

## Review of egosphere-2023-1709: “Investigating extreme marine summers in the Mediterranean Sea”.

### General review

In this manuscript, the authors introduce the new (to my knowledge) and interesting concept of *extreme marine summers*. The manuscript presents a huge amount of work from the analysis of EMS to the different drivers/mechanisms involved in such EMS with interesting results.

But, in my opinion, there is a fundamental problem regarding the definition of EMS. When reading the manuscript, I could not get a clear idea of how EMS are defined. First, the authors define EMS as “*define EMSs, separately at each grid point, as the four summers with the highest average JAS SST, i.e., exceeding the 95th percentile of the 71 available summer periods from 1950 to 2020*”. I would expect a basin definition for EMS, maybe coming after the grid point analysis. As we are talking about a season, I would expect that the categorization as extreme would affect the whole region (not with the same intensity in all grid points but with a big enough extent). Later in the manuscript the authors analyse different variables for the whole basin, calculate mean summer values,... maybe suggesting a basin-wide concept. But later there is a section named “*Differences among local extreme marine summers*”.

Another question that is not clarified is the choice of the 95% percentile. Why not 90% as for extreme air and sea temperature to define extreme events? Why 4 summers? Are there only 4 summers exceeding the percentile? Is this needed for any single grid-point or as a basin mean? Are the 4 summers the same for the whole basin?

Figure 4 shows results for three grid points where EMS occur in different years. I don't feel comfortable with a year being extreme in WMED but not in CMED. Or having a pixel suffering an EMS but not the closest one (figure 10 grey areas).

There is, of course, the possibility I did not properly understand the proposed definition. In any case, better explanation of the actual EMS definition is needed.

I also have some major concerns about the analysis of MHWs. Please, see comments below.

The authors have done extensive, very solid and interesting work on the analysis of EMS drivers. A review of the above will add value to all this work and may lead to a really interesting publication. So, I encourage the authors to send a revised version of the manuscript and please understand these criticisms as a way to improve an already good work.

From the above comments, my recommendation is that the manuscript needs major revision before it is accepted for publication.

### **Main comments**

Consider introducing the Martínez et al (2023) paper regarding the impact of detrending SST on MHWs in the Mediterranean: Martínez, J., Elisa Leonelli, F., García-Ladona, E., Garrabou, J., Kersting, D. K., Bensoussan, N., & Pisano, A. (2023). *Evolution of marine heatwaves in warming seas: the Mediterranean Sea case study*. *Frontiers in Marine Science*. <https://doi.org/10.3389/fmars.2023.1193164>

### 2.2 Methods

#### 2.2.1 Extreme marine summer definition and associated SST substructures

In lines 142-144 you define “*as the four summers with the highest average JAS SST, i.e., exceeding the 95th percentile of the 71 available summer periods from 1950 to 2020*”. Why four? Are they the only ones exceeding the 95th percentile? Why not another percentile? Have you checked if this percentile excludes some major MHW events? Please, explain clearly the process to select/exclude EMS/non EMS years. Do you choose the percentile first and then there are only four? Do you choose 4 years and look for a common

percentile? I think I understand the methodology, and mostly agree, but please explain better to avoid confusion.

EMS definition needs to be more robustly justified. What will happen when 2022 and 2023 are included in the study? 95<sup>th</sup> percentile will rise and maybe (probably) some of the 4 years will not meet this condition. Which impact can this have in the definition or the subsequent analysis. Please, if the data for 2022 are available, check this question.

### 2.2.2 Detrending time series

In this section Figure 2a is mentioned. In this figure time series for sub-basins are shown and sub-basins are defined in the figure caption. It would be better to explain the sub-basins, if they are used throughout the paper for analysis, in the methodology section and it would be good to add (up to you) a figure showing them or some indication in one of the actual figures. You should also explain how you decided the limits of the basins (22<sup>o</sup>, Sicily strait,...) Is the Tyrrhenian sea included in the WMED or CMED? Please explain.

### 2.2.3 Marine heatwave identification

For the detection of MHW events you use the fixed climatology 1983-2012, which selection is adequately explained. In a recent work Rosselló et al suggest the use of “*a moving baseline covering the 20 years prior to the year under study*”. As you try to remove the effect of multidecadal trend, the use of a moving climatology it could also be useful. This is just a suggestion you can try or at least keep in mind for future work.

Rosselló, P., Pascual, A., & Combes, V. (2023). Assessing marine heat waves in the Mediterranean Sea: a comparison of fixed and moving baseline methods. *Frontiers in Marine Science*, 10. <https://doi.org/10.3389/fmars.2023.1168368>

Lines 195-196: “*Finally, to focus on summer MHWs, we isolated events with their onset and end day falling within the JAS summer period.*” Have you checked how many events are removed by using this condition? How many MHW days are removed? In the case an event started on June 28 and lasted for 30 days, would it be considered?

