Reply to Reviewer and changes in the revised manuscript

Reviewer’s comments:

Having read the authors' response to my previous review I still have doubts about the authors' definition of EMS. Although in general they have satisfactorily answered or clarified most of the issues I raised in the review I cannot yet recommend the publication of this paper.

Thanks to the authors' detailed explanations I have clearly understood their definition of EMS but I must say that, even understanding their intention, I think their concept of EMS is incomplete to my understanding and I do not see the usefulness of its application in research on SST and MHWs in the Mediterranean.

First of all, reading the title, the reader would expect a definition that includes both temporal and spatial aspects, since it speaks of "Mediterranean Sea". The fact that the definition is based on a grid point approach without any spatial consideration, even if it is posterior, I believe detracts from the interest or usefulness of the EMS concept. The definition, as it stands, may lead to a theoretical situation, although unlikely but not impossible by the way it is defined, where the 4 years qualified as EMS may be different at each neighbour point in the grid, which makes the EMS concept less consistent. Knowing about the spatio-temporal coherence of the SST, it is possible that in a large part of the points of the grid the years qualified as EMS coincide, reducing the theoretically possible variability, but this is something that the authors should check/show in their results. In fact, in the figures representing three points in the Mediterranean there are few coincidences in the EMS years. I do not feel comfortable looking at a map of, for example, mean SST anomaly without knowing if the averages at nearby, or distant, points have been made over the same years (or a short series of years). The authors should report whether the years considered EMS are a short distribution or whether there is a large set of EMS years with large variability. A suggestion would be to evaluate the number of grid points having an EMS each year to look for years concentrating the highest number of grid points.

Moreover, the definition would only be consistent for the period under study. In the years 2022 and 2023, very intense and extensive SST anomalies and MHWs have been recorded in the Mediterranean that would most likely fall into the group of 4 detectable EMSs in the analysis. Only a slight extension of the series could change the nature of 50% of the EMS over large areas of the basin. This is a major weakness of the definition even though the authors have performed analyses that indicate that the spatial distribution or substructure of SSTs does not change, which is interesting. But the EMS years could change. From such a definition one would expect that subsequent years could be analysed and determined to be extreme or non-extreme without affecting the qualification of previous years, at least not 50% of them.

At this point, I cannot recommend the publication of this article unless the authors consider introducing a requirement for spatial coherence. In this regard I recommend the publication of one of the co-authors "Flaounas, E., et al (2021). Extreme wet seasons – their definition and relationship with synoptic-scale weather systems. Weather and Climate Dynamics", specifically section 2.2 "Spatial and temporal coherence of extreme wet seasons" for a similar analysis.

I encourage you to run the analysis suggested as an improved definition of EMS could be an interesting/important metric for SST analysis both for scientists and stakeholders.

Authors' response:

We would like to thank the Reviewer for their careful reading and the fruitful comments. In the following we reply to each one of the queries.

This study focuses on the Mediterranean Sea and therefore the title accurately provides the geographical boundaries of our research. Regarding the grid-point definition of EMSs, our approach is primarily -along with its starting point- statistical. An EMS is therefore defined as a period with a very low probability of occurrence. This allows us to investigate separately each grid point with regards to the substructures of its own distribution of SSTs. Such a statistical analysis would be valid either we considered each grid point, different sub-regions of the basin or the Mediterranean Sea as a whole.

An alternative method as the "patches-approach" in Flaounas et al. (2021) suggested by the Reviewer would be also feasible. Nevertheless, this would be a data-driven approach and thus fundamentally different since it would prohibit a statistical analysis of substructures as done in this study. Indeed, the formation of patches with coherent physical (or SST) characteristics would be a data-driven one, connecting seasonally biased SST distributions and making it hard to consider EMSs as events of equivalently low-probabilities among the patches. We agree with the Reviewer that the spatio-temporal coherence of the SST would hardly make neighbor grid points to be considered as independent but here we investigate whether the low probability of forming an EMS has a common statistical origin (e.g. due to the warmer SSTs being warmer than usual). With this, we quantify the statistical origin of EMSs across the Mediterranean basin, while the second part of our study aims to physically interpret our statistical results of the first part. To be more clear on the statistical motivation of our approach we now enrich the explanations on the EMS definition in Methods.

