Associate Editor Comments to Author:  
  
Associate Editor  
Comments to the Author:  
Dear Authors,  
thanks for submitting your manuscript to Ecology and Evolution.  
  
The manuscript "Recurrent heatwaves slow down the recovery of a phytoplankton community" presents a study on the impact of recurring heatwaves on freshwater plankton communities.   
  
Key concerns need to be addressed before to consider your manuscript for publication.   
  
The three reviewers made suggestions about several issues. For instance, the study did not measure direct functional indicators like primary productivity, relying only on proxies like chlorophyll-a and dissolved oxygen. Sampling frequency was also low. Furthermore, more details on the experimental setup and rationale are required, especially concerning the choice of heatwave intensity and experimental replication methods. The discussion needs more context.  
  
Addressing these issues would enhance the manuscript’s clarity and rigour and it would help the manuscript to be considered for publication.  
  
Please, address in details point by point each comment from the Reviewers in a separate document  
  
All the best,  
Dania

*Dear Editor,*

*thank you for the opportunity to revise our manuscript. We are also grateful for the insightful feedback from the reviewers which we have carefully considered. We have addressed all reviewers’ concerns and thereby significantly improved the manuscript. As suggested, we have removed all wording related to functioning, as we did not measure it explicitly. We also added an analysis of phytoplankton biomass and included an analysis of zooplankton data to the paper to elucidate the potential effects of grazing on the observed phytoplankton response. Overall, we consider the manuscript has improved significantly thanks to the critical insights provided by the reviewers. We sincerely hope that you will find it suitable*

*for publication at Ecology and Evolution.*

*Best wishes,*

*Francesco Polazzo on behalf of all the authors*

*All reviewers comments are copied below in normal typeface black text. Our responses are marked in blue italics.*

Reviewer(s)' Comments to Author:  
  
Reviewer: 1  
  
Comments to the Author  
In the manuscript entitled “Recurrent heatwaves slow down the recovery of a phytoplankton community”, the authors performed a mesocosm experiment in which freshwater plankton communities were exposed to three recurrent heatwaves in order to assess how these recurrent exposures could affect the functioning of phytoplankton communities, and to evaluate if the critical slowing down or the community rescue hypotheses would prevail in shaping the responses. I found the subject of this study very timely and very interesting, and I agree that more studies investigating the effects of consecutive disturbances are needed if we want to better understand the consequences of climate-change related extreme events on phytoplankton and primary productivity of aquatic ecosystems.

*Dear Reviewer,*

*Thank you for taking the time to revise our manuscript.*

*We appreciate your positive feedback and are pleased that you find the topic timely and relevant. As you rightly point out, understanding how consecutive disturbances influence the functioning of phytoplankton communities is crucial for predicting the impacts of climate change on aquatic ecosystems.*

However, I have some important concerns regarding both the study and the presentation of the results. First, while investigating the functioning of aquatic ecosystems and phytoplankton communities, the authors did not actually measure any direct functions of phytoplankton (primary productivity for example), but instead only measured chlorophyll-a and dissolved oxygen, both of which are impacted by many other factors than just primary production and therefore are not always good proxy for primary productivity. I think that actually measuring primary production (a lot of methods can be used to do so) would have helped a lot in gaining some insights in the actual response of the system.

I would have also suggested to perform more sampling events, as the sampling frequency performed in the study may have missed important events and responses in my opinion.

*Thank you very much for your valuable feedback. We appreciate your insightful comments regarding the measurement of ecosystem functioning in our study. We fully understand your concern about the limitations of using chlorophyll-a and dissolved oxygen as proxies for primary productivity, as these variables can indeed be influenced by factors beyond phytoplankton activity alone. After careful consideration of your suggestion, we have revised the manuscript, and we now do not phrase it anymore around “functioning”.*

*Unfortunately, due to the design and scope of this particular experiment, we did not perfom more sampling, which may have led us to miss important response dynamics.*

In addition, I think that not enough details are presented in the Materials and Methods for the experimental set-up, and why such a set-up was chosen. On the other hand, I found that the statistically analyses were done very carefully and well-presented.

*We have carefully revised the Materials and Methods section following the suggestion of all three reviewers. We believe that the revised version provides sufficient information on the experimental design and relative rationale.*

Finally, I think that the discussion section does not put enough context and comparison of the obtained results with other studies. I have also some specific comments that I listed below.

