

Dear recommender of PCI Ecology,

Thank you and two anonymous reviewers for the careful reading and insightful comments on our previous version of our manuscript. We have now extensively addressed your suggestions, and we are glad to submit a new version of the manuscript entitled “Three points of consideration before testing the of patch connectivity on local species richness: patch delineation, scaling and variability of metrics.”.

This new version has been made available on the BioRxiv preprint server at <https://doi.org/10.1101/640995> and we provided supplementary material (codes and virtual datasets) in the online repository <https://doi.org/10.5281/zenodo.3756712>.

We followed the recommender advice about improving the general structure and created a clear architecture based on three predictions, and we used it from title to discussion. We further addressed all the major and minor comments, as detailed in bold below. We also deeply simplified the last part of our analysis about how the variability of connectivity indices among sampled sites affect their explanatory power on species richness, which led to making a new Figure 4.

We hope that this new version will appear more structured and brings simpler and clearer messages to interested readers.

Thank you for your time and consideration,

Yours sincerely,

On behalf of all the authors, Fabien Laroche

DECISION BY RECOMMENDER

The revised pre-print has now been seen by two independent reviewers, one of which had reviewed it before. Both acknowledge the potential merit of the study but also identify points that need revision. Also in my opinion, the pre-print has improved substantially. Nevertheless, I would recommend another round of revision to improve the presentation and clarity but also resolve some methodological questions. Below I provide some thoughts additional to the reviewer reports. Many thanks for considering PCI Ecology.

First, I highly recommend shortening the introduction and re-thinking the structure of it. An introduction should generally start broad and become more specific. In the current pre-print the authors don't have a coherent story arc in the introduction yet that guides the reader from broad to specific. For example, the second paragraph L 39-48 is already very specific while the following paragraphs become broader again (for example, the next paragraph mentions another concept - the functional connectivity - without explaining it in detail). It is thus still a little difficult identifying the story thread here. As a suggestion, you could first make the point that some studies did not find support for the TIB or were questioning its validity, and that you argue that this may be related to methodological choices and to landscape properties. Then, I would actually expect the resolution of the introduction around L 89-106 while here you start with a whole new story line, discussing patch delineation, the raster perspective, γ -diversity, and connector indices. So, basically I suggest shortening the introduction by 30-50% and restructuring, highlighting only those aspects that are necessary to grasp the main objectives of the study.

We followed the insightful suggestions of the recommender regarding the general structure of introduction. The former introduction was 2092 words long. The new version is 1126 words long.

Second, the objectives of the study and the general workflow are still a bit vague/hidden. I would suggest describing a clear road map at the end of the introduction or at the beginning of the methods. What exactly is being tested and when? This may also help following the description of the methods section. In their response letter, the authors argued that they now used the sign posts "patch delineation", „index scaling“ etc. in all sections. However, I don't think this works very well in the methods part. Here, you should very clearly state what exactly is being tested (objectives) and how the corresponding analyses are done. At the moment, this is a bit confusing. For example, the authors state that they first test connectivity indices on species richness and store the resulting R2 (L 292-305), then analyzed R2 in relation to resolution and fragmentation of the landscape (L306-320), then analyzed the effect of connectivity on species richness for fine patch delineation (L 321-323) followed again by analyses of R2 (L 325-329) and so forth. This part could be much more condensed by clearly separating the simulation scenarios from the subsequent analyses.

We clarified the roadmap by setting three clear and precise prediction in the introduction (ll. 95-101). We then directly used these predictions as signposts in Material&Methods, Results and Discussions. We clearly indicated in corresponding Material&Methods sections how we tested each prediction. We clearly indicated in corresponding Results sections what was the outcome of each test. We clearly indicated in corresponding Discussion sections how we interpreted the outcome of these tests and what perspective we drew from it.

Third, the results do not describe the simulation output at all. When using virtual data, the simulated patterns should be described before using these as test grounds. If I am not mistaken, in this pre-print, the authors aim to test under which circumstances missing connectivity effects on species richness can be interpreted as divergence from TIB principles and under which circumstances the results are affected by methodological decisions. So, the first question is whether connectivity does affect species richness in all scenarios. What species richness patterns were simulated and how were these affected by dispersal? I think making clear what patterns emerged from the simulations and which of these patterns were picked up by the connectivity indices, is vital to answering the questions under which circumstances TIB is valid and under which circumstances this could be picked up by connectivity indices.

We clarified in the method section what was the output of simulations : first we made it clear that the abundances of all the species in all the habitat cells is recorded for each simulation (ll. 151-152); second we presented in a detailed separate section how the metacommunity was sampled to create a smaller dataset (ll. 186-191) where species richness was computed from the raw community composition data. Then we added a first paragraph of the result section (ll. 288-291) giving a general overview of whether connectivity is globally related to species richness in our analyses. We commented this general overview in Discussion (ll. 403-407).