Given our reply on the statistical motivation of our methodological approach, a question might be raised about whether EMSs are indeed extraordinary events with large spatial coverage and significant impact on the region. To address this question, we follow the Reviewer's suggestion and we first present the spatial extent of all identified EMSs in terms of their percentage coverage of the basin, illustrated in the bar graph in Fig. 1. As anticipated, the EMS-years are predominantly concentrated in the post 2000 period. Some exceptions during the 50-60's are largely attributed to the multi-decadal oscillation of the Mediterranean SST, as minimum SST values within the study period 1950-2020 occur during the 70's (Fig. 2a in the manuscript). None of the years within 1963-1990 have been characterized as extreme in any location within the basin.

The EMS-years presenting the greatest spatial coverage concentrating more than 50% of the grid points, are 2015, 2003, 2018 and 2012, covering 72% of the basin, 58%, 54% and 53%, respectively. These important (form a basin-wide perspective) EMSs exhibit different spatial patterns (Fig. 2). Such spatial variations among EMSs and across the basin are naturally expected, as diverse factors acting at different spatial and temporal scales within the basin may lead to above-threshold mean summer local SST values. For example, in Sect. 3.3.5 (Illustrative example of EMS 2015), we demonstrate different mechanisms operating over the western and the central/eastern basins during summer 2015 (this summer has not been qualified as EMS in only 28% of the basin). To address the Reviewer's query on the spatial coherence of SST during the aforementioned summers, Fig. 2 shows the locations experiencing the same EMS.

The spatial coherence of EMSs in Fig. 2 is rather evident, however, enclosed sub-basins and areas more susceptible to typical wind regimes (e.g., Aegean Sea and the Gulf of Lions) or areas of distinct circulation patterns (e.g., Alboran Sea), are more prone to the effect of local processes; therefore, they are further expected to differentiate from neighboring areas. For those interested in understanding the physical mechanisms underlying the spatial spread of extreme warm conditions during a specific EMS, the role of local processes in spreading or dissipating warm anomalies during that summer should be investigated across the areas of interest, in line with their objectives (beyond the scope of the current work). The per grid point analysis we follow here captures the extreme conditions experienced at local scale. Transition from EMS to non-EMS neighboring points is expected, as with any peak-over-threshold approach. This does not contradict the spatial coherence of SST, as observed in Fig. 2; it merely reflects that warm conditions are not equally extreme across different –even neighboring– areas.

We now present in the manuscript the spatial coverage of the most extended EMS conditions, (maps of Fig. 2 depicting areas experiencing specific EMSs). In this way, the reader will have a clearer idea on the EMS detection output based on this additional information on the spatial extent of the most significant EMS cases.

As previously noted in this reply document, we regard an EMS as an event of low probability. The question that the Reviewer raises with their query is whether EMSs are also to be statistically consistent with similar substructures and similar interpretations. Our analysis as correctly pointed out by the Reviewer shows that indeed EMSs are statistically consistent. The spatial distribution of SST substructures remains unchanged when we apply our analysis on the entire study period, parts of the study period, or based on the detrended dataset. In all these cases we observe the same statistical behavior: local SST substructures remain the same despite the different sets of 4 EMSs detected at each grid point. In the manuscript, this is evident through the comparison of the SST substructure fields derived from the original vs the detrended dataset (Fig. 3a-c vs Fig. 3d-f of the manuscript). The example distributions of Fig. 4 of the manuscript illustrate this in more detail. They show that the part of the SST distribution contributing the most to the seasonal SST anomaly at a specific location is the same in the original and the detrended dataset although three out of the four EMS-years differ between the two datasets (e.g., the anomalously warm colder summer days in the southern-central basin; Fig. 4c vs Fig. 4f of the manuscript). This finding is important, implying inherent characteristics of the anomalously high SST during summers in the

Mediterranean basin. More than that our results suggest that inherent characteristics of EMSs are unrelated to the specificities of individual summers or the climate-change warming signal. Therefore, including data from 2022 is not expected to significantly alter our results even if the EMSs are not the same.