*Thank you for pointing the weakness of the discussion. Indeed, the original discussion was lacking comparison with existing studies. We have made a significant effort to re-structure the discussion around the results of other published studies.*

L57; “have been hardly” should be replaced by “have hardly been”.

*Corrected, thank you.*

L69-70. Some studies have investigated the effects of consequent heatwaves on the stability (resilience, resistance, recovery..) of coastal marine systems, and some found long lasting effects of the heatwaves on the stability. I know that responses of freshwater and coastal communities can be very different, but maybe these marine studies could be mentioned here.

*We searched the literature again, but could not find studies looking at the effects of reoccurring heatwaves on stability of coastal marine systems. We could only find the studies reported below dealing with multiple HWs, but unfortunately those are of little relevance for our study.*

Smith, Kathryn E., Michael T. Burrows, Alistair J. Hobday, Nathan G. King, Pippa J. Moore, Alex Sen Gupta, Mads S. Thomsen, Thomas Wernberg, and Dan A. Smale. ‘Biological Impacts of Marine Heatwaves’. *Annual Review of Marine Science* 15, no. Volume 15, 2023 (16 January 2023): 119–45. <https://doi.org/10.1146/annurev-marine-032122-121437>.

Garrabou, Joaquim, Daniel Gómez-Gras, Alba Medrano, Carlo Cerrano, Massimo Ponti, Robert Schlegel, Nathaniel Bensoussan, et al. ‘Marine Heatwaves Drive Recurrent Mass Mortalities in the Mediterranean Sea’. *Global Change Biology* 28, no. 19 (2022): 5708–25. <https://doi.org/10.1111/gcb.16301>.

*However, we did include references to more studies (particularly from freshwater ecosystems) in the main text and discussion, some of which deal with multiple heatwaves.*

L120-122. From where do the communities used in the mesocosm experiment come from originally. This indication must be given here.

*We* *now specify that the planktonic compartment of the communities come from a large artificial pond present in the mesocosms facility, whereas the macroinvertebrates come from a river nearby.*

L128-132: How was +8°C chosen as the intensity of the heatwaves? Based on meteorological data from the area? Or in order to compare with similar studies. I know that most experimental studies often simulate heatwaves by increasing the water temperature to +3° to +°6C, at least in coastal environments.

*The +8°C HW treatment was designed in accordance to the findings of* Woolway *et al.* (2021)*, whose study specifically deals with lake HWs and project future HWs to fall within this magnitude range. Although +3°C can be considered an HW in marine systems, in a shallow lake, would hardly be considered a HW. We have added the rationale behind the treatment intensity to the main text.*

L140-142. How much volume was sampled in total for each mesocosm during each sampling event?

*We realised we did not include this important information. A total of 5 litres was sampled from each mesocosm during each sampling event, and a homogenised subsample of 250 mL was taken for phytoplankton identification and counting. As for zooplankton, the 5 litres were concentrated using the zooplankton net (mesh size: 55 µm). We have added the information to the text.*

L153-154. Chlorophyll-a is not a proxy of primary productivity, but of phytoplankton biomass. Under some perturbations, primary production from phytoplankton can change (be enhanced or depressed) while this is not seen in chlorophyll-a concentration (changes in the community composition toward taxa containing less chlorophyll-a or with different pigment ratio…). In addition, chlorophyll-a, as a proxy for phytoplankton biomass, is also affected by phytoplankton loss factors, such as grazing, viral lysis, natural cell death etc. For all these reasons, with such a low temporal resolution in chlorophyll-a concentration measurements, I am not sure that it is correct to use chlorophyll-a as a proxy of primary productivity, maybe only as a proxy for phytoplankton biomass. In general, I know that chlorophyll-a and dissolved oxygen are important marker to assess the state of an aquatic system, but I find it maybe not enough to assess the functioning of a system, especially with such a low sampling frequency.

*We appreciated the in-depth overview of why chlorophyll-a is not an appropriate measurement of primary productivity. We substantially agree with the Reviewer, and we do not phrase our study as measuring primary productivity anymore. Additionally, as suggested by Reviewer 3, we included the analysis of phytoplankton biomass alongside chlorophyll – a.*

L159-160. How was this instrument calibrated? You need to add this information here. Were some samples taken for fluorometry (chlorophyll) / Winkler (oxygen) measurements, to check about potential drift in the sensors?

