Structural controlability of pollination networks

E. Fernando Cagua[1] and Daniel B. Stouffer[1]

[1]Center of Integrative Ecology, School of Biological Sciences, University of Canterbury

December 10th, 2015

Abstract

Si sabemos que el disfrute exige lentitud, y que—mas en general—la felicidad se asocia con el ir despacio, por que corremos tanto? El arte de comer estriba en saborear cada bocado sin pensar en el siguiente, sin apresurar el siguiente. El arte de leer, en demorarse en cada palabra como si el sentido del escrito entero estuviera contenido en ella. El arte de amar, en vivir cada momento de la relacion con la persona amada como si fuese el destino de toda la historia del mundo, desde la aparicion del primer organismo unicelular hasta hoy. Y asi podemos generalizar a las demas actividades, creo, hasta obtener un arte de vivir. Para mi se resume en la palabra ahi. — Jorge Riechmann

Keywords— Control theory

Introduction

This intro was written when I though that the paper was mainly gonna be about the relative importance of species and how invasives fit there. Read with caution as I don't think that's the case anymore

Ecological communities are formed by the interconnection of several species. Therefore, changes in the abundances of one species can potentially alter the abundances of the species they interact with. For instance, in a classic example of ecosystem cascades, a reduction on the abundance of sea otters, an important predator or sea urchins, can drive a dramatic reduction on kelp abundances because the sea urchins that consume kelp are released from predation. It has been long established that some species, like the sea otter, have a disproportionate large effect in their environment relative to their abundance.

In several ecosystems the relative importance of species have been identified based on empirical observations of long term dynamics. However, in less studied, highly diverse, or where the ”keystone” role is shared by several species, it can be challenging to determine which is the set of species that influence the most the ecosystem dynamics. Alternative approaches that recognize a continuum of importance and that are less dependent on empirical observations have also been developed. Some of them are based on metrics that evaluate their position in the food web or on mass balance models of functional groups. Nevertheless, these approaches are conceptually limited to throphic interactions and in general ignore the structural mechanisms that allow or prevent the spread of perturbations in the ecosystem.

From a systems perspective, perturbations like over-exploitation, eutrophication or global warming are equivalent to management actions like culling, no-take areas or captive rearing in the sense that they have the potential to modify the abundances of one or several species in the ecosystem. Therefore identifying these key species is crucial not only to predict how these perturbations will spread trough the community but also to guide effective conservation efforts.

Recent work on the control of complex systems suggest that in principle it is possible to alter any ecological community’s composition, by modifying the abundances of just some key species. Here, we apply these theories to estimate the controllability of different ecological communities and to find the species, that due to the structural characteristics of their interactions are more likely to drive the dynamics of the community.

Methods & Results

Using six paired plant pollinator communities we explored the concepts of control theory. This is the initial data set things will be tested. If more data is considered necessary we’ll see that later.

Controllability of ecological communities

Minimum set of species

Using maximum matching to control all species in the network. We found that depending on the scheme around of the species are needed to have a grasp at the whole community (Figure 1). However, the relative ranking of networks

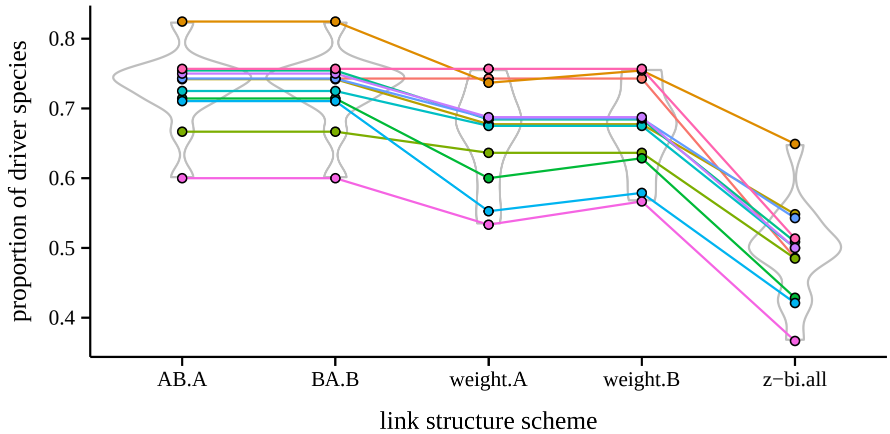

Figure 1: Proportion of species that needs to be controled for all twelve networks under different link assumptions. AB indicates that the links go from plants to pollinators; BA that the links go from pollinator to plants; "weight" that the direction of control is chosen by dependency (ties are resolved by giving priority to plants in weight.A and to pollinators in weight.B); z-bi indicates that the links go in both directions

remains relatively unchanged across schemes. So rather than focusing on the differences across schemes perhaps it would be interesting to see what characteristics make a network to require more or less driver nodes to be completely controled. Alternatively it would be interesting to see if there are systemic differences between different types of networks. An initial step would be to stay with bipartite networks of the mutualistic type (plant/frugivores) or the the consumer-resource type (host/paraiste or predator/prey). Perhaps some model networks of varying parameters?

There was no significant difference between the number of species needed to control the system in an invaded and an uninvaded ecosystem.

When comparing the number of control nodes of the observed networks to their random counterparts an interesting pattern emegerges (Figure 2). The number of control nodes is significantly larger than randomisations that mantain the degree distribution of the plants. However there is no significant difference when the randomisations mantain the degree of pollinators (or both plants and pollinators). This pattern is consistent across all schemes of link directionality. This suggest that the patterns of pollinator visitations are the structural characteristic that determines the controlability of the network. We probably should analyse other pollination networks to see if this pattern hold. We also should think about the implications of this finding.

