

Secondary extinctions in food webs: a Bayesian network approach

Anna Eklöf^{1*†}, Si Tang¹ and Stefano Allesina^{1,2}

¹Department of Ecology & Evolution, University of Chicago, Chicago, IL, USA; and ²Computation Institute, University of Chicago, Chicago, IL, USA

Summary

1. Ecological communities are composed of populations connected in tangled networks of ecological interactions. Therefore, the extinction of a species can reverberate through the network and cause other (possibly distantly connected) species to go extinct as well. The study of these secondary extinctions is a fertile area of research in ecological network theory.

2. However, to facilitate practical applications, several improvements to the current analytical approaches are needed. In particular, we need to consider that (i) species have different ‘a priori’ probabilities of extinction, (ii) disturbances can simultaneously affect several species, and (iii) extinction risk of consumers likely grows with resource loss. All these points can be included in dynamical models, which are, however, difficult to parameterize.

3. Here we advance the study of secondary extinctions with Bayesian networks. We show how this approach can account for different extinction responses using binary – where each resource has the same importance – and quantitative data – where resources are weighted by their importance. We simulate ecological networks using a popular dynamical model (the Allometric Trophic Network model) and use it to test our method.

4. We find that the Bayesian network model captures the majority of the secondary extinctions produced by the dynamical model and that consumers’ responses to species loss are best modelled using a nonlinear sigmoid function. We also show that an approach based exclusively on food web structure loses power when species at higher trophic levels are preferentially lost. Because the loss of apex predators is unfortunately widespread, the results highlight a serious limitation of studies on network robustness.

Key-words: Bayesian networks, biodiversity loss, cascading extinctions, dynamical model, food webs

Introduction

Species extinctions have reached an unprecedented rate (Barnosky *et al.* 2011), making biodiversity loss one of the most severe threats to ecosystems around the world (Reich *et al.* 2012). Often, extinctions stem from anthropogenic habitat loss, over-harvesting and climate change (Pereira *et al.* 2010), and are likely to have profound effects on important ecological services (Worm *et al.* 2006). To forecast extinction risk, we would like to estimate the probability of each and every species becoming extinct in an ecological network. Certain key species traits are likely to influence extinction vulnerability: for example, large body size, high trophic level and low density all increase the probability of extinction (Gaston & Blackburn 1995; Purvis *et al.* 2000; Cardillo *et al.* 2005; Davidson *et al.* 2009; Lee & Jetz 2011). However, species are not isolated, but rather depend on each other for sustenance, forming a complex network of ecological interactions. There-

fore, the extinction of a single species could affect other species with which it interacts, directly or indirectly (Ebenman & Jonsson 2005), potentially setting in motion a cascade of secondary extinctions through the community. These secondary extinctions can emerge from either bottom-up effects (consumers losing their resources) or top-down effects (resources responding to the loss of their consumers).

Traditionally, there have been two main approaches to the analysis of secondary extinctions in ecological networks – often called network robustness. The first line of research, originating from studies of other complex networks (Albert, Jeong & Barabási 2000), focuses exclusively on the presence or absence of consumer–resource relationships. Thus, only the qualitative network structure is taken into account. Typically, one removes species, either randomly or systematically, and tests how network robustness varies with network properties such as number of species or connectance (Sole & Montoya, 2001; Dunne, Williams & Martinez 2002; Memmott, Waser & Price 2004; Srinivasan *et al.* 2007). This so-called topological approach has the advantage of requiring only the network structure as an input: for an adjacency matrix A , whose rows and columns represent the species, a coefficient $A_{ij} = 1$

*Correspondence author. E-mail: anna.eklof@liu.se

†Present address: Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden

signifies that i is a prey of j . Although this simplicity makes it possible to analyse very large networks, the approach has several limitations. For example, in the topological case, secondary extinctions only occur when a consumer loses all of its resources: the extinction risk does not grow until all resources are lost, at which point the extinction risk equals one. Also, in the topological approach, all species are usually assumed to have the same baseline probability of extinction, whereas in natural systems some species are more vulnerable than others.

An alternative line of research attempts to explicitly model population dynamics, that is, changes in abundances or biomasses over time, for all species in the network (Ebenman, Law & Borrvall 2004; Eklöf & Ebenman 2006; Riede *et al.* 2011; Stouffer & Bascompte 2011). Using dynamical models one can capture, in addition to the purely topological extinctions, other types of extinctions. For example, through the propagation of indirect effects in the network, the primary extinction of a top predator might lead to the secondary extinction of some of its resources [top-down extinctions, for example, Ebenman & Jonsson (2005)]. Additionally, dynamic models often include an extinction threshold, a population density below which the species are considered extinct, in order to account for processes such as demographic stochasticity (Eklöf & Ebenman 2006). As such, a resource present at low values could still be insufficient to support its consumers. However, dynamical models require an extensive set of parameters, making an empirical parameterization of large food webs next to impossible. Typically, this approach has been used mostly to study synthetic webs generated using physiological scaling of species interaction strengths (Binzer *et al.* 2011). Moreover, because dynamical ecological systems are highly nonlinear, slightly different initial conditions can lead to very different outcomes. This makes it necessary to simulate numerous replicates for each parameterization. Finally, even if one were to measure all parameters correctly, this approach is difficult to extend to the study of very large networks due to limited computing power.

A middle-ground approach is to consider the probability that species will be present or absent in a complex system. Such a framework requires relatively few parameters and assumptions, yet it can account for a wide range of extinction types. One recent example of this type of model is the stochastic ecological network occupancy (SENO) model, which takes the topological network structure as well as colonization and extinction rates as input parameters, addressing the changes in species probabilities over space and time (Lafferty & Dunne 2010). Extensive simulations will converge on the actual probability of extinction for each species, but exact solutions (in the absence of top-down effects) can alternatively be found using Bayesian networks (Lafferty & Dunne 2010).

Here, we explore the use of Bayesian networks (Jensen, 1996) to directly calculate the marginal probability of species extinction in a network without requiring simulation. We add considerable flexibility in the assumptions about how consumers respond to the loss of resources.

A Bayesian network is simply a collection of random variables (here species are represented as Bernoulli random

variables determining their presence/absence) with arrows describing their conditional dependencies (feeding relationships). As such, the probability of extinction of each species depends on the state of its resources, which in turn depends on the state of their resources. The use of Bayesian networks has several advantages over the more traditional ways of modelling species extinctions. First, in Bayesian networks, one can directly assign to each species a different baseline probability of extinction. This baseline probability is then combined with the network structure to estimate extinction risk. This is useful for conservation, where lists of endangered species ('Red Lists') are often available. The main benefit from a modelling standpoint is that we need few parameters and thus few assumptions about the biological interactions.

Bayesian networks can be solved numerically very efficiently – multiple simulation reiterations are not needed, and therefore, computation time is greatly reduced. There is no need for artificial 'sequences' of extinctions, since all possible cases are considered simultaneously. Finally, as we show here, many simulation-based approaches are in fact simulating a Bayesian network that can be solved more efficiently. These benefits stress the importance of connecting ecology with the vast literature on graphical models.

