

Cover letter FE-2015-00758.R1

Dear editor,

please find enclosed a resubmission of MS entitled "*Sampling networks of ecological interactions*", FE-2015-00758.R1, which I submit to your consideration for publication in Functional Ecology. This is a second review of this MS.

ASSOCIATE EDITOR'S COMMENTS TO THE AUTHORS

- *The paper has now been seen by two reviewers, one which previously reviewed the paper and a new one added at this stage. Both reviewers see the overall merit of the article. However, they both also highlight a number of areas where the manuscript seems to give an insufficiently detailed treatment of the subject matter or where it could be improved and/or made more impactful for future readership. Given the extent of these comments, I felt that the manuscript was still at the Major Revision stage and that the revision process should be approached with that level of potential modification in mind.*

I've tried to address carefully all the queries from the editor, as well as from the reviewers. Specifically, I've paid attention to the comments by rev #2 to the first submission and tried to re-address those points that in her/his view were not given sufficient consideration in the R1 revised version of the MS.

I've reduced the length of the MS substantially, eliminating one section and resulting in an overall reduction of 20% of the word count. Yet I had several comments by the referees suggesting adding more material and references. This is always a compromise, but I really agree that the MS was overlength. I very much hope this issue is solved in this new version, which is a substantially shortened MS.

Please find below my point-by-point responses to the reviewers' (*in italics*) queries.

REVIEWERS' COMMENTS TO THE AUTHORS

Reviewer: 1.

- *Although there are some improvements here and there, most of my reservations detailed in the previous review remain valid. To be a loose collection of thoughts and literature reflections on network sampling (including ideas and explanations of interest to some readers), the manuscript may not require any fundamental changes. However, I currently don't see a consistent "framework for developing such a theory [of network sampling]" (the aim stated in the response). The overall structure of subsections and subtitles is the same as in the original submission, and this structure appears rather unclear and disorganized to me. Many of the author's*

responses to my comments stay within the narrow focus of a section or paragraph, where the statements may be correct; but not much has been done to help resolve contradictions among different parts of the manuscript (e.g. which assumptions hold for which parts, etc.), making it difficult to extract clear take-home messages. Readers may struggle to put the different pieces together to envision a framework or overall concept.

I've revised in detail this R2 submission, re-reading the useful comments made by rev #2 to the first version. I'm sorry that he/she has not appreciated much changes following her/his advice. At some points her/his comments were over-demanding in respect to the goals of the study and would require a complete re-structuring of the MS. I have eliminated a whole section on asymptotic diversity estimates and rewritten several sections trying to improve the text flow. I really can't see where the reviewer sees a lack of structure: the MS starts with a discussion of methodological issues involved in interaction sampling, then builds a conceptual connection with species diversity sampling theory to later explore the consequences of forbidden links and their important distinction with unobserved links. Finally, new approaches for missing interactions are proposed based on the recent applications of theory related to the quantification of unobserved species.

- *As stated before, the strong focus on interaction richness estimation (in text and equations) unnecessarily narrows the scope: estimating the variables in Tab. 1 (even if it were possible to estimate all of them reliably) is not sufficient for the goal "to make an inventory of the distinct pairwise interactions" (e.g. L 42 of the manuscript).*

Variables in Table 1 summarize a series of elements that need consideration when sampling species interactions. To make an inventory of actual pairwise interactions I consider fundamental to have a minimum knowledge of the the portion of forbidden links existing in a system. Note that five out of the eight variables deal with *unobserved* links, so it's a form of looking to the unobserved interactions and trying to account for them. I'm convinced this is needed for an informed inventory.

- *Exploring forbidden links based on biological information does consider interaction identities, but from reading the manuscript I cannot see how this could be combined with the richness estimation approaches that ignore identity and just consider interaction sums. In addition, in difference to the author I don't see a strong contrast between forbidden links (in interaction sampling) and species that are predictably absent from a site (in species sampling).*

The contrast is that in the case of forbidden links, the species that could potentially interact are actually present in the local community. Species that are predictably absent from a site have been considered in recent studies as dark diversity, but this component does not add up to the alpha diversity value of the

site. The contrast is biologically meaningful.