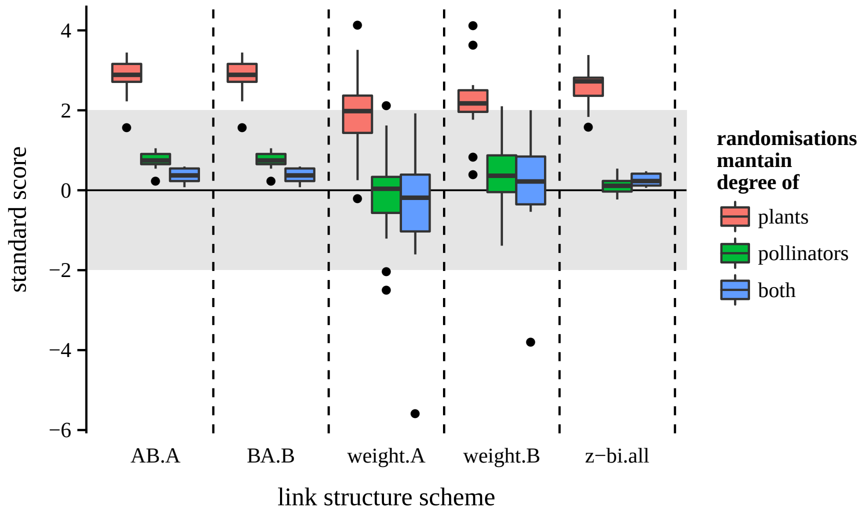

Figure 2: Difference between the number of control nodes in the analysed networks versus randomisations of the same networks. Unshaded areas suggest a significant difference. Schemes same as in Figure 1

It occured to me that it is possible that the patterns (and perhaps others observed like more phylogenetic signal in polinators than in plants) are a consequence of a greater diversity of pollinators in the interaction networks. More pollinators relative to plants imply that randomisations that mantain the degree distribution of pollinators are more constrained than those that only mantain the degree distribution of plants. It also means that statistical tests on that group have a more observations and therefore rendering significant results for smaller effect sizes. I cheked that with the 12 pollination networks of the study and there seems to be some evidence for that (Figure 3). Perhaps it would be good to check with other networks to see if it holds.

To understand the effects of the directionality imposed by the dependencies (only for the weighted method), I randomised the direction of the links. I compared the number on control species in the weighted scheme with a randomisation that mantain the interacting species but randomise the direction of dependency (Figure 4). In pretty much all cases, the number of control nodes is significantly larger in the observed networks.

Profiles of critical and redundant species

Dividing the nodes and links among critical, ordinary and redundant we can have an idea of "control profiles" of the networks. Both species and interactions

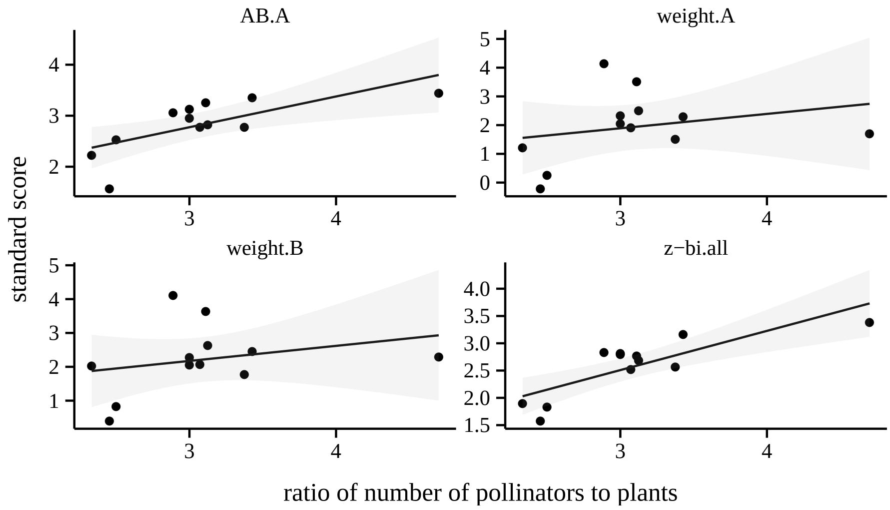

Figure 3: When considering only randomisations that mantain the degree distribution of plants, there seems to be a possitive relationship between the ratio of pollinators to plants and the Z-score of randomisations. The pattern is, however, not significant for the weighted schemes. The z-scores are identical for the AB.A and the BA.B schemes. Schemes same as in Figure 1

can be classified as redundant, ordinary and critical depending on the number of species that need to be controled to gain full control over the network once the species has been removed (Figure 5 and 6).

On the species redundancy profile (Figure 5), the profiles are also highly dependent on the chosen directionality scheme. For instance, in the case in which species have links in both directions there are not ordinary species and the number of critical species is relatively high, this is not surprising because under this scheme, there is a much smaller set of species that needs to be controled and therefore removing one of these can increase the minimum set of species needed to control the network.

However, there appears to be a relatively small variability between networks analysed using this approach. Again, as in the previous section, I think there is little that can be understood by using just this data-set. Again I think it would be most interesting to see if there are systemic differences between different types of networks where a larger variation in the profiles is expected.

On the interaction redundancy profile (Figure 6) things seem a bit more interesting because the variation across networks and across link directionality schemes is pretty small. For instance for the bidirectional scheme, although the total number of links is doubled, the relative proportion of links that are critical

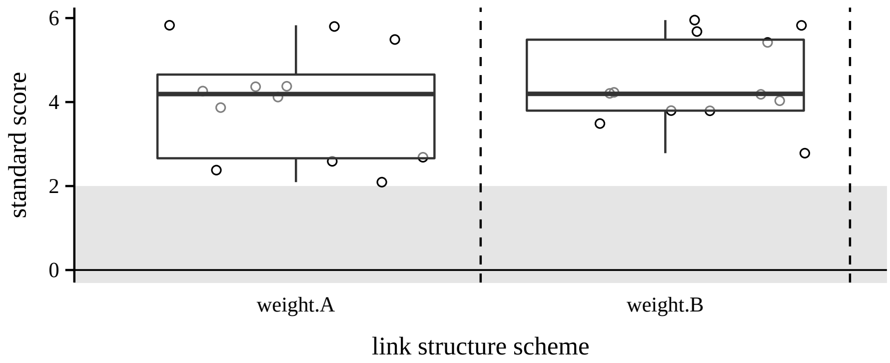

Figure 4: In the weighted case, we compared randomisations that maintain the network structure but shuffle the directionality of the dependency. Observed networks need a significanly larger number of control nodes than randomisations. The difference is more marked when there is a preference to assign dependency from plants to pollinators when dependency is equal in both directions. Z scores are based on 999 network randomisations.

and redundant is mantained.