Using Bayesian networks, we introduce a flexible method in which all the possible responses of consumers to resource loss can be modelled. To test our Bayesian network method, we parameterize a model based on differential equations that is frequently used to simulate complex food web dynamics (Berlow *et al.* 2009; Binzer *et al.* 2011). We first perform *in silico* (simulated) extinctions for the full-fledged dynamical model. We then use our method and attempt to predict the observed extinctions. We evaluate the goodness-of-fit for alternative responses of consumers to resource extinction using likelihoods. We find that a sigmoid response, in which consumers' extinction risk grows sharply after a critical fraction of resources is lost, best accounts for the observed extinctions. Moreover, adding information on resource importance further improves the forecasting ability of our Bayesian network method.

Materials and methods

BAYESIAN NETWORKS

A Bayesian network (or Belief Network) is a graphical model for the probabilistic relationships among a set of variables. More precisely, it represents a factorization of the merged probability distributions over finite sets of variables (Jensen, 1996): the probability distribution of a variable (in our case, the presence or absence of a species) can be factored in several parts, each accounting for a state of its parent variables (its resources). The nodes of the network represent the variables, and the links represent the conditional dependencies among the variables – there is a specified set of local probability distributions for each node that is conditional on the state of the parent nodes. In our case, each species (node) is a random variable (Bernoulli, taking values extant/extinct), and the feeding interactions represent probabilistic dependencies among the species. As such, we model species extinction probability as a function of the state of its parent nodes – its resources.

Therefore, we need to compile, for each consumer i , a table of conditional probabilities expressing its probability of extinction given the state of its resources (Fig. 1). This set of conditional probability tables is the Bayesian network. Next, we show how the network can be solved (i.e. the tables combined) to obtain the marginal probability of extinction (i.e. the probability of extinction once accounted for all possible outcomes) for each species.

CONDITIONAL PROBABILITIES AND RESPONSE TO EXTINCTION

Suppose that a predator i has two prey, X_1 and X_2 . We need to determine four conditional probabilities. First, $P(\neg i|X_1, X_2) = \pi_i$ is the probability that i goes extinct (\neg) given that both its resources are present. We denote this 'baseline' probability of extinction as π_i throughout the text. This quantity represents the probability that i goes extinct for causes other than the extinction of its resources (e.g. small population, overfishing, habitat destruction, disease). Then, we need to account for the probability of extinction when either of the two prey is absent. We denote these quantities $P(\neg i|\neg X_1, X_2)$ and $P(\neg i|X_1, \neg X_2)$. These two probabilities can be either equal, if losing X_1 or X_2 has the same effect on the probability of extinction of i , or different, if one resource is more important than the other. Note that under the topological approach described above, both probabilities would be equal to π_i , since only the loss of all prey will cause a species' probability of extinction to change (become 1). Note that in case of models considering stage-structured populations, each resource might be critical for a given stage (Rudolf & Lafferty 2011). For simplicity, we present only the case of species consisting on a single life stage. Finally, $P(\neg i|\neg X_1, \neg X_2) = 1$, as consumers cannot survive without resources. Also note that primary producers rely exclusively on the abiotic environment for their sustenance and thus go extinct with probability π_i , regardless of the state of the other species.

We can use these tables of conditional probabilities to model the response of each consumer to the loss of its resources. Although the method can account for any possible table of conditional probabilities, here we limit our analysis to a few simple, yet biologically plausible scenarios (Fig. 2): (i) the topological case in which the probability of extinction of a consumer is constant (π_i) unless it loses all resources, in which case the probability of extinction is one, (ii) the linear case in which the probability of extinction increases linearly with the fraction

of resources lost, and (iii) several nonlinear cases in which the extinction probability increases nonlinearly with resource loss.

Specifically, suppose that a consumer feeds on n resources (X_1, \dots, X_n). In all cases, when all resources are extant, we have $P(\neg i|X_1, \dots, X_n) = \pi_i$, and for all resources extinct, $P(\neg i|\neg X_1, \dots, \neg X_n) = 1$. For intermediate cases, suppose that the consumer lost a fraction $k/n = f$ of its resources. In the topological case, $P(\neg i|f) = \pi_i$ for all $0 \leq f < 1$ and $P(\neg i|f) = 1$ for $f = 1$. In the linear case, increasing f increases the extinction probability linearly $P(\neg i|f) = \pi_i + (1 - \pi_i) \cdot f$. For the nonlinear case, we want a flexible curve that can take different shapes depending on the parameterization. For this exercise, we use the cumulative density function of a beta distribution:

$$P(\neg i|f) = \pi_i + (1 - \pi_i) \frac{\mathcal{B}(f; \alpha, \beta)}{\mathcal{B}(\alpha, \beta)} \quad \text{eqn 1}$$

where $\mathcal{B}(f; \alpha, \beta)$ is an incomplete beta function. Using eqn 1, the response to extinctions can be linear ($\alpha = \beta = 1$), sigmoid ($\alpha > 1, \beta > 1$), inverse sigmoid ($\alpha < 1, \beta < 1$), concave ($\alpha \leq 1, \beta > 1$) or convex ($\alpha > 1, \beta \leq 1$) (Fig. 2). In Fig. 1, we contrast the tables of conditional probabilities for a simple food web in the topological and linear cases.

Note that by considering $f = k/n$, we make the strong assumption that the probability of extinction of a consumer depends only on the number of its extinct resources (k) and not on their identities. Thus, losing k resources has the same effect, regardless of their identities. Additionally, the extinction probability only depends on the resources, whereas changes in consumer densities, top-down predation rates and other network modifications are assumed to be unimportant. We call this case 'binary', as only the state of each resource (extant/extinct) influences the extinction probability of the consumer. However, if we know that one resource is more important than the others for a consumer, we can extend the analysis to account for their relative contributions. Suppose that the relative importance of a resource is θ_{ji} (e.g. the proportion of biomass flowing from resource j to consumer i per unit time compared to other resources). Then, we can define the fraction of resources lost $f_i = \sum_j \text{lost } \theta_{ji} / \sum_j \theta_{ji}$, where the first sum is taken over all the resources that are not present (j lost), while the second sum is taken over all resources of i . Thus, f_i is the fraction of resources lost weighted for their importance. We denote this case as 'flow', as opposed to 'binary'.

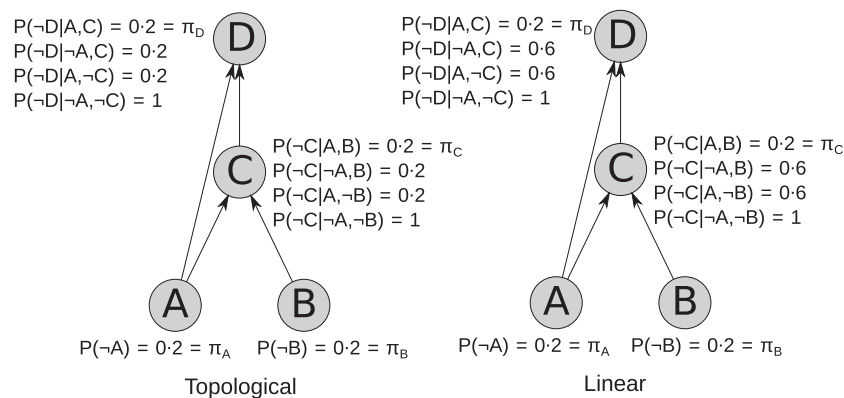


Fig. 1. A simple Bayesian network. For a simple food web, we write the conditional probability of extinction of a consumer given the state of its resources. Left: topological approach, in which the probability of extinction is the same unless all resources are lost. Right: linear binary case, in which losing resources increase the probability of extinction linearly. For example, the risk of going extinct for species C is in the topological case always 0.2 unless both A and B are lost, then the risk is 1. In the linear binary case, the risk of extinction of C depends on the state (extant/extinct) of its prey. The probability for C of going extinct is $\pi + (1 - \pi) \cdot \text{number of preys}$, in this case $0.2 + (1 - 0.2) \cdot 1/2 = 0.6$.

