This is the accepted manuscript made available via CHORUS. The article has been published as:

Topology of plant-pollinator networks that are vulnerable to collapse from species extinction
Colin Campbell, Suann Yang, Katriona Shea, and Réka Albert
Phys. Rev. E 86, 021924 — Published 31 August 2012
DOI: 10.1103/PhysRevE.86.021924
The topology of plant-pollinator networks that are vulnerable to collapse from species extinction

Colin Campbell1,*, Suann Yang2, Katriona Shea3, and Réka Albert1,3

1Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA
2Department of Biology, Presbyterian College, Clinton, SC 29325, USA
3Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA
*Corresponding author

Author addresses:
Colin Campbell
104 Davey Laboratory, #40
Pennsylvania State University
University Park, PA 16802
campbell@phys.psu.edu
Phone: (330) 550-9592
Fax: (814) 865-3604

Suann Yang
503 S. Broad Street
Presbyterian College
Clinton, SC 29325
syang@presby.edu

Katriona Shea
208 Mueller Laboratory
Pennsylvania State University
University Park, PA 16802
k-shea@psu.edu

Réka Albert
122 Davey Laboratory
Pennsylvania State University
University Park, PA 16802
ralbert@phys.psu.edu

Date of submission: March 21, 2012

Abstract:
The ability to predict the collapse of ecological communities is of significant concern in light of global patterns of rapid species extinctions. Here, we use a recently developed dynamic Boolean network-based model of mutualistic plant-pollinator community formation to investigate the stability of simulated ecological communities in the face of sequential species extinctions. We assess communities in terms of the relative change in biodiversity after species loss, and find that communities that experience a significant
loss of biodiversity differ from more robust communities according to a number of topological characteristics. Notably, we show that high nestedness, a property commonly believed to promote community stability, may in extreme circumstances promote a critical over-reliance on individual species. Furthermore, the species important to the survival of the rest of the ecosystem occupy different positions in the network than less important species. Our results suggest that network measures may be applied to real ecosystems to yield insight into both their stability and the identity of potentially critical species.

PACS indexing codes:
87.23.-n Ecology and evolution
02.10.Ox Graph theory,
87.18.-h Biological complexity,
02.70.-c Computational techniques; simulations

1. INTRODUCTION

Due to the interdependence of species in an ecological community, the loss of one species may result in a cascade of secondary extinctions across the community. At the extreme, communities may collapse in response to the loss of a single, critical species [e.g., sea otters, 1]. Understanding the dynamics of cascading extinctions is of great current interest, given the multiple and synergistic threats to biodiversity [2]. There is an increasing need to be able to forecast community collapse, especially for plant-pollinator communities, as losses of populations of the highly generalist honeybee pose a major threat to the diversity and affordability of our agricultural crops [3, 4]. Worldwide losses of populations of the honeybee, as well as declines in native pollinators, put the plants they pollinate at risk [5], and thereby endanger the remaining pollinator species in turn.

Ecological network models, which represent species as nodes and their interactions as edges, have emerged as a useful tool to explore the properties of communities that impart stability. Certain structural properties of nodes and networks may promote robustness to species loss. Communities that are most vulnerable to species extinctions appear to be those with lower density, because they exhibit less redundancy in interactions that link species [e.g., 6]. Communities with hierarchical structure in the interdependence of species (nestedness) are predicted to be robust to species extinction [e.g., 7]. Clustering, the grouping together of many species based on their interactions, also promotes robustness [e.g., 8, 9-11]. Species that are thought to be the most critical in maintaining stability are generalists, i.e. those with many edges [e.g., 6, 7, 12]. While poorly connected species are expected to be at a greater risk of extinction [13-15], the number of edges a species has does not always reflect its risk of extinction or the vulnerability of the community to the extinction of that species. For instance, the loss of species that serve as bridges between clusters, even if not necessarily high degree species, could result in considerable structural changes in the network [16, 17].

