Response to Reviewers' comments

We would like to thank both reviewers for their thorough evaluation of our manuscript and for their helpful, constructive, and encouraging feedback.

Reviewer 1:

This manuscript presents the first regional compound-event assessment for Hawaiian waters, combining a submesoscale-permitting physical-biogeochemical model with three SSP scenarios. The explicit separation of tidal and mesoscale contributions to natural variability thresholds, and the finding that spatial CE patterns are driven primarily by natural variability rather than anthropogenic trends, represent substantial new insights beyond existing global-scale CE studies. The conclusions are well-matched to the results and end with a practical call to action. Minor revisions are recommended: quantifying the permanence of acidity threshold exceedance spatially, sharpening the epistemic framing of ecosystem impact claims, and adding specificity to the observation network recommendation. Subject to these revisions, the paper is suitable for publication.

Specific comments:

1. The conclusions acknowledge that the acidity threshold exceedance becomes permanent around 2040 in all three SSP scenarios, this is a striking result. However, neither the conclusions nor the main text quantify what fraction of the domain reaches this state by decade, nor how "combined intensity" is interpreted once one variable's threshold is permanently exceeded. A figure or supplemental table showing the spatial progression of permanence is recommended.

We agree with the reviewer. In fact, we had planned to present a figure and more details on the permanence but decided that more figures and material could potentially overload the manuscript. The revised manuscript will contain a figure as shown below (Figure 1) and discuss the permanence and the effect of the seasonal cycle in preventing a permanent exceedance in more detail. The threshold is not altered by the fact that it is permanently exceeded. The definition of individual or combined intensity does not change.

The revised manuscript will state these facts clearly.

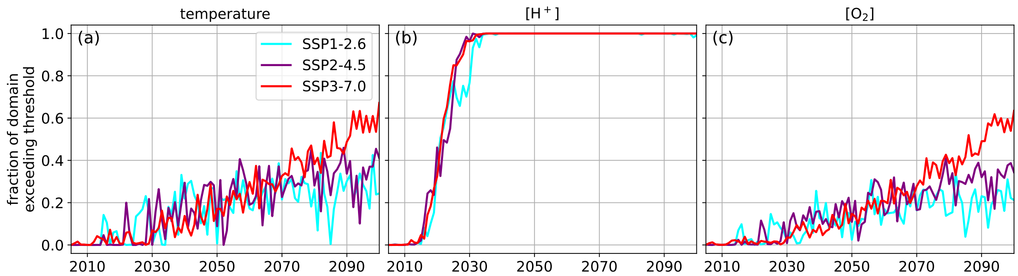

Figure 1: Fraction of domain exceeding threshold for the 0–50 m depth bin. Data shown are based on daily-mean fractions. Annual-mean shown for clarity. A mean of all three simulations is shown for SSP3-7.0. Threshold for hydrogen ion concentration ([H$^+$)] is permanently exceeded in the entire domain around the year 2040 in all three scenarios.

2. The conclusion states it is "reasonable to assume" substantial ecosystem changes under unabated emissions, which is a fair inference, but this framing is stronger than the evidence in Section 3.6 strictly supports, which relies primarily on single-stressor studies. The conclusion should more explicitly distinguish between (a) documented single-stressor effects, (b) limited multi-stressor evidence, and (c) the inference to compound events, making the epistemic steps transparent.

We will revise the statement and strictly refer to the single- and multi-stressor evidence provided by the studies.

3. The authors acknowledge that depth-averaging within bins is debatable given the strong vertical gradients present in the domain, but offer no justification for the specific bin boundaries chosen nor any test of sensitivity to this choice. This is a non-trivial concern, particularly for the 50–200 m bin, which spans ecologically distinct zones including the base of the euphotic zone and the upper oxygen minimum layer. Averaging [H$^+$] and [O2] across this range may suppress within-bin variability and systematically bias threshold amplitudes in ways that differ between the leeward and windward sub-domains, where eddy-driven vertical structure is known to differ. The authors should either provide a physical justification for why depth-averaging is appropriate

within each bin, for instance, by showing that the variables of interest are relatively homogeneous within bins at representative locations, or demonstrate that the key spatial contrasts reported are not sensitive to bin choice. A vertical profile comparison of threshold amplitudes at one representative leeward and one windward location, computed at full vertical resolution versus depth-averaged, would be a concise and informative way to address this concern.