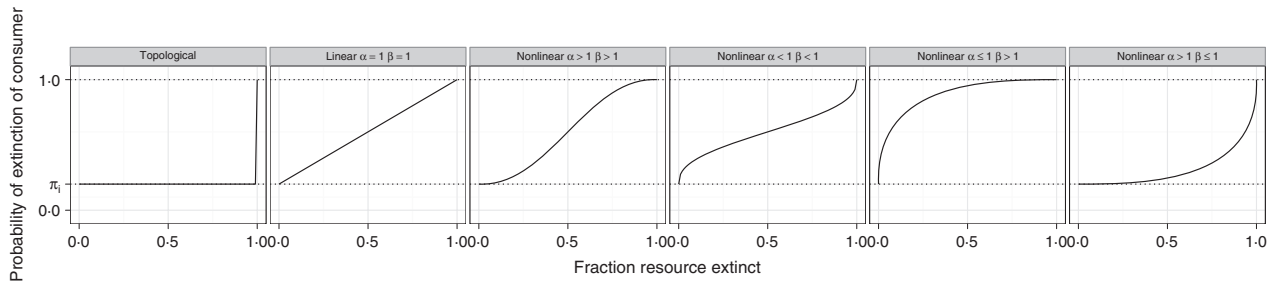


Fig. 2. Schematic description of different functional forms of a consumer's response to loss of resources. From left to right: topological, where a consumer's probability of extinction is unaffected until all resources are lost; linear, where the probability of extinction grows linearly with the fraction of resources lost; nonlinear, taking different shapes according to the parameters α and β .

APPROXIMATING THE MARGINAL PROBABILITIES

We have built a table for each species specifying its probability of extinction given the state of its resources. This defines a Bayesian network. However, the resources in turn might depend on other species in the system. To compute the probability of extinction of a species i when taking into account of all the possible states of all species, we need to combine all the tables. This is what we call 'solving' the Bayesian network. In this section, we present a naive way of doing so using simulations. In the next section, we show that there is a more accurate and efficient way of performing this task.

Defining the conditional probabilities is sufficient to perform numerical simulations, which can be used to approximate the probability p_i of extinction of a species i in the system [as shown by Lafferty & Dunne (2010)]. In practice, if one knows the state of the resources of a consumer i , it is sufficient to look up the probability of extinction of i given the state of its resources and perform a Bernoulli trial to determine whether i will go extinct. For example, take the simple network in Fig. 1 and the linear case. To test whether species A goes extinct (with probability 0.2), one would draw a random number r from a uniform distribution $\mathcal{U}[0, 1]$ and A would be considered extinct if $r < 0.2$. Assume that A is extant. Then one would test for the extinction of B (probability of extinction = 0.2). Assume that B is extinct. Then, to test whether C is extinct, one would look up the probability $P(-C|A, -B) = 0.6$ and use this value to test whether C is extinct.

In general, we need to sort the species such that, when we consider an individual species, the state of its resources has been determined: we can start from primary producers (who do not depend on other species) and check whether they go extinct. Then, we move to primary consumers (who depend only on primary producers) and determine whether they go extinct. We continue with the species that depend only on producers and primary consumers (secondary consumers and omnivores) and so on. However, a complication arises when cycles (a path that originates from one species and also returns to it) are present – cycles make it impossible to find such an order. In fact, the order can be found in linear time, using, for example, a topological sorting routine (for a detailed description of the method, see Tarjan 1972; Allesina, Bodini & Bondavalli 2005), only if the network is acyclic.

Fortunately, Allesina, Bodini & Pascual (2009) proved that not all connections in a food web contribute to robustness and that removing 'redundant' links leaves a food web acyclic. The method is obtained by finding the set of species that collectively dominate a certain consumer: all the pathways between resources and the consumer include nodes in the set. Then, one can remove certain connections within the nodes in the set, rendering the network acyclic Allesina *et al.* (2009). Interestingly, cyclic and acyclic versions of each network have the same properties in terms of secondary

extinctions. Thus, the method proposed here can be applied to any network, cyclic or acyclic.

A sketch of the simulation algorithm is as follows. As input, we need a network N composed of S species and L interactions. We remove cycles if necessary following Allesina *et al.* (2009). We also need an estimate of the baseline probabilities of extinction of all species $\bar{\pi}$ and a functional response for the consumers (topological, linear, nonlinear with parameters α and β). We can use a binary adjacency matrix A ('binary' case) or a weighted network A' ('flow' case).

First, we order the species using a topological sorting routine (Tarjan 1972; Allesina *et al.* 2005), that is, we find an order for the nodes such that all arrows point from left to right. Then, we go through each species i following the order. We compute f (or f' for the weighted case) for species i and use the value to compute $P(-i|f)$. We draw a random number from a uniform distribution $\mathcal{U}[0, 1]$; if the number is lower than $P(-i|f)$, we mark the species as extinct. This implies random extinction probabilities and that the extinction probabilities are independent for each species. Once we perform this operation on all species, we end up with a reduced network in which some species are extinct because of 'primary' extinctions (i.e. species whose resources are all present), while others are extinct in response to primary extinctions (partial or total loss of resources). Clearly, repeating the experiment many times would yield different results, as the method is probabilistic.

Suppose we repeat the extinction experiment above many times, we can then estimate the probability that species i goes extinct by computing the fraction of simulations in which i is not present in the final state. This is an approximation of the marginal probability of extinction of i , p_i . This quantity is a marginal probability since it estimates the probability of species i going extinct by integrating over all possible states of the resources of i weighted for the corresponding probability. Thus, the vector \bar{p} is the solution of the Bayesian network.

EXACT MARGINAL PROBABILITIES

The simulation approach shown above approximates the marginal probabilities, but to estimate correctly rare events, we would have to perform many simulations. An alternative is to compute the marginal probabilities analytically.

Consider a complete acyclic network composed of S species and $S(S-1)/2$ connections. In this network, once we sort the nodes (species), the probability of extinction of each species depends on the state of all the preceding species. Thus, to compute analytically the marginal probability for all species, we should consider two possible cases for the first species (extant/extinct), four cases for the second species (extant/extinct for each state of the first species), eight for

the third and so forth. This would make the exact solution of the marginal probabilities computationally unattainable for large webs, as one would need to evaluate 2^S states. Fortunately, food webs are not completely connected, and efficient algorithms exist to solve this problem in sparse networks. In fact, there is an extensive body of literature on the solution of Bayesian networks. Efficient algorithms, such as the junction tree algorithm (Jensen, 1996), can be used to rapidly compute the exact marginal probabilities. Because R (R Development Core Team 2010) is a popular choice among ecologists, we used the package *gRain* (Hjsgaard, 2011) for solving our Bayesian networks.