Do you apply any spatial extent filter to detect MHWs? Just to avoid single pixel heat spikes.

### 3.1.1 Decomposing the extreme marine summers

Lines 242-243 “*the largest contribution to the mean EMS RDA at this location comes from the warmest part of the SST distribution (higher rank days), with the exception of EMS 1999 where the middle part contributes the most.*” Has this been calculated? Is there a numerical value? I can understand from looking at figure 4a but, what exactly is “the middle part”? Do you divide the rank daily anomaly in 3 periods of 30 days? Do you divide the RDA by some anomaly threshold?

Line 244: Why “*the 60 warmest rank days*”?

I suppose you are following the methods of Roethlisberger (2020) but a clearer explanation would be good.

In this section you comment figure 3 with contributions of the warmest/coldest half to the EMS but in the analysis of the three locations you refer to “the middle part”. It would be better to use the same process for EMS analysis/characterization. Or explain what you mean as “the middle part”? Is it a percentile interval of the rank days? Is there an anomaly threshold to define the RDA parts? You also mention “the 60 warmest rank days” for 2003. Please clearly explain the methodology/process to analyse rank days series.

### 3.1.3 SST substructures 300 using detrended data

If I understood well, there are no relevant differences in the analysis from detrended SST data.

### 3.2.2 Marine heatwave properties during extreme marine summers

Line 384 “EMSs exhibit lower intensity relative to the mean summer conditions”. Do you mean “mean summer MHW conditions”?

Lines 386-387 “This suggests either the suppression of MHWs in EMSs or the existence of alternative mechanisms potentially enhancing the MHW conditions in EMSs”. Please, rephrase in a clearer way (can’t understand).

Line 389: You say MHWs in EMS are more frequent in a just 4 summers series. I’m not sure this can be statistically sound. Please, consider the count of summer days meeting MHW conditions as an alternative variable to analyse.

Lines 395-397: Please, reconsider the frequency analysis.

Line 400: Have you checked that MHWs occur in the “warmer summer days” during the 4 EMS? You can indeed have MHW events in the “colder summer days”. How do you attribute MHWs to the different parts of the RDA?

Lines 414-415 “Given, however, the smaller variability of daily SST values during summers in the EMED, we conclude that the relative role of MHW events in the formation of EMSs is more pronounced in the EMED”. I’m not sure this is strongly supported from the previous lines. It seems logical but it needs a more robust support. Please, extend this paragraph to support this affirmation.

Figure 6 (d,e,f). These figures show some MHW anomalies in the EMS. I see all grid points in the Mediterranean present MHW anomalies. Do you mean that any place in the Mediterranean experienced an MHW in EMS, at least one time? I’m not sure if these maps make sense, unless you can state that there have been MHWs in all grid points during EMS.

Table 1: In caption you say “EMS identified in 1950-2020”. Please, state that this are only 4 years.

This whole section needs to be reconsidered. I don’t think that the way the four EMS are compared to the mean 1950-2020 series is appropriate (see previous comments and figure 6) comments. Consider that the four EMS experienced strong and/or long events. You should also analyse which Med areas were affected to properly decide if basin averages are adequate. It is possible that MHW events did not affect the majority of the basin so spatial MHW averages could not be appropriate, please check and justify your assumptions.

### 3.3.4 Illustrative example of extreme marine summer 2015

Lines 658-659: What does “Summer 2015 has been the most widely experienced EMS in the basin” mean? Is it possible that part of the basin experiences EMS but not the whole basin? Does EMS refer to extreme summer for the whole basin or not?

Line 659: “non-grey locations in figure 8”. Does this refer to figure 10?

Figure 10: What are the grey areas? Non-EMS points?

Figure 10: “Non-coloured sea grid points in the Mediterranean Sea stand for locations that did not experience summer 2015 as extreme”. Does your definition apply only to grid points? Is there any EMS for the whole basin? Is there an spatial extent threshold to declare a summer as an EMS for the basin?

### 3.3.5 Air-sea heat fluxes using detrended data

Do you detrend all datasets? Wind, SWR,... ?

***Minor comments***

Lines 45-49 do not need to be a separate paragraph; they could join the previous one.

Figure 9: Numbers in legend for figures a, b and c lead to confusion, please place them in the middle of the colour interval they stand for. The caption is too long.

Figure 10: Use different colours for land/Atlantic/Black sea and Med sea areas. Same comment for legends than for figure 9.