks with more than two layers. Additionally, we can model the influence of random false negatives and positives in network data on robustness, finding that in small quantities false positives have a greater impact than false negatives on network robustness.

Our model is also capable of predicting the robustness of networks where interaction strength is weighted unevenly between different secondary species' neighbors, and species go extinct once a certain proportion of interaction strength has been lost. We have given results for robustness when interaction strength is equally distributed (maximally generalist), and when interaction strength is shared solely with one neighboring species (maximally specialist). From this, we know that maximal generalization and maximal specialization can both produce a more robust network, depending on the proportion of interaction strength that must be lost before secondary species extinction. Finally, we have demonstrated the fact that it is also possible to model interaction strength loss as opposed to simply species extinction, representing a more “gradual” extinction scenario.

These results represent a substantial advancement in analytic understanding of ecological network robustness. However, there are still many open questions. We can predict the average secondary species extinction probability for a given number of primary extinctions, but we may also want to analytically predict the possible error in robustness curves by finding the standard deviation in secondary species extinction probability for a given number of primary extinctions. Additionally, we may want to establish the true worst- and best-case scenarios for secondary extinctions, as these have not been definitively identified. For errors (i.e., false negatives and positives) in network data, our current results examine errors that occur at random, but this may not be the case in the real world. Errors may occur due to some specific reason or

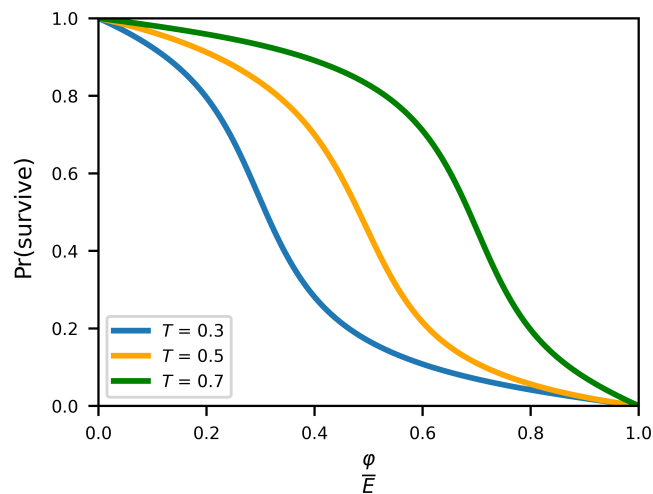


FIGURE 11 Robustness curves for interaction loss on a real-world network from a study by (Kato, 2000) at varying sensitivity thresholds T .

dynamic, and identifying what this may allow us to better mathematically model data errors and their influence.

Beyond these possible improvements, it is also important to acknowledge that in recent years, ecologists have considered properties of ecological networks that affect robustness and go beyond simpler models of species extinction. For example, ontogenetic niche shifts, where species change their diets when they undergo changes such as growing from a larva to an adult, can affect the structure and robustness of interaction networks (Nakazawa, 2015). Another consideration is how interactions can be “rewired” after species extinctions (Baldock et al., 2019; Schleuning et al., 2016), which to predict analytically would likely require combinatoric methods for sampling with fuzzy replacement (Kesemen et al., 2021). These examples highlight the fact that there is still considerable room for analytic models of ecological network robustness to develop, and there are ongoing areas of research in both ecological networks and combinatorics, which may complement one another well, so it may be useful for there to be a greater dialogue between these fields in future.

ACKNOWLEDGMENTS

We thank Jane Memmott for providing feedback on a summary version of this paper. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

C.J. was supported by Engineering and Physical Sciences Research Council Doctoral Training Partnership funding (EP/R513179/1). We thank the Research Focus Data-centric Sciences of the University of Potsdam for financial support.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.


DATA AVAILABILITY STATEMENT

Data were sourced from Bello et al. (2017) at <https://doi.org/10.1002/ecy.1818>. Additional data were sourced from Fortuna et al. (2014) at <https://www.web-of-life.es/>. Network IDs of data from Fortuna et al. (2014) are as follows: M PL 004, M PL 005, M PL 009, M PL 015, M PL 016, M PL 021, M PL 028, M PL 029, M PL 043, M PL 044, M PL 047, M PL 048, M PL 049, M PL 054, M PL 055, M PL 056, M PL 057, M PL 062. Code (Jones, 2023) is available in Zenodo at <https://doi.org/10.5281/zenodo.10403155>.