A MEASURE OF ROBUSTNESS USING BAYESIAN NETWORKS

The baseline probability of extinction π_i quantifies the probability that species i goes extinct for causes other than those represented by the network. The marginal probability $p_i \geq \pi_i$, on the other hand, quantifies the effect of the network on the probability of extinction of i . The ratio p_i/π_i tells us how much more probable the extinction of i is when the species is embedded in the network. For primary producers, $p_i = \pi_i$, while the ratio is expected to grow with trophic level.

The expected number of extant species before considering the network is $\mathbb{E}_b = S - \sum \pi_i$. The expected number of species once the network has been accounted for is $\mathbb{E}_a = S - \sum p_i$. Hence, we can measure the robustness of the network as $\mathcal{R} = \mathbb{E}_a/\mathbb{E}_b$. Naturally, \mathcal{R} ranges between 0 and 1 and depends on (i) network structure, (ii) the baseline probabilities π and (iii) the choice of functional response to resource loss.

TESTING THE METHOD

To test our method, we took a standard model for simulating the dynamics of large ecological networks. The Allometric Trophic Network (ATN) model (Yodzis & Innes 1992; Brose, Williams & Martinez 2006; Berlow *et al.* 2009) can be used to build persistent networks parameterized using allometric relationships between body size, metabolic rate and other ecologically relevant quantities. The model is commonly used in studies on secondary extinctions (Berlow *et al.* 2009; Binzer *et al.* 2011; Curtsdotter *et al.* 2011; Riede *et al.* 2011). In our experiments, we followed closely the parameterization presented in Binzer *et al.* (2011).

We built 100 networks composed of 45 species each and an expected connectance of 0.15. We numerically integrated these webs until they reached a point equilibrium or a periodic attractor. To determine whether such a state had been reached, we averaged the biomass of each species over 500 time steps and compared it with the average taken over the previous 500 steps. If no species grew or shrank more than 1%, we stopped the integration. Otherwise, we extended the integration and performed the test again. If the web did not reach a steady or periodic state in 5000 time steps, we rejected the network and started over. We checked how many species were present at the end of the integration. If the final number of species < 30 , we rejected the network. We repeated the previous operations until we collected 100 persistent webs, each composed of at least 30 species.

We chose the values in π in two ways. First, we assigned all species the same baseline probability $\pi_i = 0.2$. Second, we increased the probability of extinction with trophic level of the species. This is consistent with what is observed in natural systems (Purvis *et al.* 2000; Raffaelli 2004), since trophic level correlates with traits associated with high extinction risks (large body size, low reproduction rate and large home

ranges). This approach is also frequently used in models, for example, when simulating removal sequences (Curtsdotter *et al.* 2011). Denoting the trophic level of species i by TL_i and the average trophic level taken across all species by \overline{TL} , we computed the baseline probability as $\pi_i = 0.2 \cdot TL_i/\overline{TL}$. Although theoretically we could obtain a value $\pi_i > 1$, we did not observe this case in our simulations. In both treatments of π_i ('Uniform' and 'Trophic Level Based'), on average 20% of species are expected to go extinct primarily.

We then proceeded to simulate extinction scenarios on the network using the ATN model. This was done in order to provide reference extinction scenarios with which to compare Bayesian network models. First, we went through each species i and set it to an extinct state with probability π_i . Then, we ran the dynamical model for 5000 time steps to allow for secondary extinctions to unfold. Finally, we recorded the end state of each species in the vector $\vec{\sigma}$. We repeated the procedure 250 times and stored the results of the k^{th} replicate in a vector $\vec{\sigma}_k$.

LIKELIHOODS AND MAXIMUM LIKELIHOOD

For each network and replicate, we can compute the likelihood that one of the algorithms tested replicates exactly the ATN-simulated extinctions (Fig. 2). The likelihood therefore measures the probability that, for a given network, algorithm and parameterization, a Bayesian network would produce exactly the same results observed in the ATN model. The equation is simply:

$$L(\sigma|A, \pi, \text{algorithm}) = \prod_{k=1}^{250} \left(\prod_{i=1}^S p_i^{(1-\sigma_{ki})} (1-p_i)^{\sigma_{ki}} \right) \quad \text{eqn 2}$$

where σ_{ki} is the observed final state of species i in replicate k , p_i is the marginal probability of extinction computed for the Bayesian network, A is the adjacency matrix (or the flow case matrix A'), and π is the vector of baseline extinction probabilities.

We tested five algorithms/responses: (i) topological binary, where we perform topological extinctions and treat each resource as equally important; (ii) linear binary, in which we apply the linear response described above; (iii) linear flow, where we consider the different importance of resources (network A'); (iv) nonlinear binary; and (v) nonlinear flow. In the latter two cases, we set the parameters describing the response to their maximum-likelihood estimate $\hat{\alpha}$ and $\hat{\beta}$, found using the Nelder–Mead algorithm.

To provide a baseline case, we contrast the results with those obtained by setting each marginal probability to its maximum-likelihood estimate (MLE):

$$\hat{p}_i = \frac{\sum_{k=1}^{250} \sigma_{ki}}{250} \quad \text{eqn 3}$$

The MLE is thus obtained from the simulations of the ATN model and represents the best case possible.

All codes are available upon request.

Results

From the results of the simulations of the dynamical model, we can identify each species extinction with one of three types: (1) those species with all resources extant, (2) those with some (but not all) resources extant and (3) those with no extant resources. Clearly, types 2 and 3 can be potentially detected using our method, while type 1 cannot, as the consumer goes extinct

despite all its resources being extant. Extinctions of type 1 account for up to 22% of the secondary extinctions when we used a uniform π_i ($\pi_i = 0.2$) and up to 44% of secondary extinctions when we modelled π_i as a function of the trophic level of i (Fig. 3).

We tested five different algorithms for calculating a species marginal probability of extinction, each of which is described by two terms. The first term (topological, linear, nonlinear) describes the consumer's functional response to resource loss. Note that the nonlinear response could assume the shape of both the topological and the linear response, if this were to lead to the best likelihood. The second term (binary, flow), describes the quantity used to compute the fraction of resource lost. In the binary case, each resource is equally important, while in the

flow case, the importance of a resource depends on its contribution to the diet of the consumer. Thus, the five algorithms are as follows: (i) topological (as binary and flow are exactly the same in this case), (ii) linear binary, (iii) linear flow, (iv) nonlinear binary and (v) nonlinear flow. For each network, we computed the likelihood that a given algorithm exactly reproduced the extinctions observed in the ATN simulations as well as the maximum likelihood. We repeated the whole procedure for the case in which π_i is equal for all species ('Uniform') as well for the case in which it depends on the trophic level of i ('Trophic Level Based').