*Calibration of the chlorophyll-a sensor was done using the correspondence between the chlorophyll concentration and relative fluorescence units (RFUs) relative to different rhodamine concentrations, corrected according to the measured temperature. The oxygen probe had been calibrated using the Winkler method. This information has been added to the manuscript.*

L164. When you indicate that the model includes HW as fixed effect, do you mean that it includes water temperature?

*We included HW in our model as a factorial variable with two levels: present and absent. We now specify it in the revised version of the manuscript.*

L172-175. Would that not be a mean of evaluating resilience instead of recovery? From my understanding, the slope (trajectory of recovery) is traditionally used to assess resilience, while recovery is assessed on a specific day sometime after the end of the exposure to the disturbance. You can check Hillebrand et al. (2018), whose mathematical definitions for resistance, resilience, recovery etc. are often used for aquatic mesocosm experiment.  
Hillebrand, H., Langenheder, S., Lebret, K., Lindström, E., Östman, Ö. and Striebel, M. (2018), Decomposing multiple dimensions of stability in global change experiments. Ecol Lett, 21: 21-30. [https://urldefense.com/v3/\_\_https://doi.org/10.1111/ele.12867\_\_;!!D9dNQwwGXtA!Q4AMC5oG32-jL7v94rvVI5QpKSkh1\_wEDxkGIV4XYjbRwLMFZLRIYsYjbaqgwiNmIkfE-mJg8\_ya4NZMxCSVbmBhoQ7v$](https://urldefense.com/v3/__https:/doi.org/10.1111/ele.12867__;!!D9dNQwwGXtA!Q4AMC5oG32-jL7v94rvVI5QpKSkh1_wEDxkGIV4XYjbRwLMFZLRIYsYjbaqgwiNmIkfE-mJg8_ya4NZMxCSVbmBhoQ7v$)

*Indeed, the method we initially used shows the trajectory of recovery, which can be surely interpret as resilience. Following the suggestion of Reviewer 2, in the revised manuscript we used another metric to measure resilience, whereas the comparison of the slope was removed.*

L216. Fig 2a. Do you know why there was already a difference in dissolved oxygen concentration between treatments before the start of the experiment?

*We do not know for sure, but we assume it is due to natural variability. In any case, the magnitude of such difference is small and opposite to the observed effect after the HWs.*

L216. In Fig 2b, the colors for the points are not explained (I assumed they represent different replicates of mesocosms). Overall, I understood that you used four replicates for each treatment (control and HW). But then, I don’t understand how you ended up with four replicates for the difference between the control and the HW treatments? Did you randomly selected mesocosms to do the differences? Usually, for mesocosm studies, to assess the difference between treatments, the logarithmic response ratio is used, i.e., the mean of all replicates for the control treatment is calculated, then the mean of all replicates for the treated mesocosms is calculated, and then the logarithmic is calculated on the ratio of the mean of the treated mesocosms over the mean of the control treatment.

*We end up with 4 replicates from the difference between the control and HWs treatments because we compared the mean of the control mesocosms with each individual HW treatment replicate. We are familiar with the log-response ratio, and we have used it in the past. However, although LLR has been used in micro and mesocosm studies, it is more commonly used in meta-analyses. In micro and mesocosm experiment, the non-log ratio, has been used at least as frequently as LRR (see Baert et al 2016 Ecology, Pennekamp et al 2018 Nature, Zhao et al 2019 Ecol Lett), we thus do not see a very good reason to change this calculation.*

*We amended the figure caption, thanks for the suggestion.*

L246. Similarly, do you know why they were already some differences in chlorophyll-a concentration between treatments before the start of the experiment?

*We do not know. The mesocosms were kept in the exact same conditions. We have to assume it was natural variability, but again the magnitude is small compared to the observed response after the HWs.*

L246. Fig 3b. Same as for Fig2, the color code needs to be explained.

*It has been explained, thanks.*

L271. In this figure, I think it could be better for the reader if there was some way to indicate which days correspond to heatwave days, maybe with some distinct color or indication.

*We coloured in light red the facet of the days corresponding to a heatwave, and we have modified the figure description accordingly. Thank you for helping us improving this figure.*

L279-280. Again, you haven’t actually measured any function performed by phytoplankton, you have only assessed chlorophyll-a and dissolved oxygen, which can both be affected by other factors than just phytoplankton primary production.

*Right. We modified the manuscript accordingly.*

L330-349. All this part of the discussion does not actually discuss about the results of the study, and should therefore be either deleted or relocated to the introduction section. Overall, I found that not enough effort is made in the discussion to compare the obtained results with other studies, to put the results of the study in a broader context.