Profiles of internal/external dilation

J. Ruths and D. Ruths (2014) proposed another clever way to characterise the control profile of a network by charactering the number of nodes that need to be controled among the "mechanism" that originates it. Similar as the maximum matching algorith, the idea comes from the concept that in the absense of cycles one node can at most control another one. Then the origin of control nodes can be splitted between nodes that need to be controled because they are a "source node" those that have only out links, control nodes associated to "external dilations" and those associated to "internal dilations".

It could be interesting to compare different networks using this framework. They argue that source dominated networks should be relatively easy to control in a top-down fashion. On the other hand internal dilation dominated networks represent the signature of a network that self-regulates and is not meant to be controled in that it's not designed to visit every possible state.

However as we can see here this framework is pretty useless when the direction of control is not perfectly known (as in most real wolrd networks really...). For instance when applying it to the data-set (Figure 7), we see that when the control flows from pollinators to plants (AB.A) the network is source dominated simply because the pollinator diversity is greater than the plant's. When we asume that each guild can control each other (z-bi.all) the networks are in-

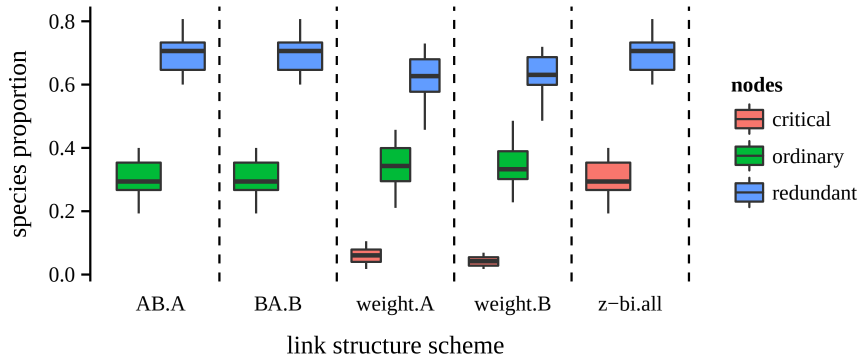

Figure 5: Proportion of species that are ordinary (when removed the number of species needed to control the network doesn't change), redundant (when removed less species are needed) and critical (when removed the number of species needed to control the network increasses). Schemes same as in Figure 1

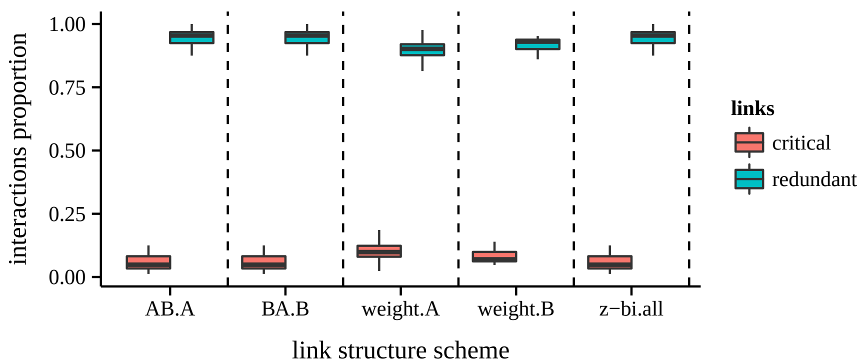

Figure 6: Proportion of interactions that are ordinary (when removed the number of species needed to control is reduced), redundant (when removed the number of species needed to control doesn't change) and critical (when removed the number of species needed to control the network increasses). Schemes same as in Figure 1

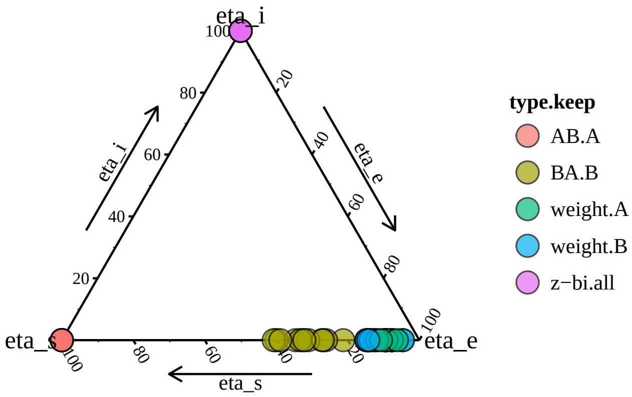

Figure 7: Control profiles of pollination networks. Control nodes are split among nodes that need to be controled because they are a source node, some that are control nodes because they respond to internal dilations (cycles) and some that are control nodes because they repsond to external dilations (sink nodes not matched)

ternal dilation dominated because it's purely composed of cycles. However the weighted version might offer a meaningful possibility to develop the analysis.

Ranking species

By frequency in control sets

There are multiple sets of species with which is structurally possible to control network dynamics. An initial posibility os to relate the importance of a species for network dynamics to the frequency at which a species is present in the control sets.

Turns out unsurprisingly that the importance of particular species and its distribution over a networks dependends on the scheme chosen for the directionality of the links (Figure 8). When pollinators control plants (AB) no plant must be controled and most pollinators have an importance close to one. On the other hand when plants control pollinators (BA) presumably because the pollinator diversity is larger than the plant diversity, the pattern is less "bimodal" and intermediate levels of importance are allowed.

