Marine heatwaves are attracting increasing interest from the research community, but their prediction lacks sufficient attention. This manuscript helps to fill this gap by analysing and extreme marine summer in the Mediterranean. The authors show the skill of a short-term forecasting system at basin-wide and regional scales. Results show the adequacy of short-term forecasting in the Mediterranean with fairly good accuracy.

The topic of the paper is interesting and the approach to it is correct, but it suffers from a certain lack of consistency in justifying the results and conclusions. I am sure that the authors have sufficient data and results to make their conclusions more robust, as the wording is sometimes vague. I think it is more a problem of how to explain the work done than of the work itself, which I think is very relevant.

We thank the reviewer for their careful reading and instructive comments which have helped improve the quality of the manuscript. We have made attempts to improve wording and messaging. Please see below for responses to the specific comments.

Through the text you are using different activity definitions. Be consistent throughout the manuscript and use a single definition if possible. If not, please make clear which one you use any time you refer to it. Suggestion, change the name of your definition to "cumulative activity", "basin activity" or similar. This is one of my main concerns as it can be confusing for the reader.

We use only one definition of activity, as was previously defined in Line 104: "Here, in order to study basin-wide spread at daily resolution, we define activity as the sum of the intensity over the area undergoing a MHW in the Mediterranean Basin". We have corrected the caption of Figure 1, which we believe was the source of confusion. We have changed "activity" to "occurrence" where necessary.

"Discussion and summary" look more like a review paper than the discussion of your results. Consider moving some references to the introduction section or extend the section by adding more discussion of your own results.

The discussion section has now been updated. The discussion of forecasting other extreme events has been moved to the introduction.

Therefore, my recommendation is to accept this work with minor revisions aimed at better clarifying the work developed by the authors. Please, see the comments below.

Other comments

Line 24: *"Zi et al., 2020"* do you mean "Li et al., 2020"?

Yes, thank you for finding this.

Please check "Benthuysen et al" as there is a typo across the text.

Thank you for finding this, the name has been corrected across the text.

Lines 79-80: *"we define activity as the sum of the intensity over the area undergoing a MHW in the Mediterranean Basin."* But in figure 1 activity is *"the daily product of area and intensity"*. Which is the actual activity metric used in figure 1? Have you used your own definition in the rest of the paper?

We apologise for the confusion. We use only one definition of activity, as was previously defined in Line 104: "Here, in order to study basin-wide spread at daily resolution, we define activity as the sum of the intensity over the area undergoing a MHW in the Mediterranean Basin". We have corrected the caption of Figure 1. We have changed "activity" to "occurrence" where necessary.

Line 94: *"and conditions remained above a third of the basin area until the decay at the end of September"* Above what? Do you mean about SST threshold? Do you mean MHW conditions were present on more than a third of the basin?

We have corrected this sentence to make it clear we are referring to MHW area being above a third of the total basin area.

Line 96, 103, 105, 155, 156, 166: Please check typo when writing *"°C"*.

Thank you for spotting this formatting issue; we have corrected it across the text.

Lines 104-105. The persistence of a minimum activity (greater than 0) is not enough to state that 2022 holds a new record for MHW activity. Please, add more justification for this statement (higher mean, higher max, greater area...) that explains the record.

Unfortunately we do not understand the confusion of the reviewer here. A justification was already included in Lines 128-134.

Line 150: Why do you choose the Ligurian Sea and Gulf of Taranto? Why not the Alboran Sea if maritime activity is the selection criteria? Why not other areas of interest? You could choose many other criteria (oceanic circulation, upwelling areas, heavy precipitation prone coastal areas, high biodiversity areas, highest MHW intensity in 2022,...). Please, better explain why using maritime activity. Later in the same paragraph biological importance of these two areas is explained.

We have already justified why these two regions are important. Lines 182-231. The reviewer is right to point out the economic and ecologic importance of other areas, although that argument can be expanded to most of the Mediterranean Sea. Unfortunately there is not enough space for more case studies in the Ocean State Report.

Lines 156-157 *"This activity is indicative of the conditions experienced by the rest of the western part of the Mediterranean basin"*. How do you support this statement? Not all the western basin experienced MHW and its intensity showed noticeable variability. Do you mean that the Ligurian Sea is a proxy for the whole WMED?

This sentence has been removed.

Line 159: Please add a value (percentage?) for *"the vast majority of days"*.

We have added the value of 89%.

Line 170: *"upon visual inspection the forecast temperature was very similar to the observed"*.

Could you add some mean bias or other accuracy measure?

We have now added a Table of Root-Mean-Square differences (Table 1), with normalised values to indicate the magnitude of errors w.r.t natural variability.

Table 1: Root-Mean-Square Differences of forecasts of summer 2022 MHW activity and atmospheric conditions (Fig, 4). Values in parenthesis are RMSD values normalised by standard deviation over the summer. Differences in MHW activity and area are relative to reprocessed satellite observations, while differences jn T2M anomaly and windspeed are relative to ECMWF analysis. Each column corresponds to a different lead time.

RMSD (Normalised)	Lead: 1 day	Lead: 4 days	Lead: 7 days
MHW Activity	0.16 (0.48)	0.20 (0.59)	0.28 (0.82)
MHW Area	8.88 (0.33)	11.65 (0.43)	16.50 (0.61)
T2M Anomaly	0.18 (0.21)	0.31 (0.38)	0.52 (0.62)
Wind Speed	0.22 (0.18)	0.52 (0.42)	0.94 (0.76)

Lines 171-176: Is the analysed accuracy basin-wide? Or for the two previously mentioned areas? You should extend this paragraph as it does not provide a clear idea of the forecast accuracy.

We have added specific values of accuracy to the two regions (e.g. Line 215), as well as to the Western Mediterranean Sea (Table 1.)

Line 180: Another definition of activity. Please check out consistency of activity definition across the text.

We have changed "activity" to "occurrence" where necessary.

Lines 180-181: *"Other contributions to this report also define 180 the MHW of 2022 as a record-breaking event, using various other definitions."* Please, better explain this sentence. It does not make sense in its actual form.

We have modified this phrase: *"Other contributions to the Ocean State Report 8 also define the MHW of 2022 as a record-breaking event, using other definitions (e.g. local SST records)."*

Line 188: Please use an acronym or the full name for the Copernicus Med Phys forecasting system instead of *"the system"*.

We have used "MedFS" where necessary.

Lines 188—192: Yours and other author results show the CMPF system is capable to forecast a range of extreme events, but this is not a conclusion of your work. This would fit better in the introduction or methods section to justify the use of CMPF data.

This point has now been moved to the introduction.

Figures

Figure 1b-c. Please, add the grey line to the legend. I assume it is satellite data.

We have added the grey line which represents reprocessed satellite data.

Figure 3: Please, change colour for Near-Real-Time Obs. And MEDFS Analysis. They are too similar and confusing right now.

We have changed colours.

Figure 4: Please, add the satellite to the legend and caption (thick grey?)

This has been added to the legend and caption.

General recommendation: Some lines on the figures look hand-drawn and some of them overlap the other so it is difficult to understand. Please, try to improve figures readability.

We have attempted to improve the clarity of the figures, especially Figures 3 and 4 which contain lots of information.