*Thank you for this comment. We completely removed this part from the manuscript, and substantially re-wrote the discussion to put our results in a broader context.*

Reviewer: 2  
  
Comments to the Author  
The authors in “Recurrent heatwaves slow down the recovery of a phytoplankton community” tested the role of critical slowing down or community rescue, as possible recovery mechanisms to recurrent heatwaves, using phytoplankton dynamics and functioning during a mesocosm experiment.   
  
The question is very important and the introduction is well written, the experimental setup is elaborate (especially the whole experimental setup described in a previous paper) but the analyses, the results and the discussion are not really convincing and leave many questions open.

*Dear Reviewer,*

*Thank you for taking the time to revise our study.*

*We are glad you found the introduction well written, and we are grateful for the insightful comments on the analyses, results, and discussion. We have revised the manuscript following your suggestions, and we believe it has notably improved.*  
  
L119-120 “artificial lagoon”- The Hermann et al 2024 paper refers to a freshwater pond (lagoons are usually next to the sea, with brackish water and different species composition). Which one is correct?

*We realised that “lagoon” refers to brackish water, and was thus misleading. It was indeed a large freshwater pond.*

L121-122 “…homogenize among experimental units for 2 months prior to the start of the experiment.”- What do you mean by homogenize? Also, 2 months can be a long time but maybe it is needed for the adaptation of the communities to the enclosures. However, during these two months, community dynamics could lead to divergence of the communities and the replicates could be different between them. Did you assess that?

*By homogenization we mean that the water was randomly exchanged among the test mesocosms to make the communities alike and prevent divergence among replicates and treatments. Thanks to this, we did not find statistical differences in community composition on day -4 (before the application of the HWs), suggesting that the mesocosms were relatively similar at the start of the experiment.*

L144 “1mL subsample”- do you think this is enough volume to assess the phytoplankton community? It seems low to me (and probably not representative enough). Was the community that dense?

*We followed the method described in:*

*Rice, E.W., Bridgewater, L. and American Public Health Association eds., 2012. Standard methods for the examination of water and wastewater (Vol. 10). Washington, DC: American public health association.*

*We first concentrated the 250 mL sample taken from the mesocosms, and took subsamples of 1 mL to count 100 cells of the most abundant taxon, and 400 cells in total (using the Sedgewick- Rafter chamber). This means, from each concentrated sample a total volume of approximately 5-15 mL was used. The analyzed volume was used to re-calculate the concentration of cells/L of mesocosm water. The method was tested with the Utermöhl and only implemented for counting after the main taxa had been previously identified. Further details of the method have been added to the manuscript.*

L148 “biovolume (μm3/org)”- I’m not very familiar with this unit of biovolume. Usually, it is per volume of water not per organism. What is this unit corresponding to? A relative biovolume of the total biomass?

*The unit "biovolume (μm³/org)" refers to the volume occupied by a single organism in cubic micrometers (μm³). This unit describes the physical space an individual organism takes up rather than a measurement relative to the volume of water*. *This is not a relative biovolume (which would refer to the proportion of total biovolume relative to a certain volume of water), but rather the absolute biovolume of individual organisms.*

*Biovolume (μm³/org) can be used to estimate total biomass by multiplying the biovolume of individual organisms by the number of organisms and the density of those (which was assumed to be same as the water density). It helps in understanding the size structure of populations or communities.*

L155 (and throughout the manuscript) – what is the logic behind the days of sampling? They are not in the middle of the HW period and in the middle of the recovery period. Would it make more sense to be just before the HW, at the end of the HW and at the end of the recovery period (just before the HW)? I know there is nothing that can be done to change this, but I would like to see the rational, as these are the time points used to address the manuscripts questions.

*We agree with the reviewer. The sampling schedule was not optimal. The reason we had to use this sampling schedule is that this experiment was part of a series of experiments carried out around Europe and needed to be done with the same experimental design. Yet, we considered it adequate for our purpose, as it had at least a sampling point before, during and after each HW, thus allowing us to test whether rescue or critical slowing occurs.*

L173-174: the difference between control-HW is basically control minus HW, thus negative values denote higher values of the HW treated mesocosms, or HW-control as suggested in figures 2b, 3b?

