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.

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.

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.

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.)
- LL. 296-297 and 304: would be good to explain shortly why only the highest R2spec values were kept
- LL. 340-341: Why focus only on *Buffer* index?
- 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"*.
- 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.

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
- 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.

Add references to LL. 147-149, and L. 329

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*.

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

Additional 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

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

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

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

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

Add “type of connectivity index” in LL. 290-291

Numbers as words or digits L. 294

L. 309: “dependent” has to be “independent”

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

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

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

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

L. 597: “low” should be “high”

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

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