Most network simulations of extinction use a topological, static approach; although this is a simple and useful way to predict a post-extinction community, a dynamic approach [e.g., 8, 18, 19] may be preferable in that it is less likely to underestimate risk [20]. To obtain more accurate predictions, models should account for
realistic changes in the community, such as species invasions or species interdependence. The importance of an extinct species for the community’s stability may depend on the extinction-invasion processes of species turnover. For example, the loss of a generalist species may be even more critical if a new, functionally redundant, generalist species is not able to invade the community. Furthermore, dynamic approaches allow for predictions due to indirect, as well as direct, effects [20, 21].

Here, we use a newly developed dynamic network model [22] to study how the loss of species impacts the stability of an ecological community consisting of plants and their pollinators. The model allows for a full dynamical analysis that completely describes the subsequent behavior of the system after a permanent species removal. This model considers the ecosystem as a network of interacting species, with the effect of an interaction determined by the physical characteristics of the species and categorized as (a) mutually beneficial or (b) beneficial for one species and detrimental to the other (Fig. 1(a)). A discrete-time, Boolean dynamic framework results in a complete map of the dynamical trajectories possible in the system (Fig. 1(b)). We use this dynamic framework to simulate repeated permanent species extinctions (Fig. 1(c)). We link the underlying dynamical processes to the topological properties of the networks, and show that networks containing single species whose extinction leads to a significant decrease in biodiversity often have properties that differ from those of more robust communities. We apply a newly developed topological measure [23] that quantifies (a) the influence of a single species upon the rest of the interaction network and (b) the influence of the rest of the network on a single species in terms of the number, length and sign (positive or negative) of paths, i.e. successive steps along edges, between pairs of nodes. We show that species whose loss leads to total collapse of the community are over-expressed in a certain area of the phase space according to this measure, and under-expressed elsewhere. These results have implications for the identification and preservation of species crucial to the survival of an ecological community.

2. BACKGROUND AND METHODS

A. Model of community assembly

A network is defined by a set of interacting elements (here, plant and pollinator species) that are referred to as “nodes” in the parlance of network theory. The interactions between nodes are represented by pairs of edges between the nodes (e.g., from A to B and from B to A), which here are either positive-directed (e.g. A -> B) or negative-directed (e.g. A -> B), depending on the nature of the interaction being represented. A path refers to a traversal from one node, along edges, to subsequent nodes.

To examine the effect of species extinctions, we expanded upon our network model of plant-pollinator community formation [22]. In this model, a community assembles from a regional pool of species that repeatedly attempts to colonize the new community. The relationships among the species are classified as either beneficial or harmful based on the physical characteristics of the species (specifically, the length of the pollinator proboscis compared to the nectar depth of the plant), and are summarized in the species interaction network. As an illustrative example, we show a species interaction network for a network of 10 species in Fig. 1(a). Each species is characterized by a binary
state variable which denotes whether the abundance of the species is above or below a threshold; for simplicity we refer to these states as present and absent. Dynamically, species successfully enter the community if the net impact of existing species is above a positive threshold, and fail to enter otherwise. Similarly, existing species persist if the net influence of existing species is above the same positive threshold, and die out otherwise. Formally, the state at time $t+1$ for a species $i$, $S_i(t+1)$, is determined by the state of every other node $j$ at time $t$ and the weight of the edge between them, $E(j,i)$:

$$S_i(t+1) = \begin{cases} 1, & \text{if: } \sum_j S_j(t) E(j,i) \geq T \\ 0, & \text{otherwise} \end{cases}$$

$E(j,i) = 0$ by definition if no edge exists between species $i$ and $j$. While a beneficial interaction does more good than a non-beneficial interaction does harm, it is difficult to rigorously quantify these effects. There exists a qualitative shift in the behavior of the system between giving equal magnitude weights to the interactions and giving positive interactions a greater magnitude weight [22]. As varying the positive weight did not have a significant impact on the dynamics of the system once it was weighted above the impact of a negative interaction, here we fix the weight of a negative interaction at -1 and a positive interaction at 4, the largest value considered in [22]. We fix $T$ at 1 for all simulations, i.e. we require a net positive influence for a species to enter or persist in the ecosystem.