*Sorry, we meant HW – controls, thank you for spotting it. We corrected it in the manuscript, even though please note that we changed this calculation following your comment below.*

L177-178: “A positive slope was interpreted as a sign of recovery, whereas a negative slope as increasing magnitude of effects and absence of recovery.” – isn’t this a hypothesis based on your results? I mean (I might be confused though, and this is related to my previous comments for clarity), you assume that a positive slope would be a sign of recovery, because before the heatwave the difference was positive. If the difference before the heatwave was negative and it became positive during the heatwave, then the indication of recovery would be a negative slope (back to the initial condition). Am I missing something? Therefore, this part should be clearer and the assumptions for recovery better defined. “ L 180-182 “In case of community rescue, we expect the slope to become close to zero after each HW, which would suggest an improved resistance to previously experienced stress.” – would we expect slope close to zero after each HW or after the second and third, as in the first, there would be no recovery?

*The hypothesis of a positive slope to be interpreted as a sign of recovery assumes that an HW would generally have a negative effect on both phytoplankton biomass and oxygen concentration. This expectation is derived from an earlier review work we published in 2022 (see Polazzo et al 2022 GCB in the reference section). However, what the reviewer is saying it correct: were our results different (positive difference), the meaning of the slope would be different. Nevertheless, we followed your next comment, and we have changed this calculation.*  
  
Is this the best way to test recovery? Could there be a better/clearer way? Could methods used in Baert et al 2016 be appropriate? (I’m not a co-author of the paper - Baert, J.M., De Laender, F., Sabbe, K. and Janssen, C.R. (2016), Biodiversity increases functional and compositional resistance, but decreases resilience in phytoplankton communities. Ecology, 97: 3433-3440. [https://urldefense.com/v3/\_\_https://doi.org/10.1002/ecy.1601\_\_;!!D9dNQwwGXtA!Q4AMC5oG32-jL7v94rvVI5QpKSkh1\_wEDxkGIV4XYjbRwLMFZLRIYsYjbaqgwiNmIkfE-mJg8\_ya4NZMxCSVbmhPGarL$](https://urldefense.com/v3/__https:/doi.org/10.1002/ecy.1601__;!!D9dNQwwGXtA!Q4AMC5oG32-jL7v94rvVI5QpKSkh1_wEDxkGIV4XYjbRwLMFZLRIYsYjbaqgwiNmIkfE-mJg8_ya4NZMxCSVbmhPGarL$) )

*The way we assessed recovery in the original submission was probably not the best in general terms. Yet, it provided an intuitive visualisation of the recovery dynamics between HWs. The way to calculate recovery / resilience in Baert et al. 2016 works well also in our case, and we have adapted it in the revised manuscript. We also used the term resilience, instead of recovery in the revised version, as such is called in Bert et al. 2016, and because our study is actually centred on resilience. Moreover, Reviewer 1 suggested that these kinds of metrics do indeed focus more on resilience than on recovery.*

L226 “time of”-something is missing

*The “of” was incorrect. It was removed. Thank you for noticing.*

Figure 3a: Chl-a in the HW treated mesocosms was different from the control even before the heatwave. This shows a trend unrelated to HWs.

*Yes, it does. We believe this is due to natural variability in the mesocosms. However, the statistical analysis and visualisation of the data give us confidence that there were strong effects of the HW treatment nonetheless.*

Figure 4. Please arrange into chronological order and define the timepoint of sampling as before the HW, during-after the X HW.

*Done*. *Thank you.*

L329-349- I find this paragraph irrelevant to the experiment and the results provided

*We removed the paragraph and modified the discussion section.*   
  
I miss from the discussion a reference to zooplankton. I know this manuscript is focusing on phytoplankton, but zooplankton was present and it must have been affected by the heatwave, affecting, in turn, phytoplankton. This should be discussed and incorporated in the paper overall, I would suggest. For example, a very recent paper (I’m not a co-author) mentions the influence of recurrent HWs to zooplankton and in turn to phytoplankton:   
Huỳnh, T.-H., Horváth, Z., Pálffy, K., Kardos, V., Szabó, B., Dobosy, P., & Vad, C. F. (2024). Heatwave-induced functional shifts in zooplankton communities result in weaker top-down control on phytoplankton. Ecology and Evolution, 14, e70096. [https://urldefense.com/v3/\_\_https://doi.org/10.1002/ece3.70096\_\_;!!D9dNQwwGXtA!Q4AMC5oG32-jL7v94rvVI5QpKSkh1\_wEDxkGIV4XYjbRwLMFZLRIYsYjbaqgwiNmIkfE-mJg8\_ya4NZMxCSVbr0ae\_zk$](https://urldefense.com/v3/__https:/doi.org/10.1002/ece3.70096__;!!D9dNQwwGXtA!Q4AMC5oG32-jL7v94rvVI5QpKSkh1_wEDxkGIV4XYjbRwLMFZLRIYsYjbaqgwiNmIkfE-mJg8_ya4NZMxCSVbr0ae_zk$)