The authors agree that the original 50–200 m depth bin is too large. We have carried out additional calculations of compound event characteristics in order to improve the vertical resolution of our projections. In the revised manuscript, we will add an additional depth bin to separate the original 50–200 m bin into two bins spanning 50–100 m and 100–200 m respectively.

While designing the study we gave balancing a meaningful, computationally feasible and comprehensive resolution in the vertical a lot of thought. From a physical perspective, using isopycnals instead of depth-bins would potentially be more appropriate since they represent similar water masses. However, isopycnals appear and disappear seasonally and also change with ocean warming. Strong temporal and spatial (mesoscale) variability within our domain aggravates the choice of appropriate "natural" depth bins since the frequent uplift and suppression of isopycnals and associated physical/BGC properties can easily reach a hundred meters and more. This is further complicated by the meridional slope of isopycnals that places similar water masses (and properties) at different depths at the northern and southern end of our domain. It must also be noted that near-surface water masses differ substantially between the windward and the leeward side of the Main Hawaiian Islands due to their different origin.

One of our main motivations has been to make the results of the paper accessible and usable for environmental scientists and conservationists that study the effects of climate change on ocean habitats. Habitats of marine organisms are typically described by (broad) depth ranges (e.g. *SeaLifeBase*, *FishBase*) **which makes us believe that presenting our results in terms of depth bins will be the most helpful and convenient for most readers.**

4. The final paragraph of the conclusions calls for an "extended observation network", which is a valuable recommendation, but does not specify what variables, depths, or temporal resolution would be needed. Even one sentence of specificity (e.g., sub-daily biogeochemical sensors in the 50–200 m range near designated MPAs) would make this actionable.

We agree with the reviewer. Sub-daily temperature/BGC sensors in coastal habitats is exactly what our study calls for to gain a better understanding of short-term natural variability.

Minor comments:

Line 74, Section 2.1: 'and' and the semicolon are redundant. Replace with either a comma before 'and', or a semicolon alone without 'and'.

This will be corrected.

Line 158, Section 2.3 : "mescoscale" should be mesoscale

The typo will be corrected.

Line 254, Section 3.3: "combined intensities in the 0–50 m depth bin virtually follow a linear relationship", here I suggest changing "virtually" to "near-linear" or "approximately linear" as is more precise.

We will use the term "near-linear" as suggested by the reviewer.

Line 394, Section 3.6: "oligotrphic" should be oligotrophic

The typo will be corrected.

Line 395, Section 3.6: "Prochlorochoccus" should be "Prochlorococcus" and should also be in italics

This will be corrected.

Line 395-398, Section 3.6: Inconsistent italics/capitalization for species names. Make sure Prochlorococcus, Synechococcus, and Trichodesmium are italicizes throughout manuscript, since standard scientific convention requires italic formatting for genus and species names.

This oversight will be corrected.

Line 401, Section 3.6: "import" should be important.

This will be corrected.

Throughout the manuscript, use Hawai'i consistently (with 'okina) , there are some instances where Hawaii without the 'okina is used (e.g., Line 202, 459).

This will be corrected.

Line 463, Conclusions: "timely" here is used to mean "important now" but reads ambiguously; "urgent" or "now essential" is clearer in scientific writing. Citation: https://doi.org/10.5194/egusphere-2026-1573-RC1

The wording will be corrected accordingly.

Reviewer 2:

The manuscript by Friedrich et al uses a downscaled regional model for the Hawaii region to study the frequency and intensity of compound extreme events over the course of the century. They use multiple runs with different SSP-RCP scenarios and evaluate the natural and emerging variability of hydrogen ion, dissolved oxygen and temperature variability at surface and subsurface, using different calculations to account for the effect of mesoscale and tidal influence in different regions around Hawaii. They find that compound event intensity is mostly dominated by acidity, that deeper waters experience more legacy effects from increased CO_2, and that accounting for tidal variability delays the onset of extreme events in nearshore regions.