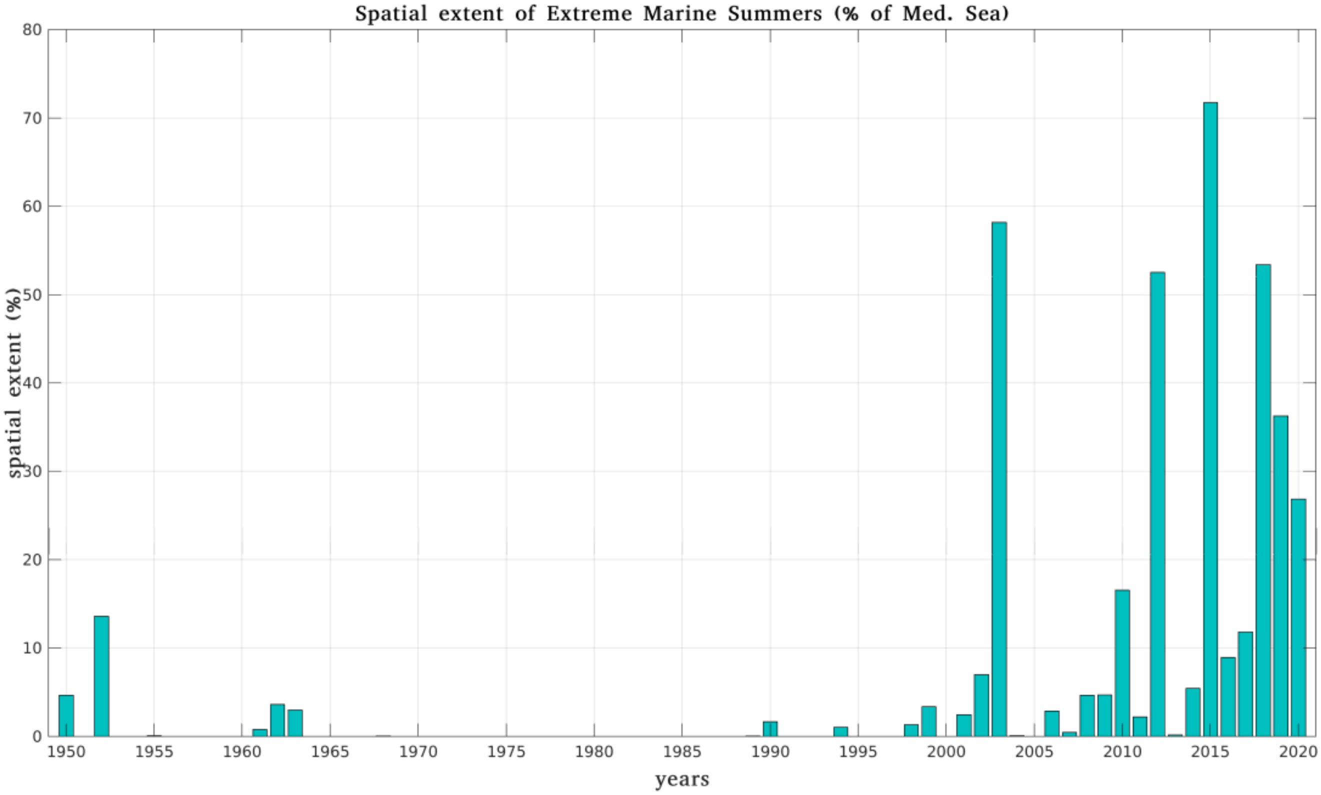

Figure 1: Percentage of Mediterranean Sea grid points experiencing EMS conditions