*We have included an analysis of the zooplankton taxa monitored in the mesocosms and discussed its potential contribution to phytoplankton change. However, there are few points that need to be mentioned in this regard. First, there are less sampling points for zooplankton than for phytoplankton, which makes causal links between changes in producers and consumers harder, even though we still described the general trend. Second, the overall effect of the HWs on total zooplankton biomass was not significant, nor it was on compositional change. Yet, we see an increase in Cladocera after the last HW. Cladocera are efficient grazers and may have contributed to the decline of phytoplankton biomass towards the end of the experimental period. We now discuss these findings in the manuscript. We have also included more references to published studies looking at the effects of HWs on freshwater planktonic communities. Thanks for the suggestion.*

Reviewer: 3  
  
Comments to the Author  
In their experimental study in mesocosms, Polazzo et al. studied the recovery patterns of phytoplankton communities from recurring heatwaves and tested whether critical slowing down or community rescue is the dominant process underlying community and ecosystem responses. Overall, I consider the study important, as developing a better understanding of community resilience to heatwaves is a very relevant in the context of global change. The research questions are relevant, and the experimental setup is adequate to address them.

*Dear reviewer,*

*Thank you for taking the time to revise our study.*

*We are glad you consider our study important and relevant.*

However, I found a number of methodological issues (e.g., details on the setup, calculation of recovery) which need to be clarified and also think that the discussion needs to be largely rewritten to better reflect on the potential mechanisms underlying the patterns. I therefore cannot recommend it for publication in its present form and listed several general and specific comments below. In my opinion, the manuscript can be reconsidered for publication after addressing these issues in detail.

*Thank you for* *pointing out the weaknesses of the manuscript. We have followed your suggestions (and the ones of the other two reviewers) to improve the manuscript. We have added extensive information in the material and methods section, changed the calculation of recovery, and re-written the discussion section. We sincerely hope that you find the revised version suitable for publication.*   
  
General comments:   
  
• The authors interpret their results in the context of recurring heatwaves. However, these heatwaves were also increasing in temperature (due to seasonality), which may have influenced recovery dynamics as warming intensity is also a key driver of community responses (see my specific comments below). This has to be at least discussed.

*The reviewer is right. In the new version of the manuscript, we have discussed how the higher absolute temperature of the different HWs may have impacted our results.*   
  
• There are some methodological issues that need some justification. For example, an explanation why the method used to analyse phytoplankton samples (i.e., not a generally recommended sedimentation method) is considered representative.

*As described above, we followed the method described in:*

*Rice, E.W., Bridgewater, L. and American Public Health Association eds., 2012. Standard methods for the examination of water and wastewater (Vol. 10). Washington, DC: American public health association.*

*The method was based on the concentration of the sampled volume and sub-sampling for cell identification and counting using the Sedgewick- Rafter chamber. The method is based on the Utermöhl method but adjusted for processing a large number of samples with varying cell density. In any case, it was previously validated based on the conventional Utermöhl approach, and only implemented after the main taxa had been properly identified. Further details on the phytoplankton identification and counting have been provided in the manuscript.*

• Importantly, some more information is needed for the calculation of recovery. While the general approach (comparing slopes relative to the control) seems valid to me, I do not understand why a time point during each heatwave were used as a starting point? The strongest effect of the disturbance can be most likely observed at the end of the heatwave. Though several possible metrics are available to quantify recovery, it is generally addressed as a dynamics following the disturbance (see e.g. Ingrisch & Bahn 2018, TREE, Martínez-De León & Thakur, 2024, TREE). Please explain the choice of your method in more detail.     
*We have changed the calculation of the recovery approach focusing more on resilience, as suggested by reviewer 2. But in doing so, we also included your recommendation, and we calculated it not using a sampling point during the HW, but always comparing before and after a HW event. We sincerely hope that you find this approach more suitable.*

• The Discussion lacks to discuss the potential mechanisms driving changes in ecosystem functions and underlying compositional shifts. Especially the first part is merely the reiteration of the results without a discussion in the context of other studies.   
For example, the authors should discuss the potential effects of the fact that the recurring heatwaves were characterized by increasing temperatures (see my first general comment).