Fourth, the discussion should start with a paragraph summarizing the overarching objectives and findings before discussing each result in detail.

We tried to follow this advice but the paragraph ended up being highly redundant with the beginning of each Discussion section. We therefore reckoned that the improvement of structure induced by prediction-based signposts were sufficient to orientate the reader through the discussion. In addition the beginning of Discussion is now less technical than it was in the previous version of the manuscript, since we added a first part commenting the general output of the simulations, as explained above.

Additional line comments

L 50: the author(s) of Ref 6 should be spelled out here as it is used as subject in this sentence. Same in L 83, L 94 (and more probably more instances).

Done (l. 48 and throughout the rest of the text).

L 52: edit „ca. 20%“

We rephrased (l. 50).

L 55: I feel that „functional connectivity indices based on surrounding populations“ as well as the „multiple life stages with contrasted requirements“ should be explained a bit more.

We added a longer explanation between parentheses (ll. 63-64).

L 285-291: Which error distribution was assumed and how was R2 computed?

We used a linear model with a quadratic predictor, as clarified at ll. 205-211. The R2 is therefore computed using standard routines from the *lm* function of R *stats* package. For readers that needs all the details of our methods, we provided the R script of our analyses in an online repository with doi 10.5281/zenodo.3756712 (ll. 285-286).

With species richness as response, you would need to use a generalized linear model (rather than a simple linear model) with a log-link.

It depends on the aim of the analysis. We explain below why we believe that a linear model is adequate and sufficient for what we aim at doing in our analysis.

Our aim here was to determine what fraction of the variance in species richness among sampled cells can be captured by a quadratic function of connectivity. This is a partition of sum of squares problem. Partition of sum of squares is a very general descriptive statistical technique, quantifying the connection among variables. It is used in a broad array of techniques, like PCA for instance, and can be applied to integer variables as well as continuous one.

We use a quadratic function because it is among the simplest non-linear forms that can fit a marginally decreasing effect of connectivity, as discussed in an earlier round of review.

Using the framework of statistical linear models is here purely instrumental and allows us to find the quadratic function maximizing the fraction of species richness sum of square explained. From these models, we only used R2spec, which summarizes the partition of sum of squares attained. For instance, confidence intervals around parameters estimates are probably not reliable, and we do not interpret them in the text.

If one wanted to compare, for instance, the quadratic functions obtained from distinct simulated datasets, then using a statistical framework tracking more precisely the discrete nature of species richness would be necessary. A general linear model would then be appropriate, as suggested by the recommender. To follow up the discussion, we believe that classic models like Poisson or Negative Binomial would not be appropriate here, given the apparent under-dispersion of the data (Fig. 1B). We would suggest using a binomial generalized linear model with a logit link function, assuming that the maximum number of species within the regional pool is known. However, we reiterate that it is not necessary here, since we just aim at computing a partition of sum of squares.

In the former version of the manuscript, we had dropped the quadratic term of the regression function in some cases based on a likelihood ratio-test, which was not really statistically grounded,

given what we just said. We removed this step and re-run all the analyses accordingly. Results were not affected.

Generalized linear models do not normally output R². (Same in L 299).

True, we used linear models as explained above.

L 306-308 and L 321-329: Similarly to above, R² is bounded between 0 and 1 and thus violates the assumption of normally distributed errors. A linear model is inappropriate in this case.

Our analyses of R²spec always amount to comparing means among categories. Given the very large amount of virtual data generated in our study, whether the underpinning individual observations of R²spec are normally distributed or not, the central limit theorem ensures that mean-estimates are normally distributed around the true mean of category and that standard errors are interpretable. In addition, there is a large body of literature suggesting that ANOVA-like statistical analyses (and in particular F-tests and t-tests) are quite robust to non-normality, see for instance the following review (the relevant section starts at p. 246):

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*, 42(3), 237–288. <https://doi.org/10.3102/00346543042003237>

COMMENTS BY REVIEWER 1

The manuscript has been deeply revised since the first submission. It now focuses on flux indices (the authors chose to remove patch-based indices from the analyses), which makes it probably a bit less ambitious than the previous version but also clearer and well organized. The streamlining is clearer now.

The neutral metacommunity simulations are better explained even if some analyses remain unclear (Results\$Index scaling and species dispersal, see below). The analyses and discussions are more conceptually grounded in terms of ecology, with an interesting contribution to the relations between cell grain and dispersal distances. This is welcome. The manuscript is fairly methodological/user oriented, which should be useful for whoever wants to use structural connectivity indices (CI).

The text is fairly clear but seems to still include syntax errors (see below, but there may be some left). I would suggest to contact a native speaker to read it.

We carefully read the text to correct syntax errors.

