

I thank the authors for the care they took addressing my comments. I reckon you did a really good job justifying your methodological choices, so I don't have any more concern regarding the robustness of the analyses.

I still have a divergent opinion though on the inter-cluster comparison, and I will try again to convince you that it might not be helpful here. I would understand if you choose to stand by your initial strategy (and I will recommend the article no matter what), but please hear me out on this because I strongly believe it can improve the readability and impact of the manuscript:

- A main strength of your manuscript, as noted by all the reviewers was that you managed to find an elegant way to analyze an unbalanced dataset. As you clearly explain at the end of the introduction, the clustering method was necessary to analyze **the longitudinal data, which is your main question (as you state it l. 123-124)**. So it would be justified at the end of the clustering analysis to retain only the clusters that include longitudinal data.
- Analyzing different climatic clusters over the same period is quite a different question than the impact of climate warming, and I don't feel like you have any predictions or hypotheses regarding this comparison. Inter-cluster differences might be related to random genetic differentiation or local adaptations regarding other factors than climate, and the absence of predictions make the results uninterpretable in my opinion. Indeed, if you find differences between clusters you can tentatively attribute them to temperature effects ; if you don't find any, you can suggest that the clusters have different thermal tolerance due to their specific adaptive history. It seems to me that if the data can't allow to discriminate between a null hypothesis (absence of impact of climate warming) and an adaptive one, then this specific question should not be addressed with these data.
- On a practical note, results from the inter-cluster comparison are hardly mentioned in the discussion (no mention of cluster 4 at all I think), and do not provide any further insight compared to Analysis (1). The paper is quite long and complex as it is (especially the discussion), and I am convinced there would be no loss of information if only Analysis (1) was reported, and you may still include Analysis (2) in supplementary material. To me the main messages are : (i) mean temperature has increased in cluster 3, but not in cluster 1, and it might be related to a decrease in fecundity; (ii) hatching rate decreased in both clusters, and might be more sensitive to extreme events than mean temperatures. I didn't feel that the specifics of cluster 2 add anything significant to these conclusions.

Minor comments:

- I think the track change version of the manuscript was updated twice instead of the "clean" revised version.
- The discussion is still quite long (over 6 pages). I like it though, I think it's full of interesting ideas and perspectives, but it might help to organize it a bit more explicitly (like, having titles for each of the main parts). Ideally, it should start with

two sections focused on the two main results: (i) response of clutch size to increase in mean temperature; (ii) increase of hatching failure, related to extreme events of prolonged high-temperatures

- Building on the 3rd comment of Dr Ilitis, I wonder whether the results concerning the three potential underlying causes of hatching failures should be detailed fully in the main text (it feels more like supplementary material to me). You can still discuss some of the nuances in the responses of abortion and parasitism to temperature, but it will make the results section (and possibly the discussion) shorter and clearer. For instance, on lines 425-429 (track changes version), the distinction between the 3 internal processes is quite confusing. First, it says that parasitism has increased (though I think it is the opposite, am I wrong?). Then, the relationship between abortion rate and temperature is highlighted, which suggest that abortion rate is the driving factor between the hatching rate decrease. But from the very next sentence to most of the following 2 pages, it is the relationship between extreme heat and sterility rate that is mostly discussed. So the whole logic behind the different arguments is quite difficult to follow, and it could benefit from a bit of further structuring, as suggested from my previous comment. In the same line, Fig 5 could be lightened by including only fecundity and hatching rate.