Review: **Investigating extreme marine summers in the Mediterranean Sea**

Major Comments:

Abstract:
In general the abstract quotes the main results of the paper. However, the way it is currently written does not form a clear and concise story, rather it presents a series of results concerning the EMSs, without showing how are these connected between them. The entire abstract needs to be re-written in a more cohesive manner, forming an "argument", which shows what is known for the SST/MHW/EMS of the Med. Sea, what is the current problem faced and what is the current study's aim/methodology/added value on the subject tackled.

Introduction:
The introduction is very long and currently structured as:
- Importance of SST (Lines 22-30)
- Past Med. Sea SST trends (Lines 31-43). Such a detailed information on the SST variability is not necessary in the Introduction of the paper. Rather on the discussion section where the writers can compare their results with previous literature.
- Future Med. Sea SST trends (Lines 45-48)
- Decadal variability of Med. Sea SST (Lines 50-64)
- MHW trends (Lines 65 - 71)
- MHW impacts on marine life (Lines 73 - 82)
- Motivation for focusing on summer Med.SST (Lines 84 - 92)
- Explanation of the EMS concept (this belongs to Methods section) (Lines 94 -105). This paragraph describes the concept of EMS very vaguely and the reader can get confused. Is the EMS based on the MHW drivers or something else? Better to describe this in Methods and not here.
- Drivers of EMS. This again belongs to Methods (Lines 107 – 114) and it needs to be clarified if it is a paragraph on the drivers of EMS or on motivation for the investigation of the air-sea heat flux effect on EMS. See comments below
- Paper objectives (Lines 116-120). This is the only clear paragraph of the Introduction so far. However, the language needs to be improved. In addition the 4 objectives of the paper can be re-arranged to:
- Structure of daily SST in EMTs
- Role of SST multi-decadal variability of in the EMS characteristics
- Role of MHWs in EMS
- Drivers of EMS

Once the objectives of the paper are clear then the Introduction structure should describe brief background knowledge on these aspects. Also the writers should define better and briefly the EMS in the Introduction, making clear that EMS is something different from the MHWs, because currently it is not clear. The reader is confused between MHWs and EMS, and whether qualitatively and quantitatively they are the same thing or not or if one is driving the other. In general, it is not clear what the relationship between the two is and what is the added value of the EMS (or what is an EMS in general) compared to MHWs. The writers need to remove some sections from the Introduction that belong mostly to methods and should devote a brief paragraph on explaining the idea behind the EMS and its substructures here.
Additionally, in many paragraphs it is not always clear whether the writers refer to MHW/SST characteristics in the global ocean or in the Med.Sea (for example in lines 84-92), whereas the Introduction section should be more devoted to conditions relevant for

the basin, as it progresses. The relevant information from the global ocean (SST) characteristics currently are given here