As a conclusion, I would say that this manuscript represents a valuable and interesting contribution to the proper use of CI. The use of simulated metacommunities in virtual landscapes allows getting conclusions that avoid observation bias that could appear in empirical studies, even if it comes with the formulation of strong ecological hypothesis which are well described and taken into account by the authors. The authors have endeavoured to take into account the comments made by the reviewers and make their work clear, even if there are still sentences that could be clarified (see below).

Thank you for this encouraging comment.

Detailed comments

L327-328: we expected the scale parameter (...) to increase

Done.

L 336-337: the sentence is not clear

We clarified.

L340-341: the choice to focus on Buffer index could be justified

We removed this section to improve conciseness and clarity of the manuscript.

L391-396: the way the authors determined the 'optimal scaling parameters' is not clearly explained and remains unclear. The following paragraph is also clumsy.

We clarified how we computed the optimal scaling parameter value at ll. 257-259.

Figure 3 is hard to read

We changed Figure 3 to improve readability.

L471: a dot is missing -> "studies). This would (...)"

Done.

L487: indices (...) that needs

Done.

L525: One may therefore

Done.

L538: compatible with TIB

Done.

L556: an decreasing exponential kernel

Done.

COMMENTS BY REVIEWER 2

The manuscript "When should patch connectivity affect local species richness? Pinpointing adequate methods in adequate landscapes using simulations." addresses the issue that an important principle of the widely applied Theory of Island Biogeography (TIB), i.e. the relationship between structural connectivity of habitat patches and local species richness, is often not reflected in the results of empirical studies. This could become a reason to generally doubt the validity of the TIB, which currently provides straightforward and easily applicable principles. Therefore, the authors aim to identify methods and conditions under which strong relationships between patch structural connectivity and local species richness are likely to occur. To do so, they use simulations of neutral meta-communities in a spatially explicit model to test the effects and sensitivity of a range of factors on that relationship. Specifically, they explore different ways of patch delineation, scaling of patch connectivity indices, index types and landscape features.

In my opinion, the authors take up a very relevant topic and have ambitious and meaningful research aims. The research questions, and specific factors and hypotheses they test are, for the most part, nicely developed and follow logically from current state and gaps of research. The chosen methodology constitutes an adequate way of answering their questions. While their meta-community simulations do constitute a substantial simplification to real landscapes and communities (e.g., homogeneous matrix resistance, equal dispersal characteristics of all species), they allow testing of a wide range of

factors with a focus on specific relative effects. The results of this study can be of great help to others who will undertake quantitative analyses of the relationship between connectivity and species diversity measures.

However, there are several aspects of the manuscript that need revision, which I will outline below. Most of the issues are related to improving the clarity and comprehensibility of the manuscript, and making it easier for the reader to follow.

Overall, it cost quite some effort to read through the manuscript due to (many!) grammatical and language mistakes and/or typos, which considerably hampered the comprehensibility. A thorough language check and revision would help to make it easier to follow, improve the readability, and could thus substantially increase the quality of the, otherwise very interesting and well-structured, manuscript.

We carefully checked for grammatical and syntax errors throughout the text.

LL. 137-146: Here, the research aim should be stated more precisely. Is it about testing whether “a stronger aggregation of the habitat map leads to stronger fluctuation of patch connectivity indices” (which it will very likely do up to very high values of aggregation)? Or about the effect of this on explaining species richness (which seems more interesting)? You could say that you test whether “a stronger aggregation of the habitat map leads to larger effects of patch connectivity on species richness”, and then include the variation of connectivity indices as an explanation of the results.

Following this comment as well as the general comment of the recommender, we now detail our aims and predictions (predictions 1-3) in Introduction.

You stated in the Introduction in paragraph about habitat aggregation (LL. 137 – 146) that a possible factor explaining weak relationship between patch connectivity and species richness can also occur in situations where “immigration does not act as a source of species diversity”. It would be good to 1. Provide a reference here to examples from literature, and 2. This aspect should also be picked up in the discussion again.

This part was probably too speculative and peripheral to appear in the Introduction, we removed it as well as corresponding inconclusive results in main text to improve conciseness of the study.

Here, additional explanation/rationale seems necessary:

- I wondered why the Buffer index was not used when testing the effect of Patch delineation. If this measure does not make sense in the case of large aggregated vector patches, this should then still be explained at least shortly (LL. 292 ff.)

We explained that Buffer indices values are not affected by patch delineation by presenting separately the computation of Buffer indices in Methods (ll. 157-159), and explicitly clarifying this point at ll. 219-220.

- LL. 296-297 and 304: would be good to explain shortly why only the highest R2spec values were kept

We explained our approach at ll. 221-226.

- LL. 340-341: Why focus only on Buffer index?

As explained above, we dropped this part of analyses in the current version of the manuscript.

