

Senior Editor Comments to Authors

I have now received the reviewers' reports and a recommendation from the Associate Editor who handled the review process of your resubmission. Copies of their reports are included below. Based on their evaluations, I regret to inform you that we are unable to publish your paper in *Methods in Ecology and Evolution* in its current form.

However, we would be willing to consider another resubmission, which takes into consideration the new feedback you have received. You will see that several comments really ask for a structural rework towards the core of the "Perspective" format of MEE, which is about the stimulation of scientific debate and the offer of conceptual advances. While the ms is giving some of that, it does not convince readers and reviewers even after 2 reviews.

We are deeply grateful for the opportunity to revise our manuscript. We have conducted a thorough round of changes. Notably, we have expanded the example greatly, by adding new data from the *PanTHERIA* database. We have re-organized the section providing an overview of embedding techniques, and reworked the table to highlight differences between the various methods. We have re-worked the conclusion to frame it around what we believe to be exciting emerging questions, which we hope will stimulate discussions around this topic moving forward.

I think that reframing how it is presented would go a long way to appeal to readers and current & future users of graph embedding. It is not usual for us to offer the possibility to resubmit after a first "reject and resubmit" decision when the reviews are of that extent and we did so because of there is a potential for the ms to have all these comments (mentioned below) addressed. A new version including all these (numerous) comments would still need much work but would really allow the ms to be a useful roadmap towards predicting metaweb (even if that means going above the word limit).

We apologize in advance for taking the last comment at face value, and going over the word limit. This was, in part, necessary to bring several elements of nuance that are present in the literature cited, but that have been asked by reviewers to be part of the text. Notably, we added an explanation of why the result of an embedding is not a functional trait. We have also re-worked the conceptual figure to give it more generality. Furthermore, we had to clarify the relationship between embedding and dimensionality reduction. All of these changes required to expand the text at several places, but we are confident that this brings the manuscript closer to the reviewers' vision.

In addition, there are several comments from the previous round that were not addressed but still created issues with the second revision; dealing with them

and using the additional perspective of this new revision will help move the ms forward.

As far as we understand it, this comment stems from Reviewer 3's in-depth reading of the original submission, our responses, and the revised submission. We do appreciate the effort invested, but this resulted in a review where it was often unclear which material was critiqued, and therefore left us with little in the way of actionable items. There are a number of claims by reviewer 3 that we dispute, and we have added our detailed responses below. It was our impression that all points from the previous round of comments had been addressed, but we would appreciate specific guidance as to which were not properly so, in order to fix it should this situation indeed present itself.

Reviewer 3 Comments to Authors

General comments

The manuscript is a resubmission of a previous version that I did not review nor see. I understand that the manuscript is proposed for publication under the "Perspective" type in MEE. Besides the current version, I had access to the first reports and the authors' answers.

We thank the reviewer for the obvious effort they put in their comments. Yet, we do want to express concern (which we will do more explicitly in the following responses) that what the comments are about is often unclear; it has been sometimes difficult for us to understand whether the reviewer was expressing criticism about the revised manuscript, our previous set of responses, or even previously published manuscripts from this working group. Nevertheless, we have done our best to take the comments at face value and address them in depth in the revision, whenever they were about the current submission. Whenever the reviewer identified issues with previous material that were also relevant for this submission, we have similarly worked hard at integrating them. We hope that the editor will appreciate the difficulty of crafting a thorough response when the reviewer comments bear on material that is not currently under consideration.

I think that the contents of this Perspective manuscript do not - as required by the journal guidelines - "stimulate [much] scientific debate". Neither does it significantly "offer conceptual advances or opinions or identify gaps". I'll now try to motivate this opinion.

We think this criticism is unfair, as we had already outlined in the responses how the *original submission* introduced novel ideas,

and notably identified knowledge/methodological gaps. We have re-worked the discussion of *this resubmission* to frame it around what we believe are important issues, including examples of analyses that can be carried out to start working on these problems. As such, we are extremely confident about the novelty of the material presented in *this resubmission*.