- *Although my concerns about dismissing the importance of sampling effects may not fully apply to the revised text, points 2 and 3 in the Summary may still create the impression that most of the unobserved links are "forbidden", while the manuscript acknowledges that the proportion of forbidden links is generally unknown.*

I insist in several parts of the MS that unobserved interactions (zeroes in the adjacency matrix) include a fraction of elements that can be adequately explained by biological restrictions of the partner species. This is the central idea of the paper. I've rephrased the 4th item of the Abstract to read "*Adequately assessing the completeness of a network of ecological interactions thus needs knowledge of the natural history details embedded, so that forbidden links can be accounted for as a portion of the unobserved links when addressing sampling effort.*", so that this important point is clear from the start.

- *For some of the revisions ("done", "edited", etc.) referred to in the response, I could not find any change in the respective text. For example, the sentences around old L 95 = new L 98 are identical, as is the sentence old L 297-299 = new L 294-296.*

For some reason I may have missed these sentences. This is now edited and corrected.

Reviewer #2.

I found this manuscript promising, but deserving of a broader contextualization – here are my main comments.

On forbidden links

- *The definition of "forbidden links" throughout seems a bit incomplete with regard to recent developments. For example, Canard et al. (2014) suggested "neutrally forbidden links", which are not (proximally) determined by life-history traits. I also think "life-history traits" would be better expressed as "functional traits". The idea that some interactions can be "forbidden" can also be viewed as a the consequence of interactions being determined by the interaction between traits and relative abundances (Poisot, Stouffer, and Gravel 2015). The question of "which interactions are forbidden" is necessarily answered by "which interactions are realized", and I would think a paragraph explaining the advances on this question would be important. This point is present in the third entry of the abstract, by the way, so maybe it is simply a matter of making it clearer throughout.*

Functional traits somehow refer to an outcome of the ecological interaction, while the forbidden link idea rests on the fact that species-specific traits (not necessarily functional) limit, constrain, or make impossible some pairwise interactions that otherwise could occur (given that the partners coexist in the same

local assemblage). The idea of "neutrally forbidden links" goes beyond the initial definition of forbidden links, yet was implicitly included in the definition. That definition rests on species-specific ecomorphological, phenotypic, behavioral, eco-physiological or life-history (e.g., migration, phenology) characteristics. All this can be considered as 'intrinsic' properties of individuals of a given species, while the idea of neutrality is more associated to population-level characteristics. The issue of relative abundances determining extremely low PIEs is discussed in the last section of the MS.

- *On natural history*

- *I am not entirely sure that natural history is a necessary information for any network-related inquiry. I realize this is a controversial opinion within ecology. Relying entirely on natural history introduces a great risk of story-telling, as in making the data fit with the pre-established narrative of our intuition about behaviors of the system. I would argue that statistics and quantitative thinking are more important: even simple approaches like variable selection would reveal using an objective basis what are the important traits, for example. If anything, we need better numerical and quantitative network models, in addition/as opposed to more natural history.*

This seems to be a personal opinion of the reviewer. I'm not advocating for relying entirely on natural history. I'm not a reductionist. I believe in the power of comparative quantitative ecology if it is based in a solid natural history background that guarantees the realism of the emerging biological inferences.

- *On the generality of the A:P ratio*

- *Specifically in page 9, the distinction between zoo and phyto centric dramatically decreases the generality of the manuscript. Is this distinction unique to plant-animal networks? What would be the equivalent in food webs? Competitive networks? The hypothesis that the A:P ratio would change is testable. I went to the IWDB, and looked at the number of top/bottom species from all bipartite interactions, presented as median, mean, standard deviation (of P/A):*

Int.	type	med	mean	stdev	host-parasite	0.25	0.27	0.05	pollination	0.34	0.71
	seed dispersal	0.58	1.25	1.55	ant-plant	0.94	0.96	0.50	herbivory	1.16	1.19 1.29

Maybe these can be discussed to give the section a more quantitative data-based vibe.