Among the weighted versions there was little difference if dependency ties were resolved for plants of for pollinators (Figure S1), again there are more species with intermediate levels of importance. Even more spread out are things

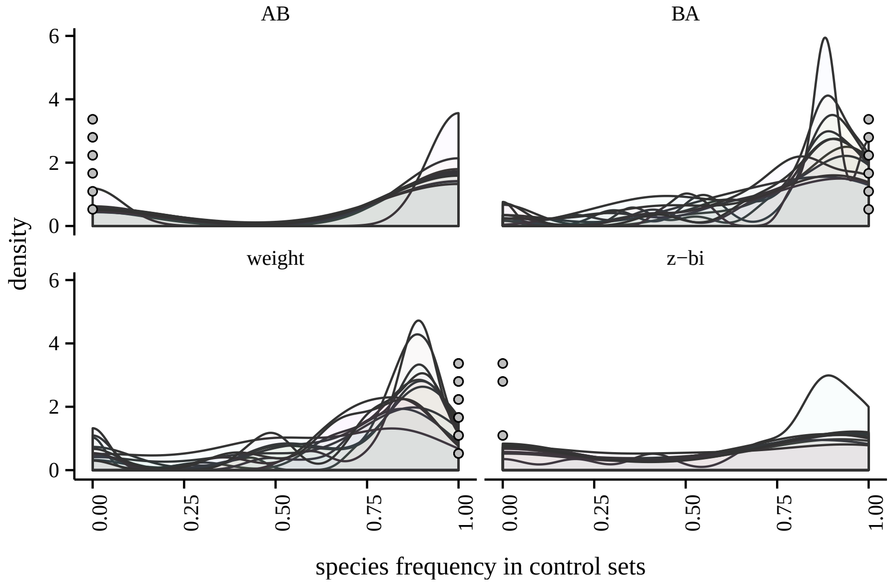

Figure 8: Distribution of the relative frequency a species is present in the set of control species among the networks. The frequency of invasive species in each of the six invaded networks is denoted with a circle. Results for the bidirectional liks (z-bi) were obatined only for the nine smallest networks, which included 3 invaded ecosystems

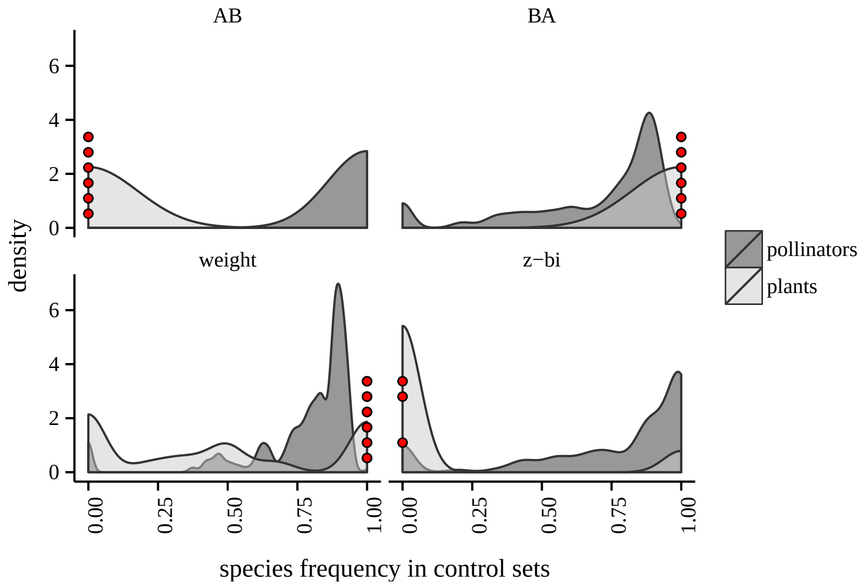

Figure 9: Distribution of the relative frequency a species is present in the set of control species by guild.

when we consider bidirectional links. I imagine that this pattern is partly because there are many many more ways to control the network in the bidirectional scheme than in the others.

In both AB and BA the relative importance of different species is almost fully constrained by the scheme. In the weighted scheme, invasive species are in all cases among the most important ones. Interestingly the pattern is opposite for when the direction can go both ways. It would be good to see then whether these patterns are a reflection of the fact that invasive species are actually dominant or an artifact arising from the fact that all invasive species are plants and that all networks are dominated by pollinators. An initial inspection on the distributrion of plants and pollinators per network for these schemes suggest that that might be the case for the bidirectional scheme but not for the weighted one (Figure 9 and S2).

We explored that option by comparing the relationship between the ratio of pollinator to plant diversity and the difference between mean imporance of pollinators and the mean importance of plants (Figure 10). Although there was a positive relationship it doesn't seem to be significative. However when removing one of the uninvaded networks in which plants were in average more important than the pollinators for the bidirectional scheme (MED4CA, Figure S2) the relationship between the ratio an the relative importance of pollinators and

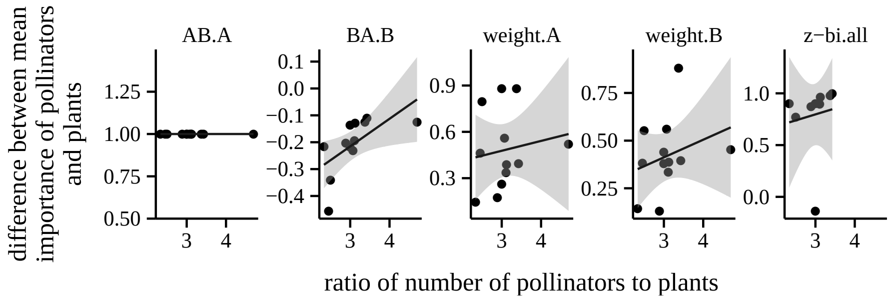

Figure 10: Distribution of the relative frequency a species is present in the set of control species by guild.

plants becomes significant in the bidirectional scheme.

By redundancy

To understand what the redundancy characterisation (splitting nodes and links between critical, ordinary and redundant) means an initial step would be to relate it to the importance given by their frequency on the control sets. However this comparison can only be done among nodes, not links.

Because I think there is enough indication that the unidirectional schemes are crap, we'll focus on the weighted and the bidirectional schemes. There doesn't seem to be a direct relationship between the frequency of a species as a control node and it's characterisation as critical, ordinary or redundant for the weighted scheme (Figure S3). However critical species in the bidirectional scheme tend to cluster towards the bottom of the frequency (Figure S4).

Given that as we saw in Figure 10 plants tend to cluster as well in the bottom of the frequency scale, it's conceivable to think that there is a difference between the profiles of plants and polinators. We explored that difference (Figure 11) and indeed there there are marked differences bewteeen plants an pollinators.