The dynamical process of species invasion is modeled in a discrete time, Boolean framework, and it is possible to deterministically map out every future state of the system given an initial state (i.e., the species which first attempt to colonize the new ecosystem). The network of all states and their transitions is referred to as the state transition network (Fig. 1(b)). Every trajectory eventually leads to a steady state (where no changes to the composition of the community occur) or a limit cycle (where some members of the community oscillate between below-threshold and above-threshold population). In [22] we demonstrated that even in communities corresponding to limit cycles most of the species (85-95%) are stable.

Real ecosystems that have not recently suffered a significant disturbance may be viewed in this model as existing in an attractor of the state transition network. We identify the attractors of an ensemble of 1000 interaction networks, each consisting of 50 plant and 50 pollinator species, via Monte Carlo sampling of initial conditions.

### B. Simulating species extinctions

To analyze the behavior of the communities after the loss of a species, we subject attractors to a random attack wherein an existing community member is forced to become extinct and prevented from recolonizing the community (i.e., it is rendered extinct not only in the ecosystem in question, but also from the surrounding communities that form the regional species pool). This modifies the state transition network and sends the community towards a new attractor, i.e., the community shifts in species composition until it reaches a stable outcome (Fig. 1(b)). We simulate sequential extinctions by repeating the random extinction procedure after the community has stabilized (Figs. 2(b), 2(d)). Eventually, the community will proceed to a total extinction state where every
species has either been rendered extinct or can no longer survive in the community. 1,000 networks of species interactions were used for the analysis, with over 340,000 unique extinction paths examined over the networks for sequential species extinctions. The same pool of 1,000 networks was used for singular extinctions, with over 90,000 unique forcible extinctions performed.

C. Network analysis

We analyzed a community network whose nodes correspond to the species present in the community and whose edges are the interactions among these community members. If the community corresponds to a limit cycle, all species that were present in at least one state of the limit cycle were included in a single meta-network for topological analysis. For most analyses there were no qualitative differences between steady state and limit cycle communities; thus the results from both types of communities were pooled. We categorized the effect of a focal species’ extinction in terms of the percentage of species in the community which subsequently go extinct (pe): "significant increase in biodiversity" (pe < -33%), "minor increase in biodiversity" (-33% ≤ pe < 0%), "minor decrease in biodiversity" (0% ≤ pe < 33%), "significant decrease in biodiversity" (33% ≤ pe < 66%), "major decrease in biodiversity" (66% ≤ pe < 99%) and "total collapse" (pe ≥ 99%). We denote the species whose loss leads to significant to total community collapse critical species. We characterize communities in terms of their maximal vulnerability to the extinction of one of their species, i.e. the maximal pe value obtained when considering a random cross-section of species in the community as a candidate for extinction. For simplicity we will refer to communities vulnerable to significant to total collapse (pe ≥ 33%) as vulnerable communities, and to the rest as robust communities. We analyzed and compared the topological properties of the species and communities in these different categories. For comparison, we considered null models of the community networks generated by the configuration model, which preserves both network density and degree distribution [24], and examined the Z-scores of the community networks relative to ensembles of 1000 of these null models. For the latter analysis we sampled more than 1% of the communities in a given pe range, or more than 500 communities, for each range of pe considered.

Although we consider every interaction as a pair of directed edges, we disregarded negative edges when analyzing strongly connected components and cycles in order to focus upon positive signal flow. Ignoring negative edges does not lead to loss of information as they can be readily reconstructed as having the opposite direction of all non-reciprocated positive edges. We additionally note that standard network measures such as density, closeness centrality, betweenness centrality, and nestedness do not distinguish between positive and negative edges; the new path-based measures applied in Figure 5 were designed in part to address this, and to provide a more holistic view of the structure of networks with negative edges. To restrict our analysis to ecologically relevant behavior, we did not consider communities with fewer than 10 active species.

A number of metrics designed to yield insight into the topology of the network and the role of individual nodes in the network now exists [e.g., reviewed in 24, 25]. We briefly define the metrics used here:
• A strongly connected component (SCC) is a subset of two or more nodes in a graph such that for every pair of nodes \( i, j \) in the SCC, a path exists from \( i \) to \( j \) and vice versa. In the context of an ecological community, a SCC may be viewed as a strongly connected sub-community.

• The density of a graph is calculated as the ratio between the number of edges contained in the graph and the total number of edges in a fully connected graph of the same size.

