Keystoneness, centrality, and the structural controllability of ecological networks

JEcol-2018-0545 - Detailed response to comments

Fernando Cagua, Kate L. Wootton, Daniel B. Stouffer

We are thankful to the assistant editor James Ross, the senior editor and David Gibson, and three anonymous reviewers for the insightful comments to improve the manuscript. In the following sections, you can find a detailed response to each comment raised during the previous round of review.

Editor comments

> "I agree that your manuscript is much improved and now only in need of relatively minor revisions. In addition to addressing the AE and reviewer's comments, please add terms related to invasive species to the keywords to aid in discoverability of your paper."

Many thanks for the encouraging words. We have included 'invasive species' as a keyword. To make space for this keyword we have removed 'controllability' which is already highlighted as part of the title.

Associate Editor's comments

> "I think this ms now makes much clearer the potential insights to be gained by applying control theory to interaction networks. In particular, the authors have done a nice job of explaining how their findings reinforce existing ideas in the literature on the stability of ecological networks while also acknowledging limitations. The reviewers identified a few remaining issues that should be straightforward to address. I also noticed a few issues:"

Thank you for the positive feedback. We have addressed all of the specific issues raised and highlight these in red text in our resubmitted manuscript.

> "*L9*: I think here and throughout much of the text, species should be possessive when followed by a noun ("species role" should be "species' role"). The possessive is currently used in some instances but is inconsistently applied throughout the ms."

Accepted and changed. We have checked through the manuscript and corrected this where relevant. Besides *species' role*, we also changed instances of *species' contribution*, *species' degree*, *species' control capacity*, among others.

> "*L25*: 'every invasive species in our dataset was part of it' Here it was unclear what 'it' refers to (the set of critical species), and the reference to a singular dataset when you have 20 networks in total was a bit confusing to me."

Accepted and changed. We have made this clearer as follows: "and included every invasive species in our dataset" (page 2, line 25).

> "*L23-26*: in these lines, you refer to critical species, but I think the distinction between having high control capacity versus being critical needs to be defined (briefly) earlier (it is stated in *L18-19* that you will quantify both control capacity and identify critical species, but it is not explained how they are distinct)."

Accepted and changed. We now identify the distinction between high control capacity and being critical on page 2, line 17. The text now reads "It also allows us to quantify a species' control capacity-its relative importance-and identify the set of species that are critical in this context because they have the largest possible control capacity".

> "*L27*: unclear what 'its' refers to."

Accepted and changed. We now explicitly state "species' degree" in page 2, line 27.

> "*L234, 235*: awkward to refer to species as 'they' and 'their' "

Accepted and changed. We have changed this to "it" and "its".

> "*L296*: 'More details in Supporting Information.' is a sentence fragment."

Accepted and changed. We have corrected this on page 13, line 302. The text now reads "See Supporting Information Section S5 for more details".

> "*L343*: I am not sure about citing Table 2 for the first time in the Discussion. Might it be more useful to cite this in the Introduction?"

Accepted and changed. We now introduce the glossary at the end of the introduction on page 4, line 102.

> "*L111, 114*: 'are comprised of' should be 'comprised'. *L339, 432*: 'at' should be 'in'. *L341*: 'allow' should be 'allowed'. *L355*: missing period."

Accepted and changed. Thank you for pointing these out; we have corrected all of them.

Reviewer comments #1

> "I'd like to commend the authors to this revised version of the manuscript on 'Keystoneness, centrality, and the structural controllability of ecological networks'. I thoroughly enjoyed the re-written manuscript, which is now considerably more accessible and clearly targets the community of network ecologists, while not ignoring the importance of the applied relevance of the findings. Overall, I don't have any major comments. I guess the language is in parts on the casual side, yet without compromising the preciseness or clarity of the text. I noticed a few typos or style errors that the authors may want to address. Congratulations on this thought-provoking methodological synthesis."

We are pleased that the reviewer enjoyed the re-written manuscript. We did too!. We are also grateful for the typos and style errors they have identified and have striven to correct them all.

> "A few references require editing to journal style (e.g. remove initials for many citations)"

Accepted and changed. We have corrected all citations. One of them, however, still shows an author's first name initial, as it is necessary to distinguish two authors with the same last name (page 10, line 216).

> "*L388*: remove repeated 'other'. I suggest that the authors re-write this sentence to improve clarity."

Accepted and changed. We have removed the second 'other'. As suggested, we have rewritten the sentence to improve clarity (page 18, line 394). It now reads as follows: "First, structural controllability emphasizes that the effect a species has on other species is not independent of the effects that the other species have in the community".

> "Reference list needs updating, e.g. Ne'eman et al. is not 2009 but 2010."

Accepted and changed. Thanks for pointing this error out. We've corrected the reference.

> "*L314*: open bracket missing. *L358*: change 'instead' to 'Instead'. *L381*: remove second comma. *Fig. S6* header of plot b). correct spelling of 'critical'."

Accepted and changed. Thank you again. We have made all these changes as indicated.