I apologize but I don't see the point with this query of the referee. In the text I merely discuss the fact that a number of available empirical networks show variable and distinct A:P ratios. And I discuss this in relation to sampling issues: i.e., whether we proceed to fill the adjacency matrix row-wise ("zoocentrically") or column-wise ("phytocentrically"). A quantitative analysis of this point is clearly beyond the scope of the study.

- *In parallel, the idea that bipartite matrices should be symmetric is derived*

from the unstated assumption that most species should be specialists or that most matrices should be (close to perfectly) nested – because these are the two situations that are the most likely conformations to yield symmetric matrices (this again, can be tested within minutes). Since species are no more perfectly specialists than matrices are perfectly nested, maybe the asymmetry of most matrices should not be surprising.

The previous point does not imply that the matrices should be symmetric. All the discussion of zoocentric and phytocentric approaches is just a reflection of how the available methods to record interactions may affect the form of the adjacency matrix.

- *As an aside, on the part about using DNA to study interactions, you may want to discuss/mention Evans et al. (2016) from this special issue.*

Thanks you for pointing out this reference. I had not access to this paper; it's now cited- as well as Bartomeus et al from this same SI.

- *On accumulation curves*

- *This approach has been used before in bipartite networks, by e.g. Poisot et al. (2012) and Gilarranz et al. (2015). I would also remove (p. 15) the reference to a personal communication by Nick Gotelli. This is not something readers can validate / reproduce and they have as such no reason to trust this information.*

Yes the approach to use accumulation curves for interaction sampling is even earlier than those references cited by the referee. These earlier references are cited in the text (Jordano 1987, Jordano et al 2010, Chacoff et al, etc.). Reference to verbal comm. deleted.

- *On estimating interactions and novel methods*

- *I was a bit dismayed not to see many mentions, throughout the manuscript, of novel approaches to infer / predict interactions. I do not believe that natural history is the only solution to this problem. There have been recent work, starting with Bartomeus et al. (2016) (from this special issue), but also Crea, Ali, and Rader (2015) (mutualistic interactions), Gravel et al. (2013) (food webs), Canard et al. (2014) (host-parasites) and even to a degree Eklöf et al. (2013) (food webs). These are all data-based methods that formulate predictions of interactions based on external data (traits, existing network structure, and abundances). These methods have a role to play in the prediction of real missing interactions.*

The theme of inference/prediction of interactions is out of the scope of this study, which focusses merely on sampling issues. Most of these studies deal with the issue of inferring a network, which is a different question to assessing the completeness of an empirical network given a sampling effort. This is the reason why I intentionally omitted this line of research. Yet some of the issues were already discussed in the section discussing mixture models. Reference to Bartomeus et al. (2016) (from this special issue) has been added.

• *I'd also like to bring the recent paper by Poisot et al. (2015) to the author's attention – there are ways, albeit not entirely mature ones, to Stop Worrying and Love the Uncertainty: maybe these methods can be used when informations about how likely interactions are to be observed are available.*

Please see my response to the previous point. Obviously there are strong space limits to discuss some important issues that those suggested by the reviewer; some of these points are marginal to the main theme of the MS, yet proper references (by Canard, Eklöf, Bartomeus, etc.) have been cited.

$$\begin{aligned} &\text{begin}\{\text{displaymath}\} \\ &\text{end}\{\text{displaymath}\} \end{aligned}$$

I very much hope you consider this revised version acceptable for publication, given the editing done and my above responses to the reviewers' queries.

Sincerely,

Pedro Jordano