Reviewer 1

Review comments for "Characterising Marine Heatwaves in the Svalbard Archipelago and Surrounding Seas" by Williams-Kerslake et al.

The authors presented a comprehensive MHW study in the Svalbard Archipelago region using TOPAZ analysis, which is validated by various moorings and OISST analysis. They presented MHWs changes in timescales of decade and season, vertical and horizontal extent, provided heat budget analysis for each MHW events, and concluded that the most MHWs are driven by the ocean heat transport. The manuscript is well written and can be published in EGUsphere after revision. My major concern is what drive the deep MHW, if the surface heat flux, how?

We thank the reviewer for their positive comments and constructive feedback. We provide detailed responses to each comment below and will revise the manuscript accordingly. With regards to the reviewer's major concern, we find that both surface heat flux and ocean heat transport contribute to the development of deep events. We will review the text to make this finding clearer.

1. L54: Use consistent time unit in L54 °C $year^{-1}$, L57 $\%y^{-1}$, L62 °C per decade

We will change the units to be consistent.

2. L110, 90th percentile. I am not sure whether the region is ice free during the summer from 1991-2022. If not, how the MHWs are defined in the ice-covered region, since water temperature changes a lot when ice is melted. E.g. the threshold, which is calculated using the temperature with ice in the early period, may be difficult to applied to the time when ice in melted in the later period. Can you test how much MHW features are changed if the threshold is set to 95th percentile?

Svalbard West has been ice-free during the summer from 1991 to 2022, so testing with different thresholds is not relevant in this context.

3. L128, equation (2), why is Tref is needed?

Tref is 0°C, therefore this could be removed from the equation.

4. Figures A1B, A2B and other figures with p-value: 1.43e-24, 1.27e-17, check and revise.

We believe the low p-values are due to a high number of observations. We will check each p-value.

5. L211-219, Figs. A5 (low correlation) and A6 (bifurcated correlation), these figures may indicate the biases of TOPAZ model in the coastal regions.

Yes, we agree. TOPAZ does not perform well on the coast at the location of Mooring M1-M2 and M4, as it is unable to resolve the cooling processes related to ice formation (L216-217). Storfjorden (M1-M2) and M4 experience intense water mass transformation due to sea ice

freezing, and the Storfjorden mooring is situated in a productive polynya. TOPAZ does, however, perform well on the west coast of Svalbard, as shown by the comparison with the Isfjorden Mouth Mooring (L203-209). TOPAZ effectively resolves the bathymetry of the Isfjorden trough, which is why we see a high correlation with the Isfjorden Mouth Mooring. We will add some clarification to the text to explain this difference.

6. L234, note that the intensity decreases in many regions although frequency and duration increase in most of the regions. Therefore, it might be helpful to use the cumulative intensity by integrating SSTA and time in units of degree-day (e.g. Huang et al. 2025, DOI 10.1175/BAMS-D-24-0337.1), which will enable us to see how MHWs intensify with time.

Thank you for the suggestion. We will add cumulative intensity to Figure R1 (below) to show how MHWs intensify with time.

7. L240-246, it might be helpful to add implications or causes of those features, e.g. warming is strong in winter than summer etc.

We are unsure what is meant here. We find that the intensity is higher in summer than in winter, but the duration is shorter, so the accumulated intensity may be larger in winter than in summer. This point is interesting; we have, however, not investigated this. Our aim was not to analyse seasonal variations in sea surface temperature (SST), but rather to compare summer MHW metrics with those of other seasons, given our focus on summer MHWs. Therefore, we believe it is best to not include this in the text, and this could be investigated in a later study.

8. L248-252, The definition of MHW differences is not straightforward: there are many regions without MHWs in left panels marked as “missing”, which results the difference in right panel are marked as “missing” or blank. Can the “missing” in the left panels be marked as “zero”? This should make the difference more reasonable. One alternative way is to assess the differences is to integrate MHWs in space and then compare their time evolution.

Thank you, we agree that the definition of MHW differences is not straightforward. Instead, as suggested, we will integrate MHWs in space for each season and then compare their time evolution. We will also add cumulative intensity to this plot. Please find the plot below:

