Review PCIEvolBiol #698
Journal: Peer Community In Evolutionary Biology
Manuscript No.: 698
DOI: https://doi.org/zenodo.8198951
Title: Spatial autocorrelation and host anemone species drive local components of fitness in a wild clownfish population
Authors: Pascal Marrot, Cécile Fauvelot, Michael L. Berumen, Maya Srinivasan, Geoffrey P. Jones, Serge Planes, Benoit Pujol
Referee: Stefan J.G. Vriend (https://orcid.org/0000-0002-9006-5988)
Date: 2023/10/19

I congratulate the authors on an interesting study and compliment them on the use of a variety of methods to elucidate the spatial dynamics and adaptive potential of a wild population of orange clownfish. Overall, I think that the manuscript is clear and easy to read for an audience that is not per se familiar with the study system. However, while reviewing this work I came across a few matters that may require more attention, which I will highlight below.

1. Adaptive potential. As you put your study in context of adaptive potential, I had imagined that, after reading the Abstract and Introduction, the analysis would include a temporal component, and that you would assess how clownfish recruitment (probability and number) would respond to changes in the environment. I want to stress that I think that adding an explicit spatial component to an analysis of LRS is interesting and elegant by itself. However, it might be misleading to provide this work in context of adaptive potential, which hinges on this dimension of *change*, which is not part of your analysis. I therefore suggest putting the work more in context of spatial autocorrelation (as is done elsewhere in the Introduction) and less so of adaptive potential, or, alternatively, include time in the analysis.

2. LRS. Interesting that the unit of your LRS measurement is the "recruit" rather than the classically used "offspring". This choice makes sense in relation to the importance of self-recruitment for population persistence as you highlight in L107-108). Up until the Methods section (L175 and further), it was however not clear to me that you defined LRS in such a way. I suggest emphasising this in the Abstract and Introduction. In particular, in L75-82, where you provide a definition of LRS that does not match yours, the manuscript would benefit from citations to other studies that have focused on a recruit-based LRS.

3. Model averaging. I appreciate that you recognise the issue with putting too much faith in any single model and instead draw conclusions from multiple candidate models through model averaging. However, if I interpret the description of the model averaging procedure used in this study (as described in L235-242) correctly, this form of model averaging can be flawed whenever there is multicollinearity among the predictor variables in the candidate models. That is, predicted responses based on the average of regression coefficients [incorrect] will not be the same as the average of predicted responses for each candidate model [correct] (see Cade, 2015, Ecology, https://doi.org/10.1890/14-1639.1). That is, the former will not take into account the covariance structure of predictor variables, whereas the latter will. So, if I understood your methods correctly, multicollinearity in your predictor variables (which is often the case in ecological studies) might be problematic for interpreting your results. One possible way forward, as highlighted by, e.g., , is standardising the parameters. In cases where this does not suffice, model averaging predictions (rather than coefficients) will solve many of the issues (also highlighted here: https://atyre2.github.io/2017/06/16/rebutting_cade.html).

Minor comments:

Title: I believe that the title covers the content of the study well, but I wonder whether it should rather read "spatial autocorrelation and host anemone species drive *variation** in local components of fitness…". *or another word that indicates a 'change'.

Data accessibility statement: Thank you for providing the scripts to Zenodo. However, the data files seem to absent. Could you add these, so that one could run the scripts?

L43-46: I find this sentence hard to read. What do you mean with "within the habitat variable" and "global habitat information"? Please clarify.

L46-48: What do you mean with "local components of fitness" here? Self-recruitment could also be seen as a component of fitness? Or does this refer to LRS? Please clarify.

L52-54: The conclusion that your findings "imply that this clownfish population is susceptible to human-induced […] modifications …" does not seem to follow logically from the findings as presented in L49-52. That is, there is no mention of 'human-induced modifications' in any of the study's aims and or findings. Please make the link between the findings and conclusive sentence more explicit.

L57: I suggest replacing "wild populations resilience" by either "The resilience of wild populations".