In their answers, the authors “point out that the submitted version of this present manuscript includes” 3 different elements. It is unclear to me whether they consider this to be the complete list of the manuscript contributions.

This specific comment in the *initial responses* was intended to highlight areas that were already in the *original submission*, specifically areas that constitute the “advances to the scientific debate” the reviewer asked for in their previous point. We used the word *include* here to unambiguously convey that this was not an exhaustive enumeration; as we are criticized on a matter of semantics, we wanted to make this unambiguous. We would like to note that a description of what the *current submission* contains is given in the abstract of the *current submission*. As the reviewer has not formulated actionable critiques about the manuscript (such as, for example, identification of areas needing a better description of what the manuscript brings), no changes were made.

Anyway, let me discuss the first of these elements. It is “an overview of embedding techniques and their application to species interaction networks”. I do not agree with this first claim. In my opinion, Table 1 that lists a large number of graph embedding techniques is not sufficient to be considered an overview.

We have edited the text to clarify that the take-home message from this table is that a large number of these methods have *not* been applied to species interactions networks, and clarified which have been applied to biological associations (without being about biotic interactions *per se*).

The authors barely address the difference between nodes embedding and graph embedding: while I can clearly see the second column of the table, the text does not contain any sentence that could help a reader not familiar with these embedding techniques to learn about this major difference.

We have added a description of the difference (and overlap) between the categories of embeddings to the main text, as well as in the table. We have further re-worked this part of the text so that the flow is more consistent, starting with why ecological networks can be embedded, clarifying the status of the values given by these embeddings, and then making the explicit point that these methods have been under-used.

The text is even misleading when (line 75) the authors say that “Their [the graph embedding techniques] main goal is to learn a low dimensional vector

representations for the nodes of the graph (embeddings)”.

We have changed this sentence. We are unsure as to whether it will satisfy the reviewer, as they have not indicated how they thought the original one was misleading.

An overview of embedding techniques should go beyond a list of references and come with (at least some lines of) an introduction to the methods and their main differences.

We would like to clarify that the purpose of this paper is to point out that a family of methods (embedding) has been under-used in the study of ecological networks, despite their potential usefulness to be related to additional data for transfer learning. This is now apparent from the sub-titles of this section of the text. What the reviewer suggests is a review or benchmark on each of the methods, which is outside the scope of a perspective manuscript, and more suitable for a review. This is written in the introduction of the *current submission*. Furthermore, there is no concise way to summarize most of these methods in “some lines”. Embedding techniques can be extremely complex, often use vastly different mathematical approaches, and it would be a disservice not to point to the original literature instead.

Going back to the manuscript contributions, the end of the paper’s introduction (from lines 37 to 40), sets a slightly different list: “In this contribution, we highlight the power in viewing (and constructing) metawebs as probabilistic objects in the context of rare interactions, discuss how a family of machine learning tools (graph embeddings and transfer learning) can be used to overcome data limitations to metaweb inference, and highlight how the use of metawebs introduces important questions for the field of network ecology.” From this, I understand that the 3 elements pointed out by the authors in their answer do not constitute the core of their contribution, which is rather described by these previous lines. This leads me to my next remark.

This comment is purely a consequence of the reviewer, with all due respect, over-interpreting the responses to the editor and using them as the basis for critiques of the *re-submitted manuscript*. From this point onwards, it is difficult for us to decide what the reviewer is actually discussing. For the sake of clarity, let us re-state the manuscript “mission statement”: viewing metawebs as probabilistic objects is likely to be very fruitful for their prediction and study, and coupling graph embeddings with transfer learning is a promising way forward, but requires us to solve problems before we get there. This is the last paragraph of the introduction in plain English; no changes made specifically in response to this point.