We found a clear ranking for the performance of the five algorithms. The best performing algorithm was always nonlinear, with more networks favouring the use of flow-based

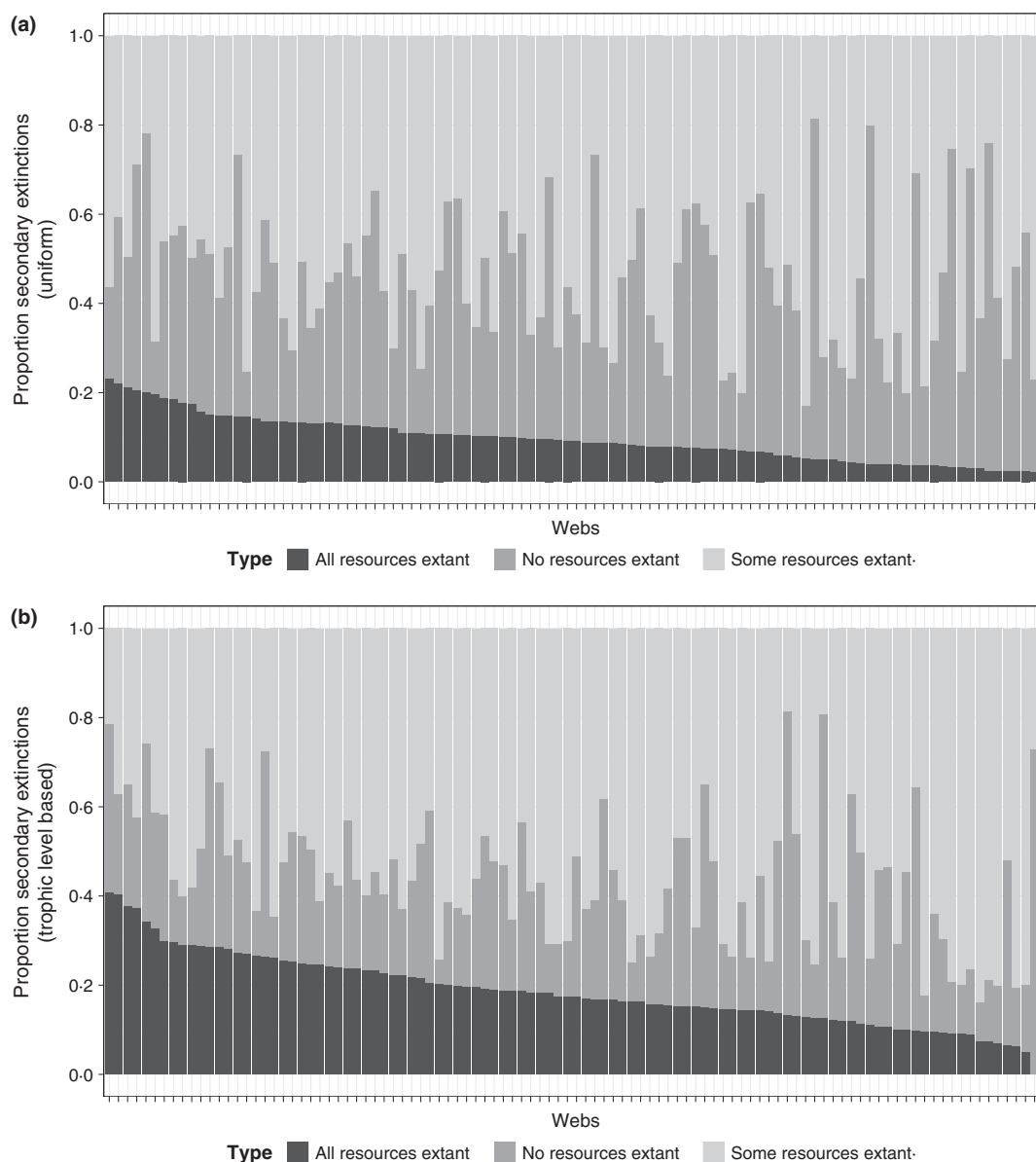


Fig. 3. Types of secondary extinctions occurring in the simulations. Each secondary extinction of a consumer can be classified according to the state of its resources: dark grey – all the resources are extant; grey – all resources are extinct; light grey – some, but not all, resources are extant. (a) proportion of each type of extinction when the baseline probability of extinction π_i is equal for all species. (b) proportions when π_i grows with the trophic level of species i . 100 webs with 250 replicates of each was used.

interactions rather than a binary description of the interactions for the resources (Fig. 4). Although these models necessarily fail to predict some of the extinctions, they still produce results close to the maximum-likelihood outcome for all networks (Fig. 5).

For the nonlinear functional forms, we searched for the maximum-likelihood parameters that best fit the observations. As such, we can plot the resulting maximum-likelihood parameter values to determine which shapes are favoured by the data. We start examining the two nonlinear binary cases, obtained for uniform and trophic level-based π_i (Fig. 6). In both cases, the most favoured shape is a sigmoid curve ($\alpha > 1$, $\beta > 1$, red), with the steepest increase in the probability of extinction (mean of the distribution) being around 50% of resource lost – the points are close to the $\alpha = \beta$ line. In many cases, the transition is quite sharp ($\alpha \gg 1$, $\beta \gg 1$, in red). A few networks favour the use of a quasi-topological response ($\alpha \ll 1$, $\beta \ll 1$, in which the probability of extinction is constant until almost all the resources are lost, in green), and a few other networks yield almost linear responses ($\alpha \approx 1$, $\beta \approx 1$). Very few cases yield convex functions ($\alpha > 1$, $\beta \leq 1$, in blue), and all but two are close to the linear case. In the nonlinear flow models (Fig. 6), we observe two main differences: (i) the transitions tend to be very sharp ($\alpha \gg 1$, $\beta \gg 1$), and (ii) there are severe deviations from the line $\alpha = \beta$, meaning that the transition does not happen when about 50% of the flow is lost, but rather at higher levels of loss.

Because for each network and replicate we ran two simulations modifying solely the π_i (using ‘Uniform’ or ‘Trophic Level Based’ approaches), we compare the robustness of each network for the two treatments (Fig. 7). In all cases, the Trophic Level-Based treatment makes the network more robust: preferentially, removing species from the top produces fewer secondary extinctions than those obtained when all species have the same probability of going primarily extinct. Compared to the results obtained using the ATN model, the algorithms based on linear response tend to underestimate robustness, while the topological approaches tend to grossly overestimate it. The two nonlinear cases closely reproduce the ATN results, consistent with the results found using likelihoods. Although the nonlinear cases are more flexible in fitting the data than the topological and linear cases, the difference in likelihoods is typically so large that any model selection technique would prefer the best performing model judged using likelihoods alone (Fig. 5).