ORCID

Chris Jones  <https://orcid.org/0000-0003-4698-5573>

Damaris Zurell  <https://orcid.org/0000-0002-4628-3558>

Karoline Wiesner  <https://orcid.org/0000-0003-2944-1988>

REFERENCES

- Aguilera, P. A., A. Fernández, R. Fernández, R. Rumí, and A. Salmerón. 2011. “Bayesian Networks in Environmental Modelling.” *Environmental Modelling & Software* 26(12): 1376–88.
- Baldock, K. C., M. A. Goddard, D. M. Hicks, W. E. Kunin, N. Mitschunas, H. Morse, L. M. Osgathorpe, et al. 2019. “A Systems Approach Reveals Urban Pollinator Hotspots and Conservation Opportunities.” *Nature Ecology & Evolution* 3(3): 363–373.
- Bellard, C., C. Bertelsmeier, P. Leadley, W. Thuiller, and F. Courchamp. 2012. “Impacts of Climate Change on the Future of Biodiversity.” *Ecology Letters* 15(4): 365–377.
- Bello, C., M. Galetti, D. Montan, M. Pizo, T. Mariguela, L. Culot, F. Bufalo, et al. 2017. “Atlantic Frugivory: A Plant-Frugivore Interaction Dataset for the Atlantic Forest.” *Ecology* 98: 3. <https://doi.org/10.1002/ecy.1818>.
- Berlow, E. L., S. A. Navarrete, C. J. Briggs, M. E. Power, and B. A. Menge. 1999. “Quantifying Variation in the Strengths of Species Interactions.” *Ecology* 80(7): 2206–24.
- Bersier, L.-F. 2007. “A History of the Study of Ecological Networks,” In *Biological Networks*, Vol. 3, edited by F. Kepes, 365–421. Singapore: World Scientific.
- Blüthgen, N., F. Menzel, and N. Blüthgen. 2006. “Measuring Specialization in Species Interaction Networks.” *BMC Ecology* 6: 2. <https://doi.org/10.1186/1472-6785-6-9>.
- Burgos, E., H. Ceva, R. P. Perazzo, M. Devoto, D. Medan, M. Zimmermann, and A. María Delbue. 2007. “Why Nestedness in Mutualistic Networks?” *Journal of Theoretical Biology* 249(2): 307–313.
- Cohen, J. E. 1978. *Food Webs and Niche Space. (MPB-11)*, Vol. 11. Princeton, NJ: Princeton University Press.
- Dawson, T. P., S. T. Jackson, J. I. House, I. C. Prentice, and G. M. Mace. 2011. “Beyond Predictions: Biodiversity Conservation in a Changing Climate.” *Science* 332(6025): 53–58.
- de Aguiar, M. A., E. A. Newman, M. M. Pires, J. D. Yeakel, C. Boettiger, L. A. Burkle, D. Gravel, et al. 2019. “Revealing Biases in the Sampling of Ecological Interaction Networks.” *PeerJ* 7: e7566.
- Fortuna, M. A., R. Ortega, and J. Bascompte. 2014. “The Web of Life.” arxiv:1403.2575. <https://doi.org/10.48550/arXiv.1403.2575>.
- Forup, M. L., K. S. Henson, P. G. Craze, and J. Memmott. 2008. “The Restoration of Ecological Interactions: Plant–Pollinator Networks on Ancient and Restored Heathlands.” *Journal of Applied Ecology* 45(3): 742–752.
- Haecckel, E. 1866. *Generelle morphologie der organismen. Allgemeine grundzüge der organischen formen-wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte descendenztheorie*. Berlin: G. Reimer.
- Ings, T. C., and J. E. Hawes. 2018. *The History of Ecological Networks* 15–28. Cham: Springer International Publishing.
- Jones, C. 2023. “Eco-analytic.” [10.5281/zenodo.10403155](https://doi.org/10.5281/zenodo.10403155).
- Jongman, R. H. 1995. “Nature Conservation Planning in Europe: Developing Ecological Networks.” *Landscape and Urban Planning* 32(3): 169–183.
- Kangas, A., T. Packalen, K. Korhonen, and J. Vauhkonen. 2018. “Sources and Types of Uncertainties in the Information on Forest-Related Ecosystem Services.” *Forest Ecology and Management* 427: 7–16. <https://doi.org/10.1016/j.foreco.2018.05.056>.