The previous reports raised the concern that the current manuscript “adds little to the previous published paper”. The authors answered that “there is not a

single instance where areas of overlap are clearly identified”. I do identify specific places that support the reports concern:

We have to front-load the response to the specific points by stating that we fully stand by our original point, namely that there is no overlap, and that no overlap was identified in the previous round of review. This attempt by reviewer 3 to re-write the history of the previous round of comments is baffling. We would like to mention to the editor that we do not understand the point of the following comments. Nevertheless, we are taking the process of peer review seriously, and have engaged with all of these comments in good faith.

The contents of the paper as described at the end of the introduction match the ideas underlying the previous published work;

We re-read both manuscripts in full, and the only “overlap” that we could find (keeping in mind that the end of the introduction of the *re-submitted version* was modified following comments in the previous round of review) is that some of the same general ideas are discussed; this is unavoidable in order to convey foundational ideas that justify the use of embeddings coupled to transfer learning: the existence of conserved backbones, invariants, or action of phylogenetic/trait-based processes. We would like to note that the reviewer is not identifying a problem here, simply stating that articles on similar concepts cover similar conceptual grounds in their introduction.

Section “Graph embedding offers promises for the inference of potential interactions” is not only “an overview of embedding techniques” but rather it is biased towards “their application to species interaction networks”, which is the topic of the previous published paper.

This is correct; it is, in fact, stated as early as the title of the manuscript (to wit: “Graph embedding and transfer learning can help predict species interaction networks despite data limitations”). Again, this comment is based on the *responses to reviewers* from the previous round, and has no actionable items related to the *current version* of the manuscript. The point of the *current submission* is to provide a higher level *perspective* on the topic, of which the *previously published paper* was a specific case-study. The novelty of this submission is in (1) the broader and more systematic coverage of the literature on embeddings, (2) a discussion of transfer learning in a more systematic way, framed by the “in sample”/“out of sample” distinction we added to Fig. 1, and (3) an enumeration of what we believe are important questions for the field to address.

Graph embedding techniques are promising for ecological networks, obviously not only in the context of inferring a metaweb. This leads us back to the main concern with this paragraph: that the authors do not even have a sentence about the major difference between node embeddings and whole graphs embeddings

shows that they are not interested in these techniques for ecological networks in general.

We would appreciate that the reviewer refrains from commenting on our “interest”, when we have been thinking about/working on/publishing with these techniques for several years. We have added information about the difference between different types of embeddings, which we hope will help the reviewer evaluate the seriousness of our interest in the topic.

Figure 1 is titled “Overview of the embedding process”, which describes only part A and B so half of the scheme. The remainder is concerned with the method from the previously published paper.

The last part of this comment is untrue, but we have lost interest in tracing comments back to discussion of the past round of reviews. The figure describes a general cycle of input/embedding/transfer/inference, with specific mentions of potential predictors to learn the embedding. We have changed the title of the figure, and based on constructive comments from other reviewers, have re-drawn it entirely.

The third element of the list pointed out by the authors in their answer is “a discussion of the remaining technical and methodological challenges associated with this approach”. Here I understood that “this” refers to “prediction through embeddings” and the title of the last section (“The metaweb embeds both ecological hypotheses and practices”) is slightly misleading as its contents are rather biased towards inference of a metaweb (after its embedding).

This is, again, a comment on the *responses to reviewers*, which requires no change to the *current submission*. Based on constructive comments by other reviewers, the last section of the manuscript has been re-organised, and the text should now flow more naturally between these emerging issues.

For the authors (line 172) “The first open research problem is the taxonomic and spatial limit of the metaweb to embed and transfer” and (line 189) “The second series of problems relate to determining which area should be used to infer the new metaweb”. These two points could have formed a perspective in the previous published paper.

We wholeheartedly agree, but they had to be cut to adhere to space constraints. No changes requested, no changes made, intent of the comment still unclear.

The last part (from line 203, “praxis of ecological research”) opens to more general considerations but again contains specific remarks related to their previous work: “Applying any embedding to biased data does not debias them” (line 206); “the need to appraise and correct biases that are unwittingly propagated to algorithms when embedded” (line 215).

We dispute this comment. The ideas contained in this paragraph have no relationship to our previous work, and this should be obvious by examining the literature cited. No changes made to the manuscript.