In the bidirectional scheme, most polinators are redundant, which means that removing that species from the system reduces the number of control nodes. Which could be explained if we take into account that there are more pollinators than plants. In the weighted scheme things are a bit more interesting, most plants are ordinary in the sense that removing them doesn't change the number of nodes that has to be controled. Wile some are critical in the sense that removing them means that one needs one more node to control the network. All invasive species were categorised as ordinary.

The redundancy is capturing different variability than the frequency in the

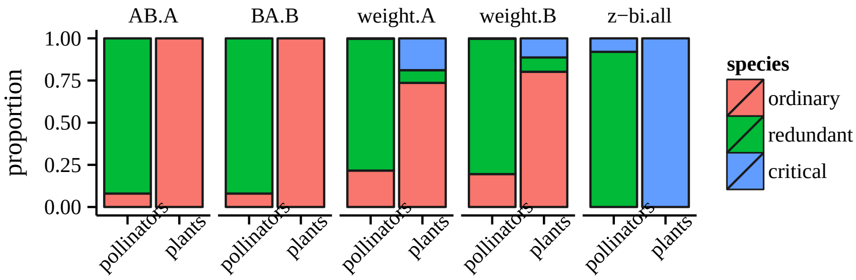

Figure 11: Proportion of pollinators and plants that are ordinary, redundant or critical among different schemes. See Figure 5 for details about this characterisation.

control sets. However they also show a clear distiction between the role of plants and pollinators.

Interaction heterogeneity

Liu, Slotine, and Barabási (2011) suggest that the controlability of a network and specifically the number of control nodes is strongly correlated to the degree distribution of the network, it is not unreasonable to think then that species with a high degree would play a special role. More recently Gibson et al. (2015) showed that what really matters is the interaction strength of the species and that strongly interacting species (those that influence other the most), and not necessarely the most abbundant, are the ones that can drive the community from one state into another. Here we explore the relationship betweeen the two metrics of species controlability, the degree and the mean interaction strength of species.

First we start by recognising that mean interaction strength is correlated with the degree of a species in a network (Figure 12). The effect is not terribly strong, particularly for pollinators, but suggest that species that interact with lots of species also tend to visit them a lot. Sadly there is no data on relative abbundances to investigate if both degree but particularly interaction strength (as infered from visitation frequency) is a consequence of mass action.

Here we only look at the effect on the weighted scheme. We can see in Figure 13 that the patterns are different for plants and polinators. For plants importancy in controling the network is positively related to both the mean interaction strengt and the degree of the species. However for pollinators that's

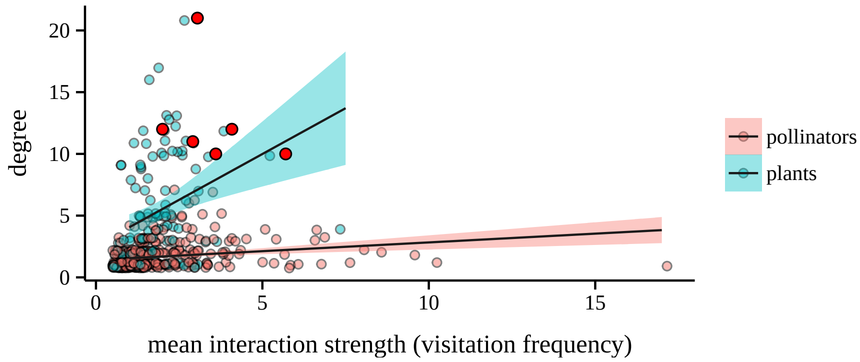

Figure 12: Relationship between the mean visitation frequency per species (as a surrogate of interaction strength) and its degree in a particular network. Invasive plants are depicted with a red circle.

not the case: the relationship between importance and mean interaction strengt is pretty weak, but there is a significant negative relationship between the degree and the importancy.

As expected plants tend to have a larger degree than polinators. But regardless of the guild one would expect that as the degree of a species increases so does the likelyhood of it being less dependent than its counterpart and hence have more outer links, which in theory should increase its importance as driver node. The fact that this happens with plants but doesn't happen with pollinators is another part of how plants and pollinators play a distinct role in teh controlability of the networks. This might happen in part because of the nested structure of the mutualistic networks. A generalist polinator will still be "nested" within a set of generalist plants, in other words a generalist plant will have more partners than a generalist polinator which drives the arrows from plants to polinators anyways. (yeah fer this doesn't make sense...)

We therefore explored the posibility that the importance is related to the contribution to nestedness. Voila! In this case both plants and polinators have a clear positive relationship (Figure 13c)! That's pretty cool but I don't know how to interpret this. Daniel help please!

Because the profiling of nodes proposed by Guimera and Amaral (2005) captures a lot of the connectivity of nodes we explore the relationship with the roles as well.