The study is well designed and well thought out, considering the limitations imposed by computational costs of conducting multiple long-term simulations at 4-km horizontal resolution and evaluating outputs at hourly/daily temporal resolution, and the methods are well explained. However, there are a few points that could be better discussed or explained. First, the manuscript would benefit from some review of water masses and mixing, and of oxygen and H+ distribution. While the text offers some information on which sources have more/less concentration of different variables, it would help to have more information on those values and a better sense of the spatial distribution of the water masses. Some information on mixed layer depth distribution (including seasonal) and rate of change would also be helpful even if it is highly variable and influenced by eddies. Was MLD distribution a factor in deciding the depth bins? If typical MLD hovers around 50m, for example, wouldn't potential long-term trends need to be discussed in the results for the 0-50 and 50-200m bins? In general, more thorough justification for the choice of depths used would be welcome. While I understand that generalizations need to be done when dealing with this amount of data, it would be good to have a better understanding of how the oceanographic features in the region guided the choices.

We will add a more comprehensive description of water masses and their effect on the spatial distribution of the variables of interest to the revised manuscript. We will also discuss the spatial and temporal variability of MLD. As stated in our response to Reviewer 1, the choice of how to present our results in the vertical has been a compromise between a number of factors. Unfortunately, the strong spatial and temporal variability within our domain, the difference in leeward and windward water masses, as well as the meridional slope of isopycnals does not allow for choosing depth bins based on oceanographic features.

The interpretation of figures at points seems a bit broad, and the figures themselves need some improvement to help the reader have a more immediate understanding of what it is they are looking at. Better placement of subplot letters, more informative titles, and labeling on axes should be reviewed. More focused suggestions and comments on these items are included below. Overall, it is a valuable contribution that expands the interpretation of compound events at scales that are of high importance for coastal environments, and it should be ready for publication with some reviews.

We will correct and improve the figures according to all suggestions made by the reviewer.

L41-43: to build this argument in the introduction it would be good to add information from observations as well. Are there references for the impact on biogeochemistry?

To our knowledge observational transects carried out through eddies such as the ones done for cyclone *Opal* and *Noah* respectively have focused mainly on productivity-related variables. Barone et al. (2019) and Barone et al. (2022) present anomalies in dissolved inorganic carbon and oxygen related to mesoscale structures / sea-level anomalies for the Hawai'i region. We believe these papers would be appropriate observational references.

L71-75: While the model validation is published elsewhere, it would still be helpful to add some information on potential biases and differences between the model and observations

We will add more details on the model performance in the revised manuscript.

L90: You might want to explain why you are using [H+] instead of pH

Following previous papers (Gruber et al., 2021; Burger et al., 2022; Burger and Frölicher, 2023; Hauri et al., 2024) we use $[H^+]$ instead of pH. The main reason for using $[H^+]$ instead of pH is the fact that pH is on a logarithmic scale which can make the application of correct statistics rather complicated. This will be mentioned in the revised manuscript.

L96-99: Perhaps this is a good place to introduce and discuss the period used for the calculations?

In the revised manuscript we will mention that our baseline period is for defining thresholds is the period 2005–2020.

L107-108: Presentable and feasible is good, but also good if it's based on oceanographic properties of the region. A discussion of general circulation, MLD distribution and variability would be helpful for the reader to understand the choices made.

A more comprehensive discussion of water mass properties, mixed layer depth variability and circulation dynamics will be added to the revised manuscript.

L120-121: So for one year of simulation?

This is correct. One year of hourly model output is used to determine threshold values based on tidal and other short-term variability that is not represented in daily averages.

L127: "individual mean intensities" – do you mean per cell?

"Individual mean intensities" refers to the three variables T, $[H^+]$, and $[O_2]$ respectively. We will correct the misleading wording.

L141-142: confusing sentence, and is there a reference for that?

Combining intensity and duration into a measure of severity is an established method that is used for example as Degree Heating Weeks. We will clarify this and add references in the revised manuscript.

L 157: spelling of mesoscale

The typo will be corrected.

L195-200: Need references here, and which water masses are those?

The main water masses involved are North Pacific Subtropical Mode Water and North Pacific Intermediate Water. This will be mentioned in the revised manuscript. We will believe that Bostock et al. (2010) will be an excellent additional reference.