L58: Just a thought whilst reading this sentence: perhaps it will provide some context to the reader if you would relate 'adaptive potential' to other widely used terms and synonyms, such as 'adaptive capacity' and/or 'adaptive plasticity'.

L59: "by means of"

L62-63: What do you mean with "harbouring adaptive potential"?

L68-82: Very clear paragraph!

L89-91: I suggest removing this sentence. The next sentence (starting with "Previous work…") would follow well from the previous sentence (starting with "Coral reef fish…").

L93: What do you mean with "that it can be achieved"? I suggest rephrasing this sentence so that it explicitly states *what* Salles et al. 2020 achieved.

L106: What is the rationale behind a percentage of 50%?

L108-111: I take from this statement that Salles et al. 2020 already found out what the ecological drivers of LRS variation in this system are. How is this different from your goal to "disentangle different ecological sources of variation in LRS…" (L104)?

L120-134: Strong paragraph.

L131-139: In these lines, you use "spatial dependency", "spatial distribution", "spatial scale", and "spatial structure", and it is not clear to me whether some of them are used as synonyms or how they relate to one another.

L135-142: I appreciate that you compare you spatially explicit model to a model without spatial autocorrelation.

L136: What do you mean with "depth"?

L145-173: Very clear description of the study system.

L181-182: Please provide a rationale for this choice here, or should I read L184-185 as such?

L195-199: Why did you choose a radius of 200 meters instead of, for example, the mean or the maximum (recruiting) distance?

L199-204: What is the maximum distance that a recruit could potentially travel? In other words, could the difference between recruiting distance and between-anemone distance rather be a result of physical constraints?

L204-210: To what extent are individuals with *S. gigantea* as a focal anemone more spatially constraint in where they can recruit? I imagine that shallow waters (or the shore itself) will inevitably force individuals to recruit to deeper depths.

L209: In L191 you introduce "focal anemone", but here you use "parental anemone". Please stay consistent in the choice of terms.

L213: I suggest citing R as well.

L217: I interpret your zero-inflated model as a zero-inflated Poisson because you refer to Poisson in L214, but it would be clearer if you could specify that here.

L227: "Because no ecological variables *were*…"

L252-255: Why did you choose to use discrete lags rather than a continuous distance measure in your correlograms?

L266-267: Why did you choose to truncate the distance matrix?

L310-318: Would it not be more natural and/or easier interpretable to put your results in terms of the probability to self-recruit and the probability to produce a (self-)recruit rather than the probability to *not* self-recruit and *not* produce a recruit?

L328-331: I think it would be easier to read this section if you would describe the results in terms of "X times more likely" rather than "X the odds".

L342-344: There only seems to follow one component following "a unique combination of", but one would expect at least two, so I suggest rephrasing. Perhaps something like "Our spatially explicit analysis of long-term monitoring data from a wild clownfish population revealed…".

L344-345: Maybe a bit more nuance would be justified. Your study clearly highlights the difference between the spatial and non-spatial models in terms of the number of recruits produced, but for the probability of (not-)recruitment, there is only one effect that seems different.

L353-354: What do you mean with "at all scales" and how could the reader check this assertion?

L419-431: It would be great if you could elaborate on the (ecological) consequences of this finding. Why would individuals have a better recruitment probability in *S. gigantea*? In addition, what other differences between the two locations where the anemones occur could explain the difference in recruitment probability?

L434: Some text appears in red.

L478-483: Interesting. This is exactly the type of environmental change I refer to in my first main point and had expected to be part of the study after reading the context of adaptive potential. As climate change is not part of your study in any shape, I suggest removing the sentence starting with "More investigations are…".

L484-486: This feels rather out of the blue, mostly because it is the first time you mention "long-distance dispersal" (Interesting nonetheless!). I suggest rephrasing the ending of the conclusion section so that it better fits with the conclusions of your work.

Figure 1: Please specify the x and y axes. Please increase the size of the text and the plot itself to increase readability.

Figure 2: The text and labels of this figure seem dragged out of proportion. Please fix this to increase readability.

Supplementary figure: Please increase the text size to increase readability.