The second element of the list pointed out by the authors in their answer is “a discussion of the properties of metawebs that make them amenable to prediction through embeddings”. I did not clearly identify which part of the manuscript corresponds to that element. I can only suppose that this refers to the paragraph that “a metaweb is an inherently probabilistic object”, but as far as I understand, the paragraph does not make a clear link between these properties and the amenability to prediction through embedding.

This is, again, a comment on the *responses to reviewers*; we hope that the changes we made in response to reviewer’s 5 very constructive comments on the text of the actual manuscript have adequately addressed it.

While I am convinced that “a metaweb is an inherently probabilistic object” and found interesting the part of the manuscript between lines 42 to 57, I did not understand how this is combined with the second half of this section (namely that “high quality observational data” can be combined “with synthetic data coming from predictive models” to “increase the volume of information available for inference”). More precisely, that ‘[the metaweb] fixes an upper bound on which interactions can exist’ is not clearly improved by a probabilistic version of this metaweb.

This section has been modified following comments from other reviewers, specifically by addressing the status of documented v. undocumented interactions, and how probabilities outputted by predictive models can help bridge the gap.

The manuscript points the need for the construction of metawebs at large spatial and taxonomic scales. The authors are not specific about what “large” is exactly. It would be interesting to be more specific on that or provide some examples. Is this a world-wide scale? A continental scale? Any scale for which aggregation of local data is necessary? Anything else? Line 192 (and below) appears the mention of “country level”; However countries are too heterogeneous in their sizes to answer my point. Also, the term ‘continental scale’ appears on line 217 but in a specific sentence and I am not convinced that this is exactly what the authors have in mind when mentioning “large” scales.

This has been addressed in the revised discussion.

The abstract contains the sentence (point 4) “[we] discuss how the choice of the species pool has consequences on the reconstructed network”. This is indeed an interesting question. But I did not see anything in the text that could refer to this.

See response above.

Minor comments

line 10: “accurate predictors are important for accurate predictions”. Indeed, but what is your point?

Our point is given in plain English immediately after the comma: “the lack of methods that can leverage a small amount of *accurate* data is a serious impediment to our predictive ability”.

line 13: replace GBIF and UICN by the full names.

Fixed.

line 73: “Graph (or Network) embedding (fig. 1)”. You should modify the reference to “fig1. A, B” because the rest of the figure is not part of the embedding process. By the way, the caption of figure 1 “Overview of the embedding process” is also misleading as (again) this title only describes half of the scheme.

Changes required by the new figure made.

I understand that the paragraph on GNN (from line 88) was added to answer a referee concern, but it is disproportionate: you use as much space not to speak about GNN as to speak about (ML) graph embeddings.

We added this paragraph in response to comments from a reviewer, and we agreed with the original point raised by the reviewer, specifically that readers may be curious about GNNs; no changes made.

I understand that the illustration of metaweb embedding was added to answer one of the referees of the first round. Nonetheless, I do not see the added value of it.

We hope that the expansion of this illustration will make its added value clearer. See responses to the other reviewers for more details.

On line 147, the authors claim to see “an inflection point around 25 dimensions”. I do not see any inflection, but I understand this is a reasonable compromise.

We have expanded the illustration, and now use the finite differences methods to identify the inflection point.

line 219: “Particularly on Turtle Island and other territories”. I did not understand why in a very general paragraph you refer “particularly” to this specific example. Maybe “for example” would be more suited.

Paragraph changed during the revision of the discussion.

Reviewer 4 Comments to Authors

General comments

The paper summarized the key challenges of inferring metawebs based on graph embedding approaches. It also highlighted the significant advantages of using graph embedding and transfer learning techniques for species interaction network prediction and other ecological problem applications. The paper provided a very important research direction of applying advanced graph embedding and transfer learning to tackle diverse inference tasks for species interaction networks.

We thank the reviewer for their kind words.

Two main questions about this paper are listed below.