Discussion

Traditionally, two main approaches have been taken to study food web robustness: the extremely simplistic topological approach and the considerably more complicated fully dynamical one. Here we applied a Bayesian network approach to a third framework that combines the simplicity of the topological approach with some important features

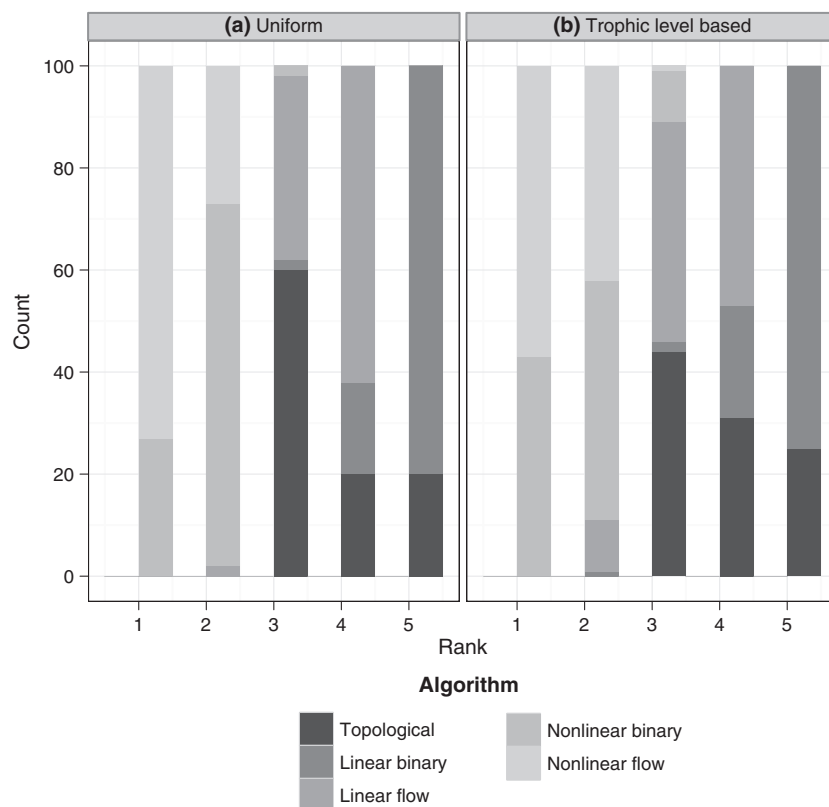


Fig. 4. Ranking of the algorithms. (a) ranking of each algorithm when the baseline probability of extinction is uniform. For the 100 simulated networks, we report how many times each algorithm is the best performing (Rank = 1), second best performing (Rank = 2) and so on. (b) ranking when the baseline probability of extinction depends on the trophic level of each species.

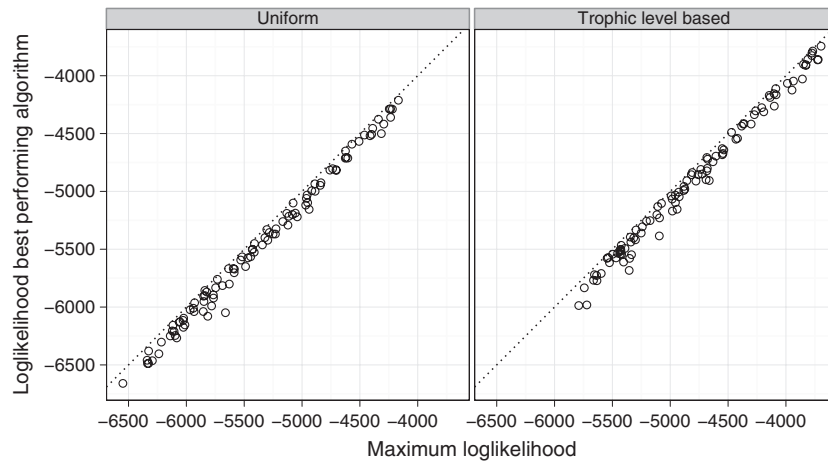


Fig. 5. Likelihood of the best algorithm. For the two cases of baseline probability of extinction, we plot the likelihood of each algorithm reproducing exactly the pattern of extinctions vs. the maximum attainable likelihood.

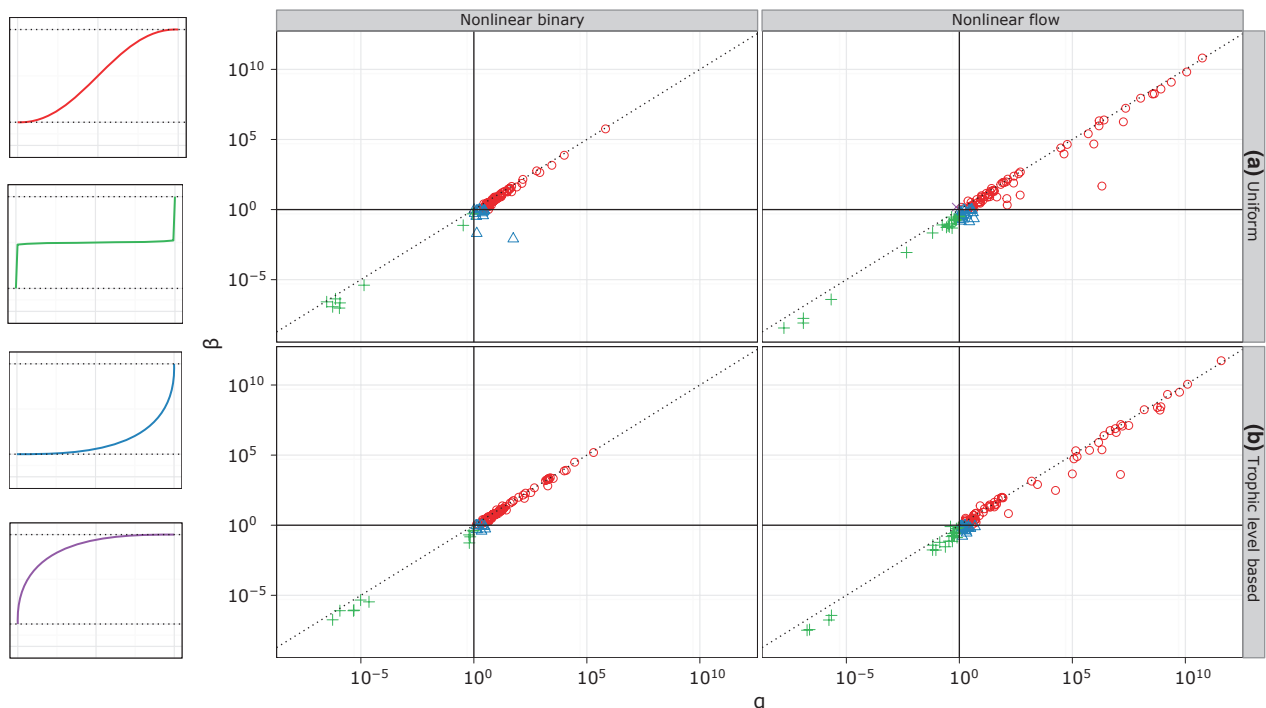


Fig. 6. Choice of functional form. α and β describe the nonlinear functional form of the consumer response to loss of resources. Red circles correspond to a sigmoid curve with $\alpha > 1$ and $\beta > 1$. Blue triangles to a convex function with $\alpha > 1$ and $\beta < 1$. Purple crosses correspond to a concave function, $\alpha < 1$ and $\beta \geq 1$. Green plus signs correspond to a quasi-topological response, $\alpha \ll 1$, $\beta \ll 1$. For illustration, the shape of the functional forms corresponding to the respective combinations of α and β is presented on the left-hand side of the figure.

typical of dynamical models, but without the extensive set of parameters needed to track population dynamics. Using Bayesian networks, when provided with a network structure and a baseline probability of extinction for each species, one can calculate the marginal probability of extinction of each species exactly.