• The betweenness centrality of a node \( i \) is the fraction of shortest paths between pairs of nodes \( j \) and \( k \) \((i \neq j \neq k)\) that contain node \( i \). A high betweenness centrality indicates a node is an important mediator of signal flow.

• The closeness centrality of a node is the inverse of the average length of the shortest path from the node to all other reachable nodes. The values are normalized to a maximum theoretical value of 1 (when the node is directly connected to every other node); high values indicate information spreads rapidly from the node.

• Nestedness is a popular metric commonly applied to ecological communities of two interacting classes of species. Nestedness measures the frequency with which pairs of species in class A share interactions with species in class B. While nestedness is a broadly studied metric in ecological systems, there are several metrics which quantify nestedness in different ways. Here, we use the NODF metric proposed by Almeida-Neto and colleagues, which is intended to reduce possible bias introduced by network size [26].

• The path-based relationship measures are a set of recently published measures that accumulate the interactions between pairs of nodes as a weighted sum of the number and sign (positive vs. negative) of paths between pairs of nodes [23]. The role of a node in the topology of the network is assessed in terms of the effect of the node on every other node in the network (the node’s influence) and the effect of every other node in the network on a node (the node’s susceptibility).

3. RESULTS

On average, forcing the extinction of a single, randomly selected species leads to the subsequent extinction of 14% of the other community members (Fig. 2(a)). In only 6% of the cases does a single extinction have no further impact on the community, and in 9% of the cases the removal of one species facilitates the invasion of other species, leading to a net increase in biodiversity. The number of extinctions required for complete collapse of a stable community follows a skew-normal distribution, with at most slightly fewer than half of the total number of species in the regional species pool driven extinct before total collapse (Fig. 2(b)). While sequential extinction of about one half of the nodes in the original community is the most probable route toward total collapse (Fig. 2(d)), in 0.02% of cases, removing a single species is sufficient to lead to the collapse of the entire community. While rare, this extremum of behavior suggests the existence of a broader category of critical species, and of communities vulnerable to the extinction of a single species. We thus perform a systematic study of the effects of single extinctions on ecological community composition (see Methods). While the removal of a species with an average characteristic length generally caused more damage than a species with an
extremal characteristic length, the effect was subtle (ranging from 2.5% to 5% average species loss); we therefore turn to topological properties of the species and their communities.

Species whose removal causes a significant loss in biodiversity directly or indirectly support the beneficial interactions for many other species in the network. Due to this support, removal of the focal species causes negative interactions to dominate and most other species to go extinct. We find that in instances where a single extinction causes damage to the network, there is a negative correlation between the amount of damage caused and both the number of SCCs present in the community and the number of species present in the SCCs (Pearson correlation coefficient with equal weight for each bin = -.65, p<1E-7) (Fig. 3). Thus, highly vulnerable communities have few, lightly populated SCCs compared to robust communities; i.e. vulnerable communities have a small core of interdependent interactions. Communities with greater nestedness generally require more extinctions before total collapse occurs (Fig. 2(c)), and larger nestedness corresponds to more resilience to species loss after a single extinction, up to nestedness values near 15 (Fig. 4(a inset)). However, very high values of nestedness increase the damage done to a community after the loss of a single species, in some cases leading to the total collapse of the community (Fig. 4(a)). The network density follows a similar distribution (Fig. 4(b)).

We tested the dependence of this behavior on both density and degree distribution through a configuration null model, which preserves these properties [24]. As shown on Fig. 4(a), the nestedness Z-scores relative to this null model are positive and high, meaning that the nestedness of the community networks is multiple standard deviations higher than the nestedness of an ensemble of null models. This indicates that nestedness is dependent upon network structure separate from the degree distribution and density. In addition, the Z-scores show a trend similar to the nestedness values: there is a statistically significant negative correlation between the percent extinct (pe) values and Z-score for the three pe values below and including 0%, and positive correlation for pe>0% (weighted Pearson correlation coefficient with magnitude >.7 and p<1E-7).

Species that cause an increase (decrease) in biodiversity have lower (higher) values of both betweenness (Fig. 4(c)) and closeness centralities (Fig. 4(d)) compared to the central "minor decrease in biodiversity" category (two-sample unpaired t-test, p<1E-7 in all cases).