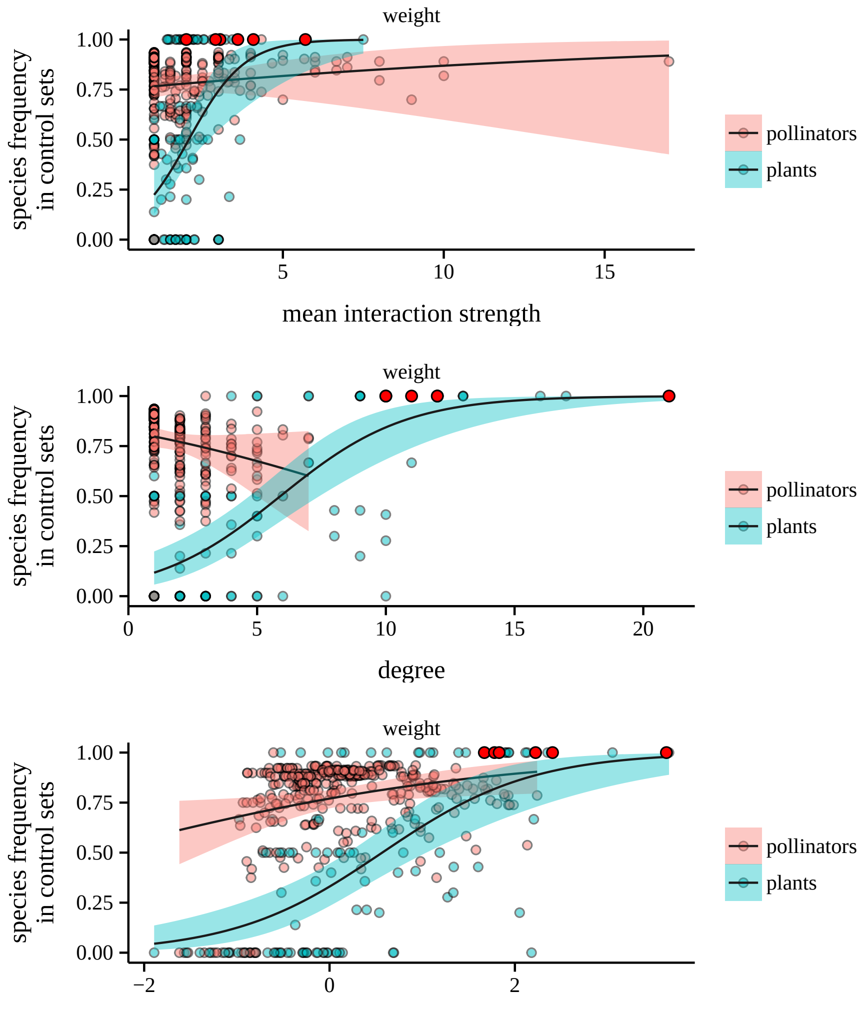

Figure 13: Relationship between the mean visitation frequency per species (as a surrogate of interaction strength) and its degree in a particular network with its presence in the control sets. Invasive plants are depicted with a red circle. Contribution to nestedness were calculated using 99 network randomisations.

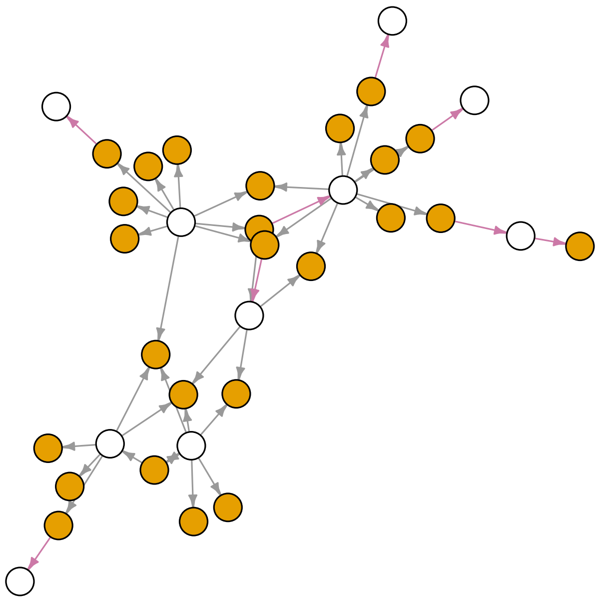

Figure 14: Example network. Plants and polinators are shown as white and yellow circles respectively. Redundant links are shown on grey and critical links are shown on purple.

Link redundancy and species

We got some links that are redundant and other that are critical from the control point of view. I wanted to have a look at what drives that. I haven't done the formal analysis yet but a visual inspection of the network is reveals a pretty strong pattern.

Critical links tend to occur predominantly when control direction goes from pollinators to plants, this is when the plant is more dependent in the pollinator than the pollinator is on the plant (which is much less common than the other way around), and generaly involve specialist plants (Figure 14).

Discussion

For instance, restoration projects that aim to drive an invasive species out of the community, to reestablish a particular ecosystem service or in general modify the ecosystem state have had a limited amount of success. We argue that quantifying the relative dynamic importance of species the set of key species is the first step to evaluate current management efforts as well as to design targeted and informed

Acknowledgmenets

The authors thank Dr. Takeuki Uno for the insight provided to find the set of all maximum matching algorithms, and Bernat Bramon and Marilia Gaiarsa for feedback in early stages of the project. EFC acknowledges the support from the University of Canterbury Doctoral Scholarship, the University of Canterbury Meadow Mushroooms Postgraduate Scholarship, a travel grant from the European Space Agency and a Rutherford Discovery Fellowship (to DBS). DBS ackloledges the support of a Rutherford Discovery Scholarship, administered by the Royal Society of New Zealand.

References

Gibson, Travis E et al. (2015). "On the Origins and Control of Community Types in the Human Microbiome". In: pp. 1–84. arXiv: arXiv:1506.05171v1.

Guimera, Roger and Luis A. Nunes Amaral (2005). "Functional cartography of complex metabolic networks". In: *Nature* 433.7028, pp. 895–900. DOI: 10.1038/nature03286.1..

Liu, Yang-Yu, Jean-Jacques Slotine, and Albert-László Barabási (2011). "Controllability of complex networks." In: *Nature* 473.7346, pp. 167–173. DOI: 10.1038/nature10011.

Ruths, Justin and Derek Ruths (2014). "Control profiles of complex networks". In: *Science* 343.6177, pp. 1373–6. DOI: 10.1126/science.1242063.