*We have re-written the discussion section, adding more references and highlighting the potential mechanisms that affect the phytoplankton response (including phytoplankton-zooplankton interactions, see below).*

Also, they seem to entirely neglect the potential effect of grazers. There is a growing number of studies that shows that warming/heatwave effects on phytoplankton are often resulting from food web interactions. See e.g. Kratina et al. (2012, Ecology), Velthuis et al. (2017, Ecosphere), Garzke et al. (2019, Plos Biol), Vad et al. (2023, Global Change Biol), Huynh et al. 2024, Ecol Evol), and more.

Given zooplankton data seem to be available in a parallel study by Hermann et al. 2023 (cited by the authors), I strongly recommend to at least discuss their potential effects underlying phytoplankton dynamics based on existing studies (I suggested some above, but I recommend checking others as well). The most straightforward would be to include zooplankton biomass (or biomasses of major taxa) in the analysis.   
*Following your suggestion, we have included in the analysis the zooplankton data, and analysed the potential effects of grazers on phytoplankton responses.*

Specific comments:   
  
L60: Check also the study by Vad et al. (2023, Global Change Biol) who investigated resistance and recovery also using a planktonic model system so the conclusions of the study can be relevant for your work.

*Thank you for suggesting this study. The results are only partially relevant to our study.*   
  
L68-69: Yes, applying a single heatwave is certainly one of the potential reasons as the effect of multiple perturbations may accumulate over time. But besides frequency, the duration and intensity of heatwaves are also very important and context-dependent (i.e., critical threshold temperatures vary across species, communities and ecosystems). I suggest to briefly mention this here and also discuss it in the Discussion in the context of your study, especially as you had heatwaves with increasing temperatures (i.e., intensity).

*We agree this is an important point. We have included new text in the discussion section where we discuss how different intensity in the HWs could have impacted our results.*

I agree that most studies focusing on these aspects are based on terrestrial or marine ecosystems, but there are freshwater examples as well and I suggest referring to some. See e.g. Seifert et al. (2015, Ecology and Evolution) on the effect of different heatwave intensities on laboratory aquatic communities.

*Thank you for suggesting this very interesting study. We have cited it in the discussion of the revised manuscript, as well as other papers referring to the same topic.*

L99-101 and 113-114: These sentences are redundant; I suggest merging them.

*Done. Thank you.*   
  
L119: Can you please also add the dimensions (diameter, height) of the mesocosms? What was the water level set, i.e. above the sediment layer? (I can see that these details are added in another paper by the authors, but I suggest to briefly mention it here as these are basic information which help to better understand the experimental setup)

*We have added this information.*   
  
L121-122: Are you sure that during these 2 months the communities became more homogenous? I would rather expect that the mesocosm communities will be similar right after filling (due to the same inoculum) and gradually diverge resulting in increased beta-diversity among them. Did you test this (e.g. by comparing group dispersions at the beginning vs the end by betadisper() in R). If not, I suggest being more cautious with your formulation.

*We randomly exchanged water among the test mesocosms to make the communities alike and prevent divergence among replicates and treatments. We did not find statistical differences in community composition on day -4 (before the application of the HWs), suggesting that the mesocosms were relatively similar at the start of the experiment and that the homogenization tasks (water exchange) worked well. However, we did not include a larger number of sampling points within the acclimatation period (only day -4). In any case, we have revised the sentence to be more accurate.*

But to test the heterogeneity among mesocosm communities would be a good addition to your manuscript, as it will inform about how different the communities were before the start of the experiment.

*As mentioned above, we do not have data for the initial phase. However, we did test for compositional differences on day -4, and did not find significant differences between communities, indicating that the community structure were very similar at the start of the experimental period.*  
  
L123: If I understood well, the present study is based on a subset of a larger mesocosm experiment detailed in Hermann et al. (2024). If true, please mention this more explicitly and explain which subset of the larger experiment were used here. Also, I checked the description in Hermann et al. and found some inconsistencies between the two studies, e.g., the 2 vs 3-month acclimatization period. Please doublecheck this.   
*We now more explicitly state that this is a subset of a larger experiment, and we have amended some minor inconsistencies. Thank you.*

L143-145: When estimating phytoplankton cell densities, sedimentation methods like the Utermöhl-method are generally preferred particularly for the smaller species or in samples with low densities so that densities are not underestimated. Can you provide a justification, why you used this method and why you consider it representative for your case? Also, how many cells (or colonies) did you count per sample (usually 400 is recommended)? Adding these details are important to be able to judge the representativity of your methodology for quantifying phytoplankton composition.