Fig.1 is a good diagram that shows the whole pipeline with the input graph adjacency matrix, output graph embedding and combined with transfer learning technique. I would also suggest the authors to include some experimental results based on graph embedding and transfer learning for specie interaction inference with real dataset.

We have greatly expanded the illustration, and as a result have split it into two components: one showing the more statistical considerations, and the other showing the ecological ones.

Please update the article information for the paper (Xu, M.. Understanding graph embedding methods and their applications. SIAM Review, 2021) in the reference section.

Thank you - we now cite the published version rather than the version on arxiv.

Reviewer 5 Comments to Authors

General comments

This manuscript consists in a revised version (re-submission) of a perspective paper dedicated to the potential contribution of graph embedding to metaweb prediction.

The authors provided a pedagogical illustration on a host-parasites system, aiming at predicting (in a probabilistic way) the links of this bipartite network using Random Dot Product Graph embedding. This illustration implements key elements of Fig. 1 and might invite the reader to use or develop the embedding framework on various datasets. The figure showing the decrease of the loss with the rank of the embedding is particularly welcome since it shows to what extent network structure can be reasonably summarised in few dimensions using a specific embedding.

We would like to thank the reviewer for pointing out that the rapid decrease in L_2 loss is a good thing; we have added this information to the description of the case study, and further clarified this point by presenting the L_2 loss (for interaction/pairwise level quality) and the cumulative variance explained (for network-level quality), as a pre-requisite for an ecological interpretation. We had to split the figure in two following further points by the reviewer, which resulted in new information in the figures and the text. We are extremely grateful for the improvement this resulted in.

However, I got a bit surprised not to find the ecological interpretation of this embedding in terms of response and effects traits. I think the authors should better try to link ecological theory in general and ecological hypothesis associated with machine learning methods throughout the manuscript. I understand that the manuscript is centered on methods to predict metawebs with somehow incomplete sampling or knowledge. Consequently, as many research papers on applied machine learning in ecology, ecological hypothesis and theory are a bit behind the scene.

This is a fair point, we thank the reviewer for bringing up this concern; we have added a paragraph to the section on embeddings, to explain why we do not think the values outputted by the models *are* traits. To summarize: they are not directly related, as causes or consequences, to organismal fitness. This does not prevent their use in predictive settings (indeed, we have clarified this at several places in the text, with examples), but it would be problematic to assume that they are “functional” traits. We think that *latent values* is a more appropriate descriptor, and is how we refer to them throughout the text.

I think that a perspective paper should clarify possible ecological interpretation of machine learning methods. If the manuscript is clear and enlightening on the probabilistic metaweb approach, ecological hypothesis associated to link predictions are much more obscure. The following points should be somehow addressed in the manuscript to get additional perspectives:

What are the interpretations of Random Dot Product Graph embeddings in terms of latent traits?

In addition to the paragraph we added in response to the previous point, we have decided to split the illustration figure in two, in order to show how the position of species in the embedding sub-space relate to ecological information. Specifically, we extracted body mass information from PanTHERIA, and show that the position of hosts alongside the first (main) dimension of the embedding is predictive of parasite richness, but not of body mass; we also tie this result to existing literature. This is, in a way, a comforting result: there is information in the network that is not contained in body-mass,

and approaches that can use both (such as embedding and transfer learning) are therefore likely to access *more* information about what makes interactions.

What are the hypothesis behind link prediction using other information (traits, phylogeny as in Strydom et al 2022)?

We thank the reviewer for asking this important clarification; we have greatly expanded the section on the use of node metadata to clarify which hypotheses are associated to broad families of predictors, alongside references to published literature showing the importance of these predictors. We would like to point out that doing so goes beyond the scope of the previously published paper, and is a more general overview of ‘picking the right predictor for the right hypothesis’, up to and including hypothesis testing, as we now mention in the text. This is now additionally reflected in the conceptual figure, which separates ecologically relevant predictors from the rest of the latent variables used in the process, which should clarify what is an “ecological” data and what is created through the embedding process.