In topological models, the only input parameter is network structure, and a species goes extinct whenever it loses all of its resources (Dunne *et al.* 2002). The probability of extinctions does not increase with the fraction of resources lost – a

consumer with only a small fraction of its resources remaining is as viable as one that has suffered no loss. This assumption has repeatedly been criticized as unrealistic and contrasts starkly with the biological understanding of extinction risk (Purvis *et al.* 2000; Eklöf & Ebenman 2006; Srinivasan *et al.* 2007). In our Bayesian network method, this issue is addressed by explicitly modelling the functional form of the response to loss of resources, that is, how the probability of extinction of a consumer increases with increasing loss of resources. We tested three functional forms – topological, linear and nonlinear.

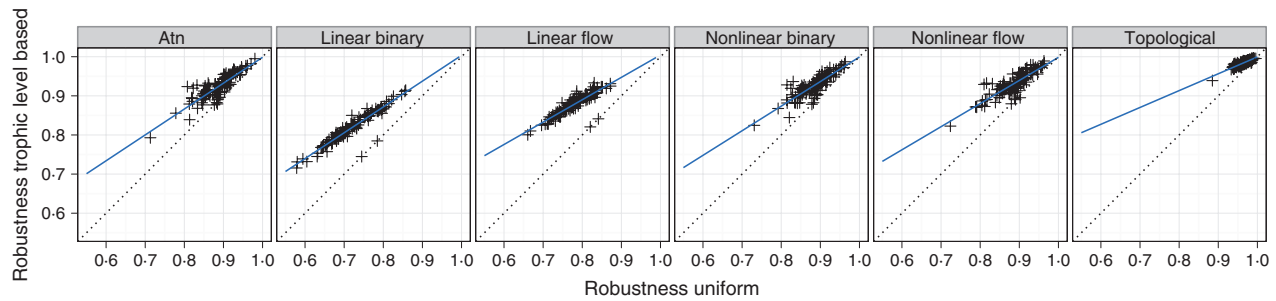


Fig. 7. Robustness and choice of baseline probability of extinction. For the simulations (ATN) and the five algorithms, we plot the robustness obtained when the π_i is equal for all species (x -axis) vs. the robustness obtained for the case in which the π_i depends on the trophic level of i .

In analysis of food web robustness, it is often assumed that all nodes are equal – each species has the same probability of undergoing extinction or the probability is solely determined by network properties (Dunne *et al.* 2002). Although the simplicity of the approach is appealing, this extreme simplification is in contrast to empirical results. One of the strengths of our Bayesian network method is therefore the possibility of including more precise ecological knowledge about species extinction risks. Each and every species is assigned a baseline probability of going extinct, which is taken into account in the likelihood evaluation, and the variation in extinction among species does not need to follow a sequence. This is useful since we know that some species are more vulnerable than others, regardless of their pattern of interactions, large-bodied species, species depending strongly on particularly habitats, overexploited or rare species all likely to go extinct for causes other than network structure (Purvis *et al.* 2000). Using the method presented here, this information can be taken into account, for example, by assigning the baseline probability of extinction using lists of endangered species. For practical applications, there is a need to bridge the gap between theoretical ecology and conservation biology, and including results from conservation-oriented research into algorithms for the analysis of networks is a first step in this direction.

We compared our results from the Bayesian model with the results obtained from the widely used ATN model (Berlow *et al.* 2009; Binzer *et al.* 2011, for example) and assumed that extinctions in the ATN model were ‘true’ extinctions. There are two main reasons for this choice. First, experimental data in which replicate extinction experiments in large networks are performed through manipulation are completely lacking – although sorely needed. Second, we wanted to contrast our result using the results from a commonly used dynamical model that included several ecologically relevant parameters. The ATN model meets this criterion. We find that the Bayesian network approach predicts the majority of the secondary extinctions that the ATN model generates. However, secondary extinctions of consumers whose resources are all extant cannot be predicted using our approach. These extinctions can, in the dynamical model, be caused by purely top-down interactions, for example, disruption of predator-mediated coexistence, or still be bottom-up effects in which resources decrease to a level insufficient to

support the consumer (Ebenman & Jonsson 2005). In such cases, the network structure itself would not change (the resources would be diminished, but would not become extinct), and thus, our method would predict no increase in the extinction risk. Our simulations show that these extinctions are quite rare (20%) when all species have the same probability of going primarily extinct, but more abundant (44%) when species at the top of the food web are more likely to go extinct than those at lower trophic levels (Fig. 3). Because in natural ecosystems apex predators and large-bodied species have high probability of extinction due to anthropogenic effects, this result highlights a severe limitation of purely network-based studies of robustness. Thus, when we assume that all species are equally likely to go extinct, we tend to grossly underestimate the robustness of the system (Fig. 7), while using a purely topological algorithm, we would encounter the opposite problem (Fig. 7).

An important advantage of the Bayesian network approach is that it does not require simulated extinctions. The most common methodology for analysing food web robustness has then been to simulate primary species removals – sequentially remove species, either at random or based on some topological properties (Dunne *et al.* 2002; Srinivasan *et al.* 2007) and record the number of secondary extinctions following each removal. Whereas in the topological approach each primary extinction can have only one outcome [unless the consumers have some degree of adaptability and can form new interactions if resources are lost, see Staniczenko *et al.* (2010) and Thierry *et al.* (2011)], the outcome of dynamical models depends on the initial conditions of the model. Therefore, there is a need for numerous replicates for each removal (Eklöf & Ebenman 2006; Curtsdotter *et al.* 2011; Riede *et al.* 2011). This makes the simulations computationally intensive and restricts the choice of network sizes. The Bayesian network algorithm, on the other hand, takes into account all the possible outcomes along with their probabilities, without the need to compute them all. This is promising given that data on large ecological networks are appearing more frequently in the literature (Dunne, 2006).

On the other hand, Bayesian networks have some disadvantages. The first is that top-down effects cannot be implemented in a Bayesian network. If these are important in nature, possible alternatives are dynamical models or the SENO model

(Lafferty & Dunne 2010). In addition, simulations produce a large number of possible scenarios that might occur in nature, while Bayesian networks produce the means of these scenarios. For some questions, a range of scenarios is of interest, such as the identification of alternative states or expected correlations among species within a community. Of course, these can be simulated sampling outcomes from Bayesian networks as well.

In order to further incorporate biological realism in Bayesian network approaches, we include information on the energy flow between each consumer–resource pair. These flows weight the importance of the links, giving a higher importance to interactions between two species with higher energy flow, an approach often taken in ecosystem studies (Banašek Richter *et al.*, 2009; Ulanowicz 2009). This is a possible way to move beyond purely topological structures and start accounting for interaction strengths. In addition to flows, network analysis should take into account how resource needs change from one life stage to another. This has been successfully simulated in past studies (Rudolf & Lafferty 2011; Lafferty 2012), and Bayesian networks can easily incorporate this through specification of the marginal probabilities. However, whenever life stages introduce cycles in the network, one would need to use approaches that can deal with ‘loopy Bayesian networks’ (Jensen & Nielsen 2007).

The earliest topological studies (Dunne *et al.* 2002) already stressed that the lack of a mechanism for consumers to adapt to loss of resources by ‘rewiring’ to other resource species could potentially lead to overestimating the number of secondary extinctions following the removal of single species. This mechanism has now been investigated (Staniczenko *et al.* 2010; Thierry *et al.* 2011), and it has been shown that rewiring increases food web robustness. Although rewiring is not included in the present exercise, it could be included in our method, for example, by modelling the conditional probability of rewiring.