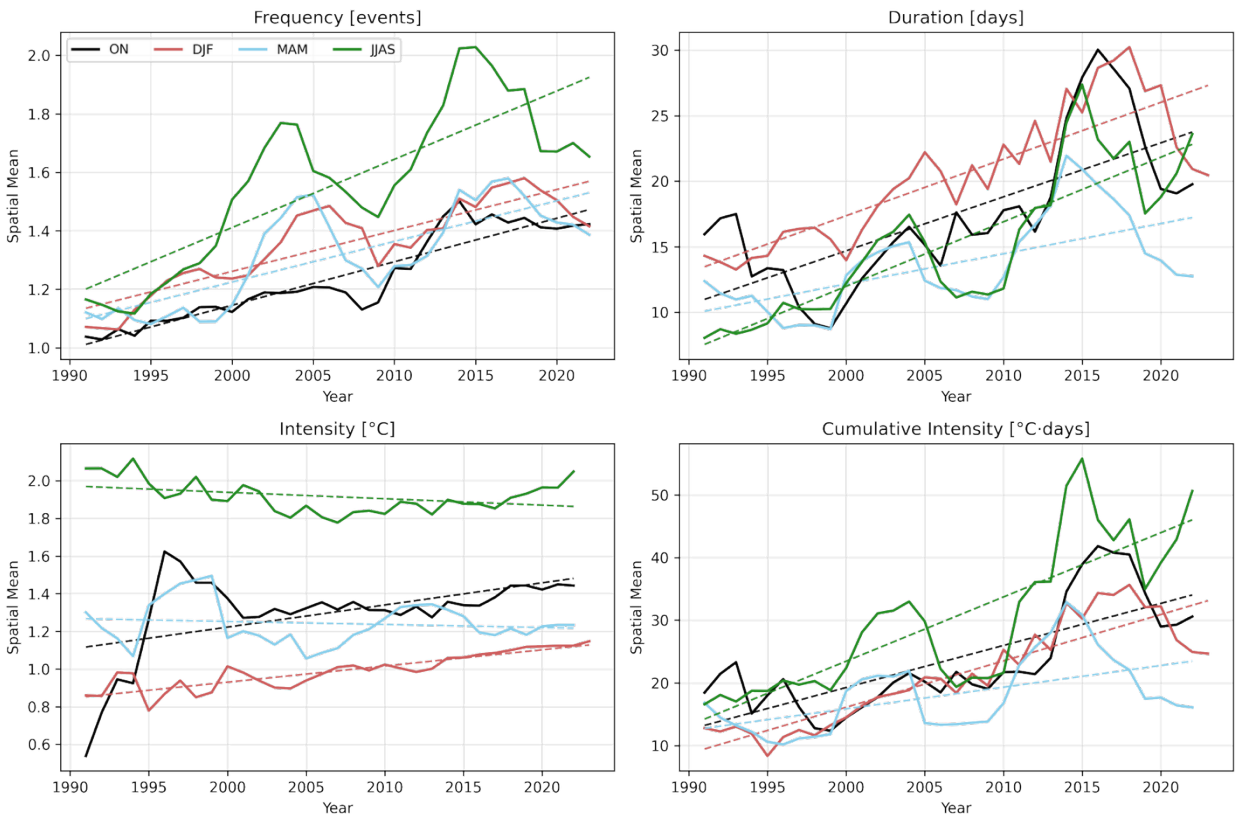

Figure R1: Spatially averaged MHW frequency (events), duration (days), intensity (°C) and cumulative intensity (°C days) for the Svalbard Archipelago and surrounding seas (69-82°N, -10-35°E) for Autumn (ON), Winter (DJF), Spring (MAM) and Summer (JJAS). Data is smoothed using a 5-year running mean. MHWs are not analysed north of the sea ice edge (sea ice concentration ≥ 15%). Dashed lines represent the linear fittings.

9. L279, "Note that MHWs are not analysed north of the sea ice edge (sea ice concentration ≥ 15%)." This might be noted much earlier in definition in section 2.1.

Thank you for your comment. We will add *"To generate maps of MHW metrics in Svalbard West and surrounding seas, MHWs were also detected individually for each grid point in the seas surrounding Svalbard. For this analysis, MHWs were not analysed north of the sea ice edge (sea ice concentration ≥ 15%)"* in Section 2.1.

10. Figures 5, 6, "peak date of each MHW". How is this defined? MHW evolution may not be synchronized in different regions, and therefore it is not straightforward to define "one" MHW within a large region (more than one grid point). What the black dots represent?

The peak date of each MHW is the date of peak intensity (maximum SSTA) for the MHW events detected using SST averaged over Svalbard West. For the peak date for each Svalbard West event, we then see in the larger mapped area (Figures 5, 6) if the grid cells exceed the 90th percentile. We understand the limitations of basing the peak date for the larger area on the peak date taken from the spatial average of Svalbard West, as the peak date is likely to differ for each grid cell. However, we chose this method since the MHW events we identify are based on Svalbard West. We will clarify this in the text.

11. L315, 332, “With the exception of events in 2016 and 2017 (deep events),” Does this imply that the deep MHWs are driven by the surface heat flux, which is hard to imagine. If not, what drive the deep MHWs? “With the exception of events in 2016 and 2017 (deep events),” why?

In 2017, the surface heat flux (SHF) anomaly is only slightly larger than the ocean heat transport (OHT) anomaly, suggesting that both SHF and OHT contributed to the event, with SHF playing the stronger role. In 2016, despite a net negative OHT anomaly driven by anomalous heat export at the northern boundary, an 8-TW positive anomaly still entered the region through the southern boundary. Thus, even though substantial heat was lost overall, the anomalous heat import—together with SHF—could have contributed to the development of the MHW event. Thus, the deep events are not solely driven by the surface heat flux. We will make this finding clearer in the text.

12. Figure 10, suggest exchange the dotted with solid lines, which will highlight the MHWs.

As mentioned by Reviewer 2, this section on the impact of heat advection on MHW events requires further analysis for us to be able to make concrete conclusions. As a result, we have decided to remove this section.

13. Section 4, Discussion, the discussion is lengthy and should be shortened.

Thank you. We agree that the discussion needs to be shortened. The section on heat advection will be removed since this will no longer appear in the results. We will also remove the section on sea ice decline since our results do not focus on sea ice. Additionally, we will shorten the section on the ecosystem impact of MHW events. We have also noted that there is a lot of repetition of the results and methods in the discussion. We will move any methodological justifications to the methods section. We will also remove any repetition of how analyses were performed, specific numerical values, figure references that were already presented in the results and detailed descriptions of individual events, unless they are directly needed for interpretation. This will ensure that the discussion only focuses on the main findings and how they fit into the broader context.