What structures are considered in the different embeddings? Table 1 mentions several embedding techniques and separates node from graph embedding. However, even two node embeddings algorithm can have different interpretations. For example, tsne is based on neighbor (local structure) whereas node2vec relies on random walks, so paths in the network (global structure). Both could be interpreted in ecological terms.

This is an important clarification to request, and we thank the reviewer for bringing it up. The strength of embeddings is that they are robust to network structure *and* remove the need to know specific information about the network beforehand. We have added this information in the section on embeddings, but also discuss it (with new references to the literature) in the specific context of RDPG in the illustration. Some methods (RDPG among them) can preserve structures at different scales, perhaps giving some clues as to why they perform well on the few examples they have been tried on so far; this is speculation, and not part of the manuscript.

Predicting metawebs with these two methods do not hold the same hypothesis on species interactions. How does it affect potential applications? I think Table 1 must be enriched more (maybe add a figure or split it) in order to be a roadmap and not simply a catalog. It should mix and bridge embeddings, ecological interpretations and even some illustrations to guide the reader in this machine learning jungle. Such clarification should be also present in the manuscript

We have pointed out in the text, in a few places, that examining the “best” embedding method for a given problem is important; this is coming from existing literature. One issue with what the reviewer suggests is that the embedding techniques are not always easy to

map onto ecological considerations. This is a good thing; indeed, this is because these techniques have been developed for general problems on graphs, and ecological networks are very likely to fall within the range of graphs that these methods will handle (in fact, this is demonstrated by the references in the table, specifically the column on applications). We think that one way this table can serve as a roadmap is by putting more emphasis on what has *not* been done, *i.e.* the majority of embedding techniques have not been applied to species interaction networks. We have updated the legend of the table to make this point stand out more, and given this argument a sub-section in this part of the manuscript.

To what extent co-occurrence data should be considered in embedding approaches? Co-occurrence is considered for statistical association networks but probably not in the same manner than the deep learning model of (Strydom et al 2021). So, to what extent co-occurrence data can be used to predict interaction? Table 1 mentions statistical association methods and JSMD. The authors should be more clear in Table 1 and in the manuscript on the link/differences between association methods (that predicts associations from co-occurrence) and link prediction using embeddings. The term statistical association is present in Table 1 but not in the text, it can be quite confusing for the reader.

We have clarified this point in the table, by explaining that some methods have been used for network-like problems, without being applied to species interaction data. Whether predicting co-occurrence qualifies as a network-like problem is a contentious point (and one we, mostly, disagree with), but the important point is that these methods can get data from node properties, and generate predictions that are, so to speak, network-shaped. We also touch upon the issue of co-occurrence in the illustration, and have added references highlighting why co-occurrence alone is not a good predictor to the introduction. Nevertheless, it *can* be added to the predictor data, and we have updated the conceptual figure in this way. When approaching this methodology from the perspective of different sources of information representing different hypotheses on what drives interactions, adding information on co-occurrence can help disentangle its role.

I think this manuscript could be considered for publication in *Methods in Ecology and Evolution* but it must offer broader perspectives than Strydom et al. To do so, the authors should try to address somehow the previous listed points.

We thank the reviewer for their encouraging words, and for the suggestions that led to the revision of the illustration, and indeed most of the manuscript. We hope that the revised version will meet their expectations.

Minor comments

L20: Local networks capture alpha-diversity of interactions but also beta diversity since interactions from the metaweb can be absent from local networks species absence.

This is correct, but does not seem relevant to the paragraph; no changes were made as we are already past the word limit.

L24: Yes, I agree that Saravia et al. 2021 shows that local network structure (represented by network metrics blind to species identity) does not differ from the one expected from a null model (Trophic Theory of Island Food webs). However, the manuscript focuses on alpha-diversity metrics. In terms of species and interaction composition, they can still differ even if they have more or less the same structure. It think this point should be clarified. Indeed, otherwise, it somehow states that we do not need to focus on local composition since local networks have the same structure as the metaweb.