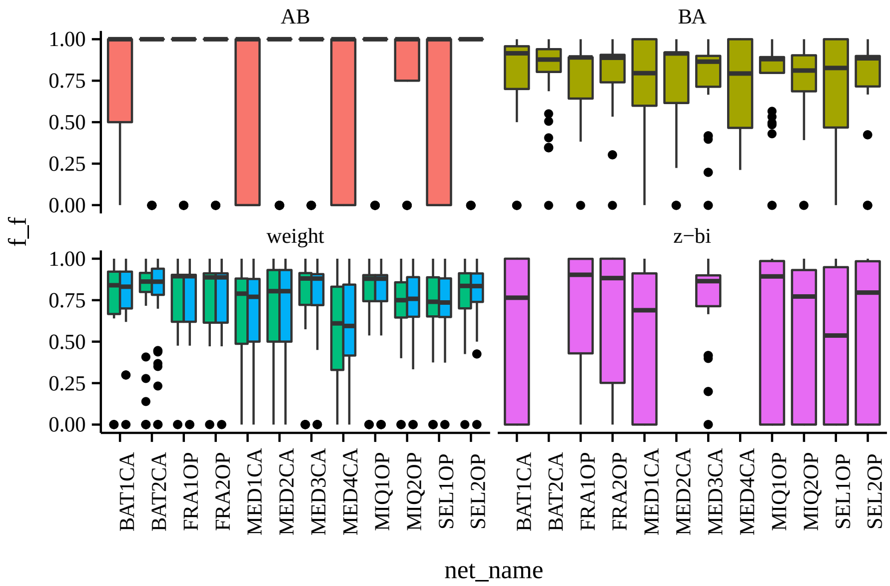

Figure S1: Distribution of the relative frequency a species is present in the set of control species among by network

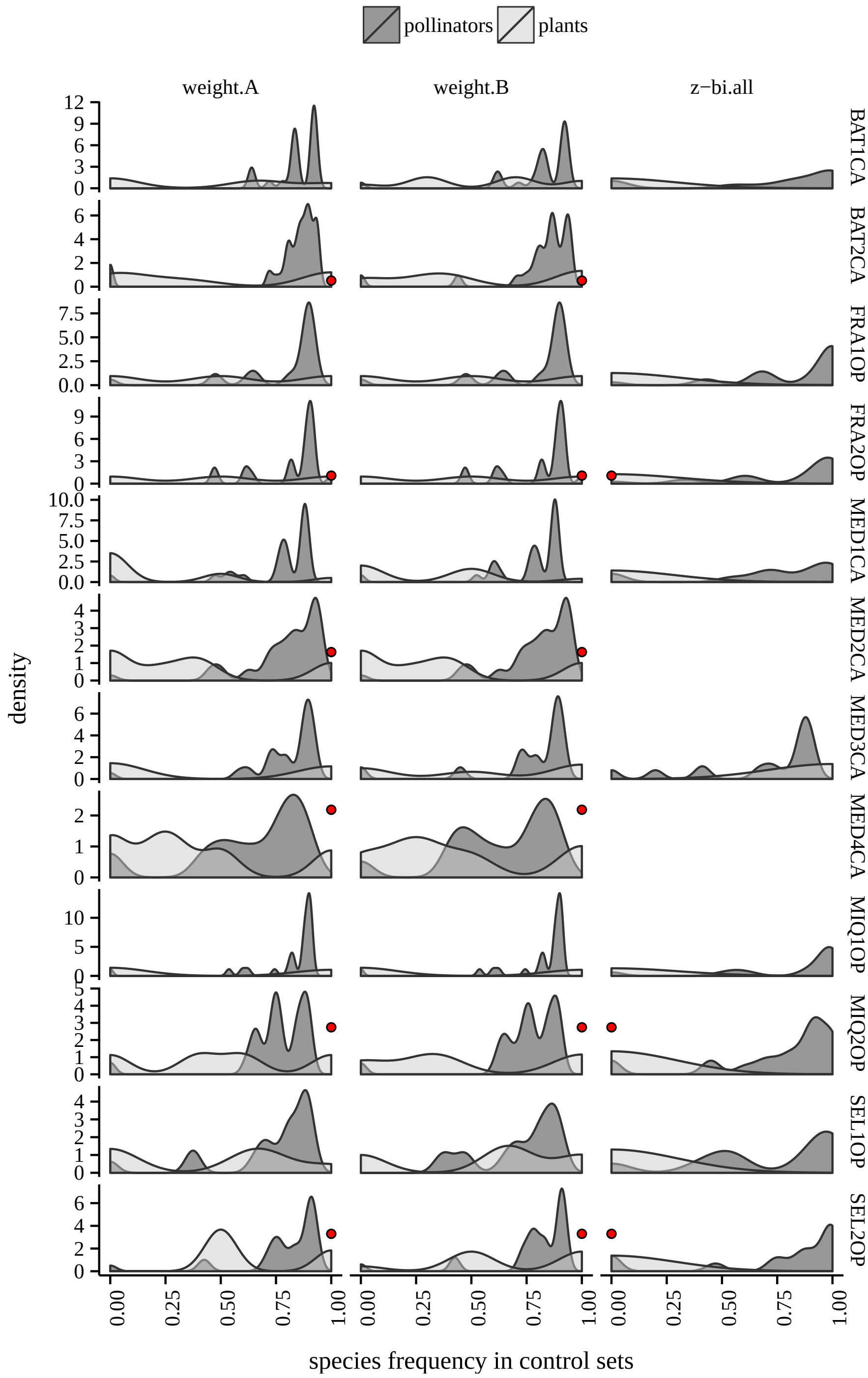

Figure S2: Distribution of the relative frequency a species is present in the set of control species among by network partitioned by plant or pollinators. Invasive species is depicted with a red dot.

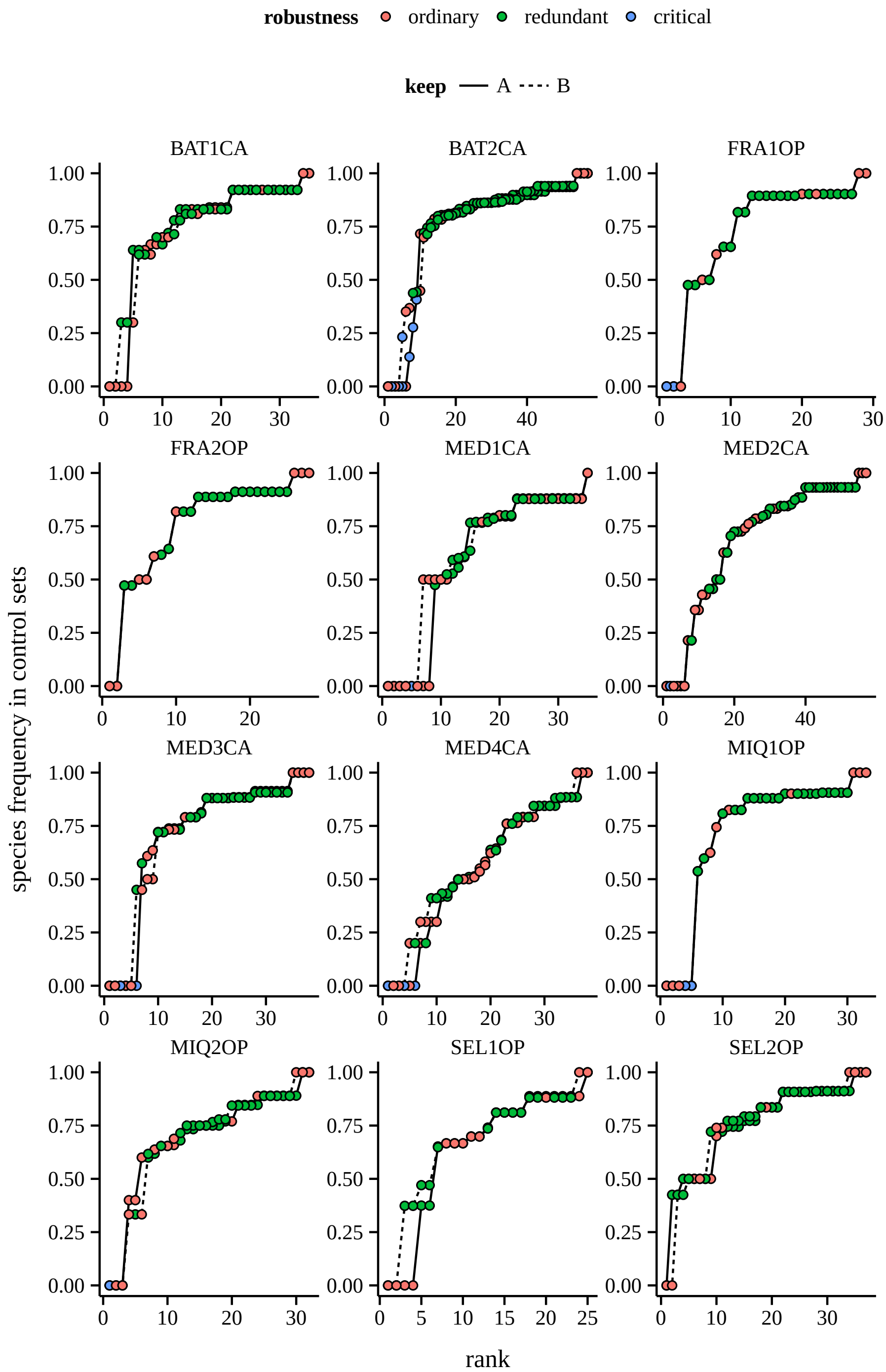

Figure S3: Relationship between species' frequency in the control sets and the redundancy categories for the weighted scheme.

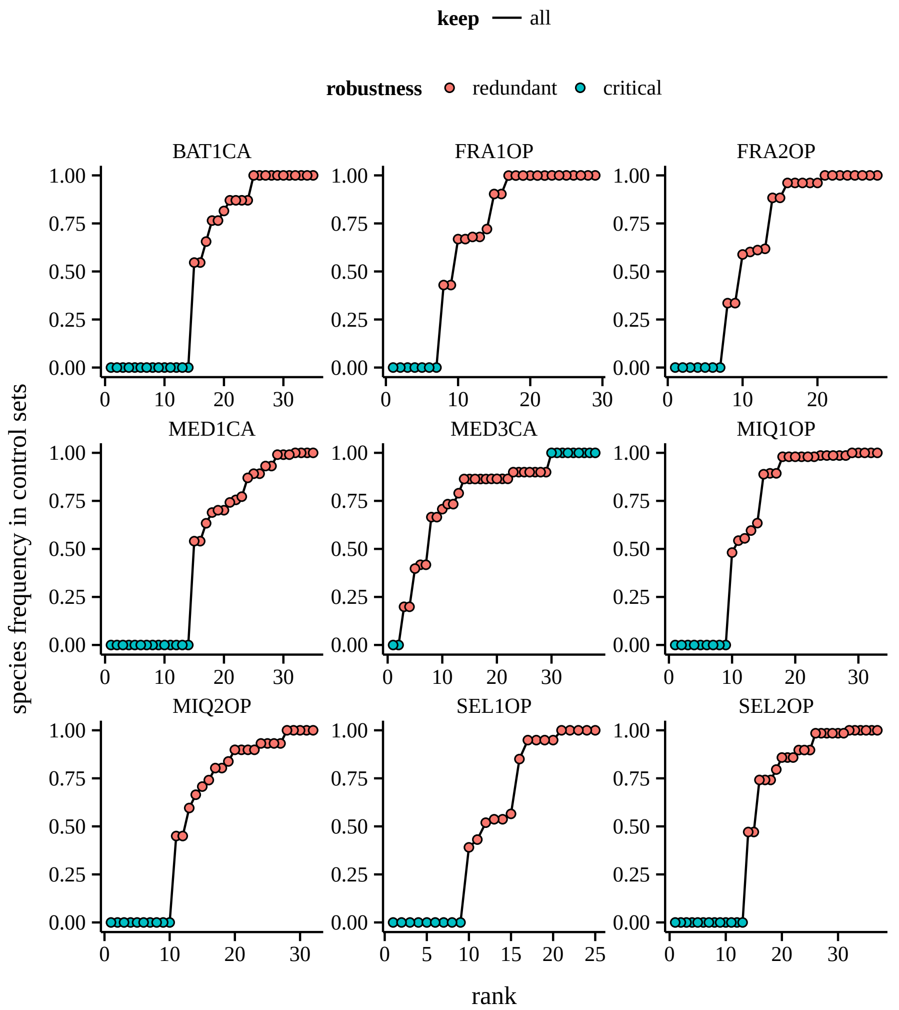

Figure S4: Relationship between species' frequency in the control sets and the redundancy categories for the bidirectional scheme.