- Kato, M. 2000. "Anthophilous Insect Community and Plant-Pollinator Interactions on Amami Islands in the Ryukyu Archipelago, Japan." *Contributions from the Biological Laboratory, Kyoto University* 29(2): 157–254.
- Kesemen, O., B. Tiryaki, Ö. Tezel, E. Özkul, and E. Naz. 2021. "Random Sampling with Fuzzy Replacement." *Expert Systems with Applications* 185: 115602.
- Lamperty, T., and B. J. Brosi. 2022. "Loss of Endangered Frugivores from Seed Dispersal Networks Generates Severe Mutualism Disruption." *Proceedings of the Royal Society B* 289(1984): 20220887.
- Maiorano, L., A. Montemaggiore, G. F. Ficetola, L. O'connor, and W. Thuiller. 2020. "Tetra-eu 1.0: A Species-Level Trophic Metaweb of European Tetrapods." *Global Ecology and Biogeography* 29(9): 1452–57.
- Memmott, J., N. M. Waser, and M. V. Price. 2004. "Tolerance of Pollination Networks to Species Extinctions." *Proceedings: Biological Sciences* 271(1557): 2605–11.
- Montoya, J., S. Pimm, and R. Sole. 2006. "Ecological Networks and their Fragility." *Nature* 442: 259–264.
- Morales-Castilla, I., M. G. Matias, D. Gravel, and M. B. Araújo. 2015. "Inferring Biotic Interactions from Proxies." *Trends in Ecology & Evolution* 30(6): 347–356.
- Nakazawa, T. 2015. "Ontogenetic Niche Shifts Matter in Community Ecology: A Review and Future Perspectives." *Population Ecology* 57(2): 347–354.
- Pellissier, L., C. Albouy, J. Bascompte, N. Farwig, C. Graham, M. Loreau, M. A. Maglianesi, et al. 2018. "Comparing Species Interaction Networks along Environmental Gradients." *Biological Reviews* 93(2): 785–800.
- Pocock, M. J. O., D. M. Evans, and J. Memmott. 2012. "The Robustness and Restoration of a Network of Ecological Networks." *Science* 335(6071): 973–77.
- Ramazi, P., M. Kunegel-Lion, R. Greiner, and M. A. Lewis. 2021. "Exploiting the Full Potential of Bayesian Networks in Predictive Ecology." *Methods in Ecology and Evolution* 12(1): 135–149.
- Schleuning, M., J. Fründ, O. Schweiger, E. Welk, J. Albrecht, M. Albrecht, M. Beil, et al. 2016. "Ecological Networks Are more Sensitive to Plant than to Animal Extinction under Climate Change." *Nature Communications* 7: 12. <https://doi.org/10.1038/ncomms13965>.
- Solé, R. V., and J. M. Montoya. 2001. "Complexity and Fragility in Ecological Networks." *Proceedings: Biological Sciences* 268(1480): 2039–45.
- Tylianakis, J. M., E. Laliberté, A. Nielsen, and J. Bascompte. 2010. "Conservation of Species Interaction Networks." *Biological Conservation* 143(10): 2270–79.
- Valiente-Banuet, A., M. A. Aizen, J. M. Alcántara, J. Arroyo, A. Cocucci, M. Galetti, M. B. García, et al. 2015. "Beyond Species Loss: The Extinction of Ecological Interactions in a Changing World." *Functional Ecology* 29(3): 299–307.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Jones, Chris, Damaris Zurell, and Karoline Wiesner. 2024. "Novel Analytic Methods for Predicting Extinctions in Ecological Networks." *Ecological Monographs* e1601. <https://doi.org/10.1002/ecm.1601>