In conclusion, our study shows that a Bayesian network approach can capture the majority of the secondary extinctions in food webs without the need for tracking population dynamics. Our hope is that this method will be a useful complement to existing tools for analysing the robustness of food webs and other ecological networks, reducing the gap between food web theory and conservation-oriented research.

Acknowledgements

We thank S. Pawar and L. Dee for valuable comments, K. Lafferty and P. Staniczenko for insightful suggestions and comments. AE and ST are supported by NSF EF-0827493 and SA by NSF DEB-1148867. AE was additionally supported by SOEB.

References

Albert, R., Jeong, H. & Barabási, A.L. (2000) Error and attack tolerance of complex networks. *Nature*, **406**, 378–382.

Allesina, S., Bodini, A. & Bondavalli, C. (2005) Ecological subsystems via graph theory: the role of strongly connected components. *Oikos*, **110**, 164–176.

Allesina, S., Bodini, A. & Pascual, M. (2009) Functional links and robustness in food webs. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **364**, 1701–1709.

Banašek-Richter, C., Bersier, L.F., Cattin, M.F., Baltensperger, R., Gabriel, J.P., Merz, Y. *et al.* (2009) Complexity in quantitative food webs. *Ecology*, **90**, 1470–1477.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B. *et al.* (2011) Has the earth’s sixth mass extinction already arrived? *Nature*, **471**, 51–57.

Berlow, E.L., Dunne, J.A., Martinez, N.D., Stark, P.B., Williams, R.J. & Brose, U. (2009) Simple prediction of interaction strengths in complex food webs. *Proceedings of the National Academy of Sciences*, **106**, 187–191.

Binzer, A., Brose, U., Curtsdotter, A., Eklöf, A., Rall, B.C., Riede, J.O. *et al.* (2011) The susceptibility of species to extinctions in model communities. *Basic and Applied Ecology*, **12**, 590–599.

Brose, U., Williams, R.J. & Martinez, N.D. (2006) Allometric scaling enhances stability in complex food webs. *Ecology Letters*, **9**, 1228–1236.

Cardillo, M., Mace, G.M., Jones, K.E., Bielby, J., Bininda-Emonds, O.R.P., Sechrest, W. *et al.* (2005) Multiple causes of high extinction risk in large mammal species. *Science*, **309**, 1239–1241.

Curtsdotter, A., Binzer, A., Brose, U., de Castro, F., Ebenman, B., Eklöf, A. *et al.* (2011) Robustness to secondary extinctions: Comparing trait based sequential deletions in static and dynamic food webs. *Basic and Applied Ecology*, **12**, 571–580.

Davidson, A.D., Hamilton, M.J., Boyer, A.G., Brown, J.H. & Ceballos, G. (2009) Multiple ecological pathways to extinction in mammals. *Proceedings of the National Academy of Sciences*, **106**, 10702–10705.

Dunne, J.A. (2006) The network structure of food webs. *Ecological Networks: Linking Structure to Dynamics in Food Webs* (eds M. Pascual & J.A. Dunne), pp. 27–86. Santa Fe Institute Studies in the Sciences of Complexity. Oxford University Press, Oxford.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.

Ebenman, B. & Jonsson, T. (2005) Using community viability analysis to identify fragile systems and keystone species. *Trends in Ecology and Evolution*, **20**, 568–575.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community viability analysis: the response of ecological communities to species loss. *Ecology*, **85**, 2591–2600.

Eklöf, A. & Ebenman, B.O. (2006) Species loss and secondary extinctions in simple and complex model communities. *Journal of Animal Ecology*, **75**, 239–246.

Gaston, K.J. & Blackburn, T.M. (1995) Birds, body size and the threat of extinction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **347**, 205–212.

Hjsgaard, S. (2011). *gRain: Graphical Independence Networks*. URL <http://CRAN.R-project.org/package=gRain>. R package version 0.8.7.

Jensen, F.V. (1996) *An Introduction to Bayesian Networks*, vol. 36. UCL press, London.

Jensen, F.V. & Nielsen, T.D. (2007) *Bayesian Networks and Decision Graphs*. Springer, New York.

Lafferty, K.D. & Dunne, J.A. (2010) Stochastic ecological network occupancy (seno) models: A new tool for modeling ecological networks across spatial scales. *Theoretical Ecology*, **3**, 123–135.

Lafferty, K.D. (2012) Biodiversity loss decreases parasite diversity: theory and pat terns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 2814–2827.

Lee, T.M. & Jetz, W. (2011) Unravelling the structure of species extinction risk for predictive conservation science. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **278**, 1329–1338.

Memmott, J., Waser, N.M. & Price, M.V. (2004) Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **271**, 2605–2611.

Pereira, H.M., Leadley, P.W., Proença, V., Alkemade, R., Scharlemann, J.P.W., Fernandez-Manjarres, J.F. *et al.* (2010) Scenarios for global biodiversity in the 21st century. *Science*, **330**, 1496–1501.

Purvis, A., Gittleman, J.L., Cowlishaw, G. & Mace, G.M. (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **267**, 1947–1952.

R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>. ISBN 3-900051-07-0.

Raffaelli, D. (2004) How extinction patterns affect ecosystems. *Science*, **306**, 1141–1142.

Reich, P.B., Tilman, D., Isbell, F., Mueller, K., Hobbie, S.E., Flynn, D.F.B. *et al.* (2012) Impacts of biodiversity loss escalate through time as redundancy fades. *Science*, **336**, 589–592.

Riede, J.O., Binzer, A., Brose, U., de Castro, F., Curtsdotter, A., Rall, B.C. *et al.* (2011) Size-based food web characteristics govern the response to species extinctions. *Basic and Applied Ecology*, **12**, 581–589.

- Rudolf, V.H.W. & Lafferty, K.D. (2011) Stage structure alters how complexity affects stability of ecological networks. *Ecology letters*, **14**, 75–79.
- Sole, R.V. & Montoya, M. (2001) Complexity and fragility in ecological networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 2039–2045.
- Srinivasan, U.T., Dunne, J.A., Harte, J. & Martinez, N.D. (2007) Response of complex food webs to realistic extinction sequences. *Ecology*, **88**, 671–682.
- Staniczenko, P.P.A., Lewis, O.T., Jones, N.S. & Reed-Tsochas, F. (2010) Structural dynamics and robustness of food webs. *Ecology Letters*, **13**, 891–899.
- Stouffer, D.B. & Bascompte, J. (2011) Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences*, **108**, 3648–3652.
- Tarjan, R. (1972) Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, **1**, 146–160.
- Thierry, A., Beckerman, A.P., Warren, P.H., Williams, R.J., Cole, A.J. & Petchey, O.L. (2011) Adaptive foraging and the rewiring of size-structured food webs following extinctions. *Basic and Applied Ecology*, **12**, 562–570.
- Ulanowicz, R.E. (2009) The dual nature of ecosystem dynamics. *Ecological Modelling*, **220**, 1886–1892.
- Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S. et al. (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science*, **314**, 787–790.
- Yodzis, P. & Innes, S. (1992) Body size and consumer-resource dynamics. *American Naturalist*, **139**, 1151–1175.

Received 5 February 2013; accepted 11 April 2013

Handling Editor: Jessica Metcalf