L215: would be good to have some general characterization/description of seasonal cycle

A more detailed explanation of the effects of the seasonal cycle will be added to the revised manuscript.

L217: Need to be a bit more specific on how the figure shows that

This is indeed misleading as the "lateral advection" cannot be seen in the figure. We will rephrase the sentence.

Figure 4: I understand changing the colorbar to see the signal, but keeping the colorbar consistent would make it easier to compare depths

We will revise the figure and try to keep a consistent colorbar. It should be noted, however, that the difference between the depths bins amounts to a factor of three for temperature and an order of magnitude for [H^+] respectively. So, one colorbar for all depth bins might not be feasible for all variables.

L247: might be good to at least point to the section where intensity/severity are defined

A reference to Section 2.2 will be added.

L269: how could potential changes in MLD affect the distinction between 0-50 and 50-200?

We have already carried out additional calculations to improve the vertical resolution of our compound-events projections. In the revised manuscript the 50–200 m depth bin will be replaced by a 50–100 m and a 100–200 m depth bins for better resolution. The MLD varies seasonally and spatially between $\sim$10 m and $\geq$100 m across our domain. However, the projected warming has overall very little effect on the MLD evolution over the course of the century. We will add a more comprehensive discussion of MLD changes on our results in the revised manuscript.

L287: spelling of patterns

Will be corrected.

Figure 10: letters on panels hard to find, y-axis on e/f need labeling

We will check and improve all figures with respect to their readability and comprehensibility.

L351: At the surface, are you referring to the propagation between 2086-2089? It is easier to see it at depth indeed

This is correct. At the surface, the seasonal cycle obscures the propagation to some degree. The latter becomes clearer at deeper layers where the seasonal cycle is less relevant.

L354-355: I'm not sure how to interpret that from the figure given the lags and the faster response of deeper layers

We will clarify this statement in the revised manuscript. Our main point is that – once the anthropogenic trend has shifted environmental conditions strongly enough – compound events can be created by (westward) propagating mesoscale features such as anticyclonic eddies. In this feature, it is mainly isopycnal deepening and associated warming that (on top of the anthropogenic trend) leads to an extreme event. Therefor, there is no longer response time of deeper layers.

L357-359: The information from the velocity checks and interpretation seems reasonable, but it's hard to see information on eddy activity from meridionally averaged plots

We believe a Hovmoeller plot is a good and established way to illustrate the spatial propagation of anomalies. However, we are open to other suggestions.

L360: perhaps change lower depths to deeper waters?

This will be corrected.

References

Barone, B., Coenen, A. R., Beckett, S. J., McGillicuddy Jr, D. J., Weitz, J. S., and Karl, D. M.: The ecological and biogeochemical state of the North Pacifi c Subtropical Gyre is linked to sea surface height, J. Mar. Res., 77, 215–245, 2019.

Barone, B., Church, M. J., Dugenne, M., Hawco, N. J., Jahn, O., White, A. E., John, S. G., Follows, M. J., DeLong, E. F., and Karl, D. M.: Biogeochemical dynamics in adjacent mesoscale eddies of opposite polarity, Global Biogeochemical Cycles, 36, e2021GB007 115, 2022.

Bostock, H. C., Opdyke, B. N., and Williams, M. J.: Characterising the intermediate depth waters of the Pacific Ocean using δ13C and other geochemical tracers, Deep Sea Research Part I: Oceanographic Research Papers, 57, 847–859, 2010.

Burger, F. A. and Frölicher, T. L.: Drivers of surface ocean acidity extremes in an Earth system model, Global Biogeochemical Cycles, 37, e2023GB007 785, 2023.

Burger, F. A., Terhaar, J., and Frölicher, T. L.: Compound marine heatwaves and ocean acidity extremes, Nature communications, 13, 4722, 2022.

Gruber, N., Boyd, P. W., Frölicher, T. L., and Vogt, M.: Biogeochemical extremes and compound events in the ocean, Nature, 600, 395–407, 2021.

Hauri, C., Pages, R., Hedstrom, K., Doney, S. C., Dupont, S., Ferriss, B., and Stuecker, M. F.: More than marine heatwaves: A new regime of heat, acidity, and low oxygen compound extreme events in the Gulf of Alaska, AGU Advances, 5, e2023AV001 039, 2024.