*As described above we counted 400 cells (or colonies) per sample (using the Sedgewick- Rafter chamber). This was done by taking subsamples (1 mL) of the concentrated sample, and measuring about 5-15 mL per sample (depending on the algae density in each sample). We have provided further details on the method and the validity of the same.*

L150-151: Please provide a reference for this.

*We have added a reference.*   
  
L154: Yes, chl a can be used (and is widely used) as a proxy for biomass, but it depends on taxonomic composition and the physiological state of cells. Why not using community fresh weight as a proxy for biomass as it was anyway estimated and it would probably represent a more reliable proxy for community biomass? Or both chl a and fresh weight?

*Following your suggestion, we now use both chlorophyll a (as a proxy for photosynthetic activity) and phytoplankton biomass estimated as fresh weight. Thanks for the suggestion.*  
  
L155: You mention here (and later in the results) that DO concentration was recorded and analysed. As DO concentration is temperature-dependent, i.e., the solubility of oxygen decreases with increasing temperature, even without underlying community changes a decrease is expected. Did you account for this? DO saturation is a measure that is corrected for temperature.   
*Thank you for this comment. We initially did not account of the effect of temperature on the solubility of oxygen. Now, we included in the calculation only days where water temperature was the same in control and HW mesocosms (so in between HW events), to eliminate the effect of temperature on oxygen solubility and to compare only similar situations.*

L156: ‘during the experimental period’ – please specify at which frequency

*Added, thank you.*   
  
L159-156: So is chl a determination based on field fluorescence measurements? (i.e. no dark adaptation was involved?) Also, can you explain how fluorescence was converted to concentration?

*Yes, it was done based on fluorescence measurements and a regression model that correlates relative fluorescence units (RFU) and chlorophyll-a concentrations. Such correlation model is the one used for its calibration. Further details have been provided in the text.*L167: ’to ensure’ – it is not really ensuring, but to visually inspect, please rephrase

*Rephrased.*   
  
L174-177: I don’t understand why recovery was calculated as a slope from the middle of the heatwaves and not from the end of the heatwave (where the strongest effect can be expected) Please at least provide some justification for your method and support it with references, if possible. See also my related general comment.

*We have changed this calculation following your suggestions and those of reviewer 2 regarding this important point. Thanks.*  
  
L256-257: I suggest considering showing the temporal and treatment-specific differences by a distance-based multivariate analysis (e.g. NMDS or dbRDA, with different colours indicating treatments). It could be either one analysis/plot, but if temporal turnover masks (i.e., is overall stronger than) the treatment effects, the key dates can be shown by separate plots. Taxa/species labels may also be indicated (or only the most important ones based on your SIMPER analysis). In my opinion, this would be a clearer overview of compositional changes than the present barplots (which can be shifted to the supplement).

*We have included a NMDS plot showing compositional change in phytoplankton over time, which provides a very clear and useful visualisation of the compositional dynamics. However, we also kept the bar plots, as they complement the information about relative changes in taxa biomass.*   
  
L258-261: I suggest moving the table with the PERMANOVA analysis from the appendix to the main manuscript. I believe this would help to more readily see and understand the results on community composition. Also, supplement the table with more complete summary statistics (incl. dfs, R2 values) to better understand the models and the treatment effect sizes.

*Moved to the main text and implemented.*   
  
L275-310: These parts are to a large extent only the re-iteration of the results and do not provide any information on the potential mechanisms. I suggest including a general first paragraph which briefly summarizes all the key results, while discuss the specific results (on DO, chl, and composition) in separate paragraphs in the context of findings by other studies. Please also see my general comment on the discussion.

*The discussion has been rewritten in line with the proposed changes.*   
  
L329-330: I feel that this statement is somewhat contradicting with the discussion at L316-320. Also, to me it is not clear on what the author’s conclusion of no evidence for critical transition is based on. There is no analysis performed to test this. What would be the definition/threshold of a critical transition in the context of your study? This has to be made clearer.

*This sentence has been removed. Please see new text in the discussion.*  
  
L523-524: group ‘B’ – should be group ‘A’

*Thank for spotting it. It has been corrected.*