This is correct, and we addressed this point in the original submission, in the next paragraph (which is focused on backbones of networks). No changes were made.

L63: Does it “generate” or uncover the core rules associated to species interactions? I do not really understand the point of “generating rules”. For me, statistical methods try, using abstract representations, to uncover/formulate biological rules.

This is a good point, and we thank the reviewer for pointing it out. We have rephrased this sentence to clarify that observational data are used for the inference of rules, and that the resulting augmented dataset is used for analysis.

L79: Yes, embedding methods can exhibit structural invariants in ecological networks. However, what characteristics of the networks should be considered in the model (links, paths, motifs...)?

This is an interesting point; we would like to apologize, first, for the lack of clarity, as we were talking about structural invariant in *networks*, not in their embeddings. We have clarified the paragraph. Second; the invariants in structure are not necessarily revealed by the usual network measures, but are indeed *structural*, *i.e.* they can be seen through the shape of the network. We have clarified the paragraph to make it clearer that the lower dimensions of an embedding should capture these invariants, and improved the transition to the work of Eklof et al. showing the low-dimensionality of food webs.

L83: If the choice of the embedding matters for the result, why not providing to the reader some sort of roadmap of these embeddings techniques for different applications?

The literature suggests that the algorithm should be selected on a problem-specific basis, and we have changed the paragraph to reflect this. The table has been changed in response to other comments.

L85: For the moment, Table 1 is more a catalog than a roadmap.

This is correct – we are unsure where the notion of a “roadmap” comes from, as we label the table an “overview”, and do not argue in the manuscript that the table can serve as a roadmap. Nevertheless, we are confident that there is value in the table as it is, as our survey of methods revealed that a number of them have not been used in network ecology; indeed, a number of references in the table are network-adjacent, as they describe the use of embeddings for JSDMs or statistical associations. In keeping with the response to the previous comment (the need for a case-by-base appraisal of embedding techniques), this is particularly problematic; we have edited the caption of the table to reflect this.

L86: Ok, here comes different network representations (latent traits or random walks). How do we interpret these methods in ecological terms? Which one is the most suitable in the various potential applications.

As we have mentioned in other places in the manuscript, this is context-dependent, and existing literature suggests that one should explore different embedding techniques.

L99: Bohan et al. 2017 mentions several associations methods to build networks. If I am correct, these associations methods are not mentioned in the manuscript. The authors must clarify this point.

The methods in Bohan et al. are not based on embedding (and deal with specific categories of datasets that are far from widespread anyways); we have not edited the manuscript here.

L115: The authors should clarify the ecological hypothesis associated to such approach and the potential limits of predicting interactions using node metadata.

We have expanded this paragraph significantly, by explicitly stating what the assumptions in using these predictors are, and further by showing how the relationship between phylogeny/traits and embeddings can be used in the context of testing hypotheses about the drivers of the structure of networks.

L134: I think this illustration is relevant but it must be enriched in order, for example, to compare different embedding or incorporating meta-data. To what extent will different embedding techniques provide different networks? Will the loss with the rank be similar if you use link (as in the RDGP model) or path based embedding?

We have greatly expanded the illustration, notably to bring into consideration more ecological elements. The potential to perform

the analysis that the reviewer suggests is limited by data availability, and would be outside the scope of this manuscript.

Moreover, Stochastic Block Model can also be used to perform link predictions (Gaucher et al. 2019; Link Prediction in the Stochastic Block Model with Outliers. stat.) and is related to RDGP, I think it deserves a mention here.

We have added a discussion of this article in the section on metawebs; an important feature of embeddings is that they do not assume a specific structure for the network in order to capture emergent features.

L136: What are the ecological hypothesis associated to RDGP model in terms of response and effect traits?

We have addressed this comment by adding a paragraph about the status of latent values w.r.t. response/effect traits, which we can summarize as “the latent values should not be discussed as functional traits”. Furthermore, we have added a justification of why RDGP is an interesting embedding technique in the first paragraph of the section on the illustration.