In Figure 5, we group all instances where less than 99% of the community is lost, and compare them to the cases of total community collapse. Species in the robust communities occupy a wider range of node influence and susceptibility values than those in the communities vulnerable to total loss in part due to the larger sample size (4.94×10⁶ c.f. 882), but the underlying distributions are not different (unpaired t-test for equal means, p>0.6 for influence and susceptibility distributions). However, the distribution of highly critical species is shifted towards positive influences (p<10⁻⁴) and negative susceptibilities (p<10⁻¹) compared to the distribution of other species, suggesting that highly critical species serve as strong providers and weak receivers of mutualistic services in the community. This result is weakened when expanding the range to accept instances where less than 99% of the community is lost (e.g. instances of at least 66% community loss); i.e. the effect is only clear in cases of extreme community collapse.
4. DISCUSSION

We used a dynamic model of plant-pollinator communities to study the effects of forced species extinctions on the persistence of the rest of community. While ecosystems are typically robust to the extinction of a small fraction of their species, some suffer significant or total collapse after the removal of only a single species. The topologies of vulnerable communities (wherein $\geq 33\%$ maximum species loss is observed after single species removal) differ from robust communities according to a number of properties. Vulnerable communities have higher average betweenness and closeness centralities, higher density and nestedness, and fewer strongly connected components than robust communities. Overall, these results suggest that vulnerable communities feature a highly interdependent core group of species that acts as a mediator of mutualistic services provided in the community.

Critical species, that is, those species whose removal cause significant to total community collapse, tend to have very high betweenness and closeness centralities compared to other species. These features paint critical species as key mediators of mutualistic interactions in the community. This view is corroborated by the application of the path-based measures of [23] on critical species in the extreme cases of total community collapse. These highly critical species tend to provide more benefit (influence) to other species in the network than other species in the community, but are more likely to have negative susceptibilities; i.e. while these highly critical species benefit many of their partners, their partners do not strongly benefit the critical species themselves. The distribution of influences and susceptibilities are peaked near zero due to the fact that most interactions are positive (beneficial) for one species and negative (detrimental) to the other. The relative position of the highly critical species in phase space takes into account long-range interactions but loosely correlates to many edges that are positive for the interacting species and negative for the critical species, and few with the opposite orientation.

In mutualistic interaction networks, topologically important species (such as generalists) are hypothesized to maintain biodiversity [27], and simulations of the extinction of these topologically important species support this view [7, 12, 13, 18, 28-30]. Our assessment provides further evidence for the functional role of critical species and vulnerable communities, specifically the role of topologically important species in maintaining biodiversity. However, our work also highlights the significance of the relationship between node properties and network topology. The combination of high density, high nestedness, and few, sparsely populated strongly connected components in vulnerable communities suggests that they tend to have a densely connected core and a loosely connected, tree-like periphery. Critical species are positioned at the dense core of the community; their removal has a cascading effect through the remainder of the network. The structuring of interactions within a community into nested subsets around a core protects species on the periphery of an interaction network [31], as well as minimizing competition [32]; however, as we have shown, the loss of species at this core can have dire effects. We propose that while increased nestedness promotes stability when nestedness remains low overall, there is a tipping point beyond which a community relies too strongly upon core species: their presence is critical to the continued survival of
the community. If critical species are at higher risk of extinction [see 30], prioritizing their conservation becomes crucial.

The behavior of mutualistic ecological communities is highly complex, which makes the characterization of their dynamics and the development of management strategies challenging [33]. The methodology discussed here integrates dynamical and topological analysis to determine the vulnerability of the entire community to collapse due to the extinction of a single species. We find that vulnerable communities have properties that distinguish them from robust communities; additionally, we find that critical species in vulnerable communities have properties that distinguish them from the other species in those communities. Importantly, the techniques discussed here are testable in real ecosystems, and as such collectively constitute a powerful analytic tool for identifying both real ecological communities that are vulnerable to significant or total collapse, and the constituent species whose continued existence is necessary to prevent such collapse.