Reviewer comments #2

> "I think the concepts presented in this manuscript are a useful addition to the literature on management of ecological networks. I agree with the authors that it is important that

> the discussion surrounding what constitutes a 'keystone' species should go beyond centrality metrics. I found the results and concepts to be generally well presented, particularly graphically, and the authors present a thorough sensitivity analysis to various assumptions.

We are very thankful for the positive feedback but also for the issues highlighted. We have aimed to address them all.

> "The methods are mostly clear but there is no description of how the 'likelihood of being a superior node' is calculated."

Accepted and changed. Thanks for pointing this out. We have modified the sentence mentioning 'the likelihood of being a superior node' in the first paragraph of the section S3.4 to "...and whether it is a superior node..." to avoid confusion. Then we define the 'probability of being a superior node' at the end of the section: "It is important to note that a species that is found to be a superior node under a particular control configuration will be so in all possible control configurations for the particular non-reciprocal network. We, therefore, denote the average of σ across all non-reciprocal versions as the probability of the species being a superior node in its community".

> "I would suggest including clarification on the limitations of using control capacity to summarise species' importance, and examples to help the reader build an intuition regarding the control capacity metric. As pointed out, networks that give rise to a species having a high control capacity can be quite distinct to networks that give rise to a species having a high value of particular centrality metrics. Providing an example in the Discussion of, for instance, those species that have a control capacity of 1 but a low degree would be very useful."

Partially accepted and changed: We think that more studies are needed before we can accurately build an intuition of what constitutes a species with a high control capacity, especially in regards to other kinds of networks. As such we have added a clarification in page **??**, line **??** that states "As much potential as any metric or metrics to summarise species' importance might appear to have, it's clear that we also need more empirical studies in different types of networks in order to build intuition and ground truth their usefulness". We also considered the suggestion about including a specific example, but we ultimately concluded that none would be particularly didactic for a reader unfamiliar with the exact empirical context.

> "Jargon occasionally obfuscates the concepts being introduced and there are places where I thought the language used was vague (esp. if terms being used have particular meanings depending on the field you're coming from). I would suggest clarifying, or rewording, the following: 'desired control trajectory' - does this mean 'final state'?""

Accepted and changed: We have attempted to clarify this sentence as follows: "Finding out exactly how

difficult it is to control a network depends strongly on the particularities of the desired control trajectory (i.e. the path to the desired final state) as well as the dynamical relationship between nodes." (page 9, line 191).

"*L1364*: 'dynamic persistence' "

Accepted and changed: Thank you. To avoid confusion we have removed the word 'dynamic' and now the sentence reads "... our results suggest that structural controllability is closely related to the persistence... " (page 17, line 370). This preserves meaning without sacrificing precision.

"*L1370*: 'dynamic stability' "

Accepted and changed: Contrastingly, in this case, the word dynamic distinguishes a particular facet of stability. We mean the ability of the system to return to an equilibrium state after a perturbation. We now include this definition on page 17, line 376 to make the concept clearer.

"*P9*: 'ecological objective' - simply 'objective'?"

Accepted and changed: We have changed this to simply 'objective', as suggested (page 9, line 190).

"*L405*: 'central in their community' "

Accepted and changed: We've changed the wording to avoid ambiguity. Now the text reads "... to attain high centrality in their communities... " (page 18, line 411).

"*L406*: 'inherently different. The sentence spanning lines 407-410 is quite confusing - please clarify.

Accepted and changed: We rephrased this sentences as follows: "We found that it is not that invasive plants have some different mechanism for influencing the community compared to their native counterparts. Both native species and invasive plants tended to attain a high control capacity if they were important to network persistence, were abundant, and depended little on other species" page 18, line 412.

"*L430*: 'marginally affected by the topology of their interactions' "

Accepted and changed: We have rephrased this line as: "By accounting for network dynamics (even if in a simple way), structural stability incorporates more ecological realism, especially in the extreme scenario in which the structure of interactions within the community only marginally affects the community's state" (page 19, line 435).

"*L436* and *L441*: 'dominate the state [of their community]' "

Accepted and changed: Because we don't have a proper definition, we have modified the conclusion paragraph and now the affected sentences read: "Collectively, critical species have the ability to strongly influence the state of their community and therefore are likely to be highly relevant for ecosystem

management and conservation. While useful, centrality metrics—which have often been used as a proxy for keystoneness—fail to identify some of these species, highlighting their limitations when we fully embrace the notion that ecological communities are dynamical systems. Paine (1969) showed nearly 50 years ago that a single species can sometimes fill this role on its own" (page **??**, line **??**).

"*Discussion*: 'species' strength"'

Accepted and changed: In the Methods section, we define species' strength as the sum of a species' visits (page 11, line 255). For brevity, we did not define it again in the Discussion, but now we have improved the text and briefly describe it when we mention it there too. The text now reads "When controlling for a species' visitation strength (the sum of a species' visits), which is indirectly a proxy of its abundance. . . " (page 17, line 382).

"Minor typos: *Figure S11* caption: 'likelyhood' -> likelihood. *L388*: 'other other'. 'Driver node density' (supp p12) -> 'size of the minimum driver node set'?"

Accepted and changed: All typos have been corrected as indicated.