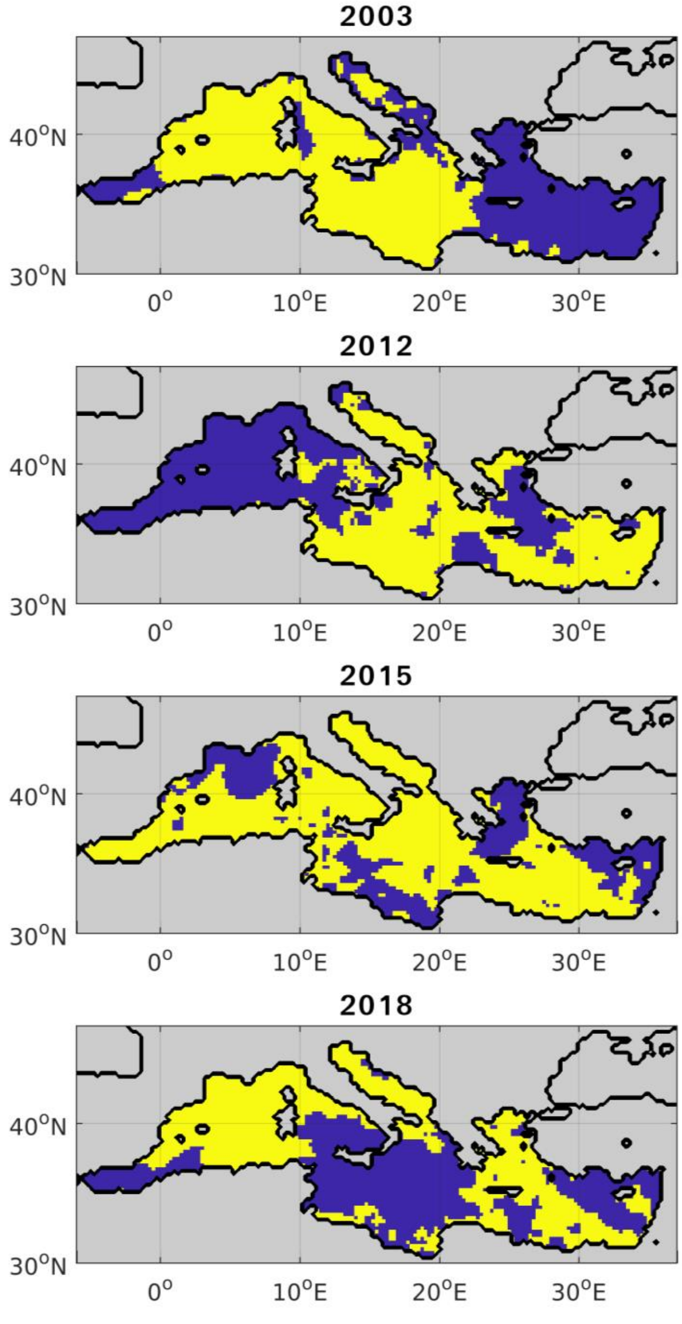

Figure 2: EMSs concentrating more than 50% of the Mediterranean Sea grid points: 2003, 2012, 2015 and 2018 (from top to bottom, respectively). Yellow stands for areas experiencing EMS conditions and blue for non-EMS conditions

Changes in the revised manuscript:

To be more clear on the statistical motivation of our approach and provide additional information on the spatial extent of the most significant EMS cases, the explanations on the EMS definition and detection in Methods have been further enriched as follows.

Lines 147-154 (in the revised manuscript):

(…) This approach captures periods with a low probability of occurrence experienced locally, allowing for a statistical analysis of their properties in a consistent way across the basin, based on the methodology of Röthlisberger et al. (2020). In this way we investigate separately each grid point with regards to the substructures of its own distribution of SSTs. To offer the reader an understanding on the EMS detection output regarding the spatial extent of the most significant EMS-years, Fig. B1 illustrates the spatial extent of EMSs concentrating at least 50% of the MS grid points (2015, 2003, 2018 and 2012, covering 72% of the basin, 58%, 54% and 53%, respectively).

Figure 2 (depicting areas experiencing the most extended EMSs) has been included in the supplementary figures of Appendix B (Fig. B1)