- In the method section, it would be good to provide some more rationale and/or references for the specific settings you use, particularly in the paragraph about “Neutral metacommunity simulations”.

There is no real justification for the absolute values of parameters used in our simulations. Note that we do not specify the true size of a cell either. Here we were interested in selecting parameters with contrasted values, enabling us to observe contrasted patterns, but the absolute values are not interpretable.

- In the Discussion (LL. 546-572) the authors talk about similarities and differences in the performance of three connectivity index types used (Buffer, dF, and dIICflux). The high (and apparently expected) similarity of Buffer and dF raises the question why both of them were included in the study, or at least why this aspect was not mentioned already earlier in the Introduction (potentially with a hint to the remaining differences and why they are still worth testing). Otherwise it comes a bit surprising that two indices are compared which are known to be so highly correlated.

As we explained in Discussion, the tight correlation between Buffer and dF indices was not a well-established consensus in the pre-existing literature. In particular, Miguet et al. suggested that there might lead to distinct empirical conclusions. Besides, even if these are tightly correlated indices, it is important mentioning that the quantitative relationships between the scaling parameter and the species dispersal distance differ between the two types of indices.

There are some issues where I disagree with statements made by the authors:

- LL. 398-402: I would not agree. In Figure 3C there is a marked difference between the peak and neighboring scaling values, even though it is not at the border of the value range

Yes, this was an over-statement. We removed this sentence from results and used a more nuanced formulation at ll. 431-433 of Discussion. In addition, we added information about standard deviation of R2spec in figure 3 to help interpreting the magnitude of mean R2spec between scaling parameters.

- LL. 477-484: I do see considerable differences between R2spec values between low and high dispersal in Figure 4, particularly for index dIICflux, so do not find this statement convincing that “using too fine mesh size is harmless”. In real systems there could be even larger differences in dispersal levels between species, so this could potentially have a large effect.

We removed this part of discussion, which was indeed too distant from what our results can really support.

Add references to LL. 147-149,

This section does not exist anymore.

and L. 329

Done.

LL. 342-355: I had a hard time following your approach here, particularly the rationale for using the residuals of separate models based on Hurst coefficient and habitat proportion again in models that use these same variables, and why dIICconnector was not used more directly in a model to explain R2spec.

We removed this section from the manuscript.

The resolution of the graphs is quite low and made it hard to read them.

We improved the resolution of our figures.

Minor comments

I recommend to format terms used for describing the connectivity indices (Buffer, dF, dIIcflux) or approach of patch delineation (coarse vs fine) in the same style. Currently it is a mixture of writing with and without quotation marks, or using Italics (but not always). It could help to be consistent here to make it easier for the reader to follow

We harmonized terms format throughout the text.

Naming of the connectivity indices: e.g. in Table 1 (L. 275) Flux is used, while throughout most of the manuscript dF is used.

We corrected in table 1.

Consistency of using γ /gamma. LL. 151, 153, 159

Gamma diversity is not mentioned in the manuscript anymore.

L. 251: sub-heading should be "Patch connectivity indices" (to stay consistent with terminology used)

We used consistent subheadings throughout the text (see our answer to the recommender above).

L. 283: not clear how this number 28 comes up (maybe add in brackets how you reach that number here)

It was a mistake, the correct number is 21, and we add the detail of the calculation at ll. 184-185.

Add "type of connectivity index" in LL. 290-291

This sentence does not exist anymore.

Numbers as words or digits L. 294

We harmonized throughout the text.

L. 309: "dependent" has to be "independent"

Done (l. 238).

L. 324: add in brackets here "(4 for dF and dIIFlux, 5 for Buffer)" - this will help the reader to keep track

Done (l. 267).

LL. 391-392: in Fig. 3 A it looks like this should be 4 times (instead of 8)

When species dispersal equals 0.25, 0.5, 1 cell, corresponding optimal scaling equals 2, 4 and 8 cells. Therefore we are confident that it is an eight-fold relationship.

LL- 426-428: Not true that average R2spec is reported for each combination

Now that we added a caption, it is true.

L. 493: make sub-heading consistent with previous chapters "Index scaling and species dispersal"

Sub-headings are now consistent throughout the text.

L. 597: "low" should be "high"

This sentence does not exist anymore in the manuscript.

L. 647: here “fragmentation” is used, while throughout the entire rest of the manuscript “aggregation” was used.

We used “aggregation” throughout the text.

The description of dispersal characteristics of species currently varies between “high/low dispersal”, “high/low dispersal ability”, “high/low dispersal distance” (e.g., LL. 96, 106, 371, 373, 390. As I understand, in most cases dispersal distance is meant (and not rate, or likelihood, or success), so using precise terminology would help to make this clear.

We used “dispersal distance” throughout the text.