ACKNOWLEDGEMENTS

We are grateful to an anonymous referee for valuable suggestions. This material is based upon work supported by the NSF under a grant awarded in 2008 to S.Y., and partly supported by NSF grants DEB-0815373 to K.S. and CCF-0643529 to R.A. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the NSF.

WORKS CITED


FIGURE AND TABLE CAPTIONS

FIG. 1. (a) An example community where plants (pollinators) are represented with diamonds (circles). Incoming pointed (flat) tips represent beneficial (detrimental) interactions. (b) A portion of the state transition network that corresponds to the interaction network shown in (a). Every state is a unique combination of present (1) and absent (0) species, listed in the order pollinator 1, 2, 3; plant 1, 2, 3, 4. The state is also represented with a diagram of the interaction network, where “present” species are shaded. The state transition network terminates at a steady state wherein every species is present in the community. (c) If plant 3 is forcibly removed from the steady state shown in (B), the community eventually progresses to the “all dead” steady state wherein no species are able to persist, i.e. the community collapses completely.
FIG. 2. (a) The distribution of the relative community size after re-stabilization following the forced, permanent extinction of a single species. The relative community size is obtained by dividing the size of the community after re-stabilization by the size of the community immediately after the forced extinction. The distribution is approximately normal, with a peak corresponding to a subsequent extinction of 14% of the remaining species. In some cases the extinction is shown to increase biodiversity (relative size is in excess of one). Circles correspond to simulation results and the line is a best fit Gaussian. (b) The distribution of the number of extinctions required to eliminate every species in a stable community with at least 10 species originally present. The distribution is skewed-normal, with a maximal value slightly less than half of the total number of species in the network’s regional species pool (see Methods). (c) The median number of extinctions required for complete collapse of a robust community, as a function of the community’s nestedness. The positive correlation indicates robustness promotes stability over the range of nestedness shown. (d) The distribution of the number of extinctions required to cause complete community collapse, shown as a proportion of the maximum number of species present in the community. Circles correspond to simulation results and the line is a best fit Gaussian.

FIG. 3. The fraction of species present in a strongly-connected component (sub-community), for communities categorized by the fraction of remaining species that are lost in the equilibration process subsequent to the permanent loss of a single species (see Methods). Circles indicate average values, with circle radius corresponding to the number of communities represented. Error bars correspond to the standard error of the mean. All data points are left binned. (inset) The distribution for the number of strongly-connected components follows a similar trajectory, with near a single strongly-connected community observed in instances where a single extinction drives the entire community extinct.

FIG. 4. Network- (a, b) and node-level (c, d) properties of mutualistic ecological communities when subjected to a single-species elimination, as related to the fraction of species that are lost in the subsequent equilibration process. Black circles indicate average values, with circle radius corresponding to the number of species or communities represented. Error bars correspond to the standard error of the mean. All data points are left binned. (a) Communities that experience minimal loss in biodiversity have the lowest values of nestedness, with a slight increase for communities that experience an increase in biodiversity and a significant increase for communities that experience significant loss of biodiversity. Gray squares correspond to the nestedness Z-scores relative to configuration null models, which preserve the density and degree distribution of the source community. (a inset) A complementary view, with the axes switched and data binned according to nestedness value rather than percent extinct. The conventional view of increasing nestedness leading to more robust networks is shown to hold for low values of nestedness. (b) Higher network density is correlated with significant changes to biodiversity. (c, d) Species whose extinction causes greater loss in biodiversity tend to have greater betweenness and closeness centralities c.f. species in robust communities. Values are normalized to the maximum values observed in the source network.
FIG. 5. (color online) The distribution of node-network influences and susceptibilities for noncritical species in robust communities (blue circles), vulnerable communities (yellow squares), and critical species (black triangles). (a) A broad view at the 0.001%, 99.99% range for species in robust communities. Adjacent curves indicate that the distributions of noncritical species in robust communities are centered near the origin for both measures. (b) A zoomed view showing the 3%, 97% range for species in vulnerable communities. The dashed lines indicate the median values for critical species, and the adjacent data and curves show the distributions and best-fit Gaussians to the data in both dimensions. The distribution of critical species is shifted towards positive influences and negative susceptibilities compared to the distribution of other species.