"*Figure S1b* is referred to (supp. p4) but doesn't exist in the figure"

Accepted and changed: Thanks for noticing this issue. We now refer to the figure as "Figure S1 (center panel)".

"*References*: several references use the term 'plantPollinator', or '. . . ' "

Accepted and changed: We have corrected the reference that included 'plantPollinator' in the title. We have left the references that include an ellipsis in the author list as they are, as the Journal of Ecology Author Guidelines indicate that this is the preferred format when the reference has more than seven authors.

Reviewer comments #3

"In the literature on ecological networks, the term keystone species has often been tossed around without any clear definition. I was pleased to read that here the authors have developed a cleaner, more effective way to identify keystone species in ecological networks than methods that have been proposed before. I think the authors have made the correct choice in emphasizing the conceptual aspects of the approach, and do so quite well. The background on control theory was complex but manageable to read and understand in this context."

We are very thankful for the positive feedback.

> "Overall I found the paper well-written and sound, but there were a few questions I had that could improve clarity in the methods used. In the glm that used nD as the response variable, is there any concern that including N (number of species) as a predictor may cause a problem when the response is D/N? I can understand why this was included, but I can also imagine this resulting in models that show an effect of N when there is none (I could be wrong on this)."

Thanks for the observation. As you mention, we believe it is ok to include the sample size as a predictor in this model for three reasons. First, there are several instances when the probability of something occurring increases or decreases with sampling effort, etc. It *does* imply a non-linear relationship though. Second, if there were a relationship between N on D, then using D/N (as opposed to of D) as the response variable and N as a predictor would actually reduce the size of the effect of N, rather than amplify it. If on the other hand, N and D are completely decoupled, then using D/N as the response variable should affect the intercept but not the coefficient of D. Third, in any case we only found a comparatively small effect of N on D. Although it might be relevant statistically it seems to be ecologically insignificant.

> "In the species level glm, it was not entirely clear if models were fit to each network individually or if all species were lumped together (the former is what I think was done based on my reading of the text). Regardless I wonder if it would change the results if a mixed effects model was fit to lumped data, using network as a random effect."

Accepted and changed: We are sorry for the confusion. In the species level GLM, we lumped all species from different networks together. We have modified the text in page 11, line 251 and page 12, line 263 to make this clearer. Moreover, we actually did test for random effects in the model. We tested random effects that allowed for a random intercept for the network (but also its site, and the study it comes from). In all cases, the variance among groups was zero. This means that likely there are no differences on the distribution of control capacity between networks (or site and study) and therefore it was unnecessary to include random effects in the models presented in the paper. We have corrected this omission and describe the random effects in page 12, line 264.

> "How were the model sets derived? Were all possible sub-models compared or was a step-wise approach used?"

Accepted and changed: We actually compared all possible candidate models. We have edited the manuscript on page 12, line 266 to include this information.

> "Reciprocal links were included in the directed graph when dij = dji. Would it have any effect if there was a little wiggle room there, e.g., when the two are close but not equal?"

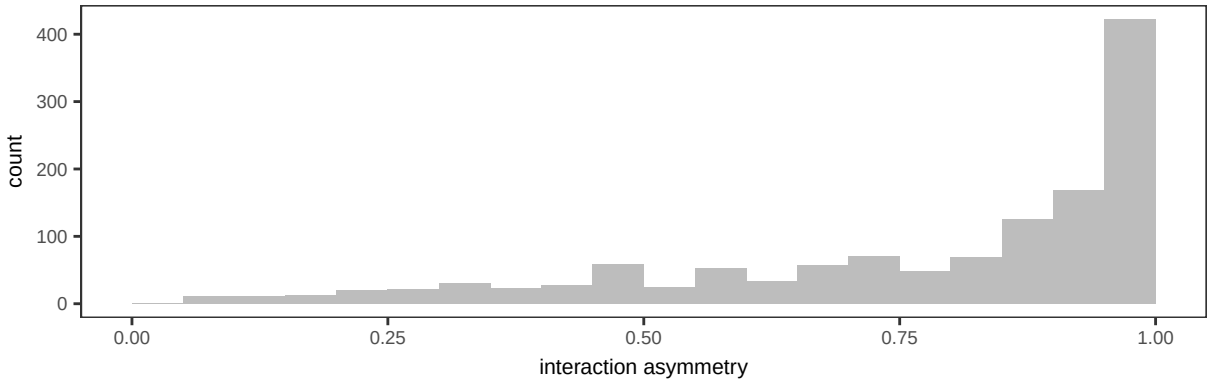

Figure R1: Histogram of the interaction asymmetries between species. Previous research has shown that mutualistic interactions are often highly asymmetric in natural communities. Indeed, only 16.5% of interactions have asymmetries smaller or equal than 0.5; 1% are smaller than 0.1.

Rather than add yet another layer of complexity to our manuscript and given the rapid turnaround of the revised manuscript, we did not conduct this added sensitivity check. However, we note that only a very small fraction of interactions had a small asymmetry (see Figure R1). We, therefore, don't expect our results to change considerably even when we include reciprocal links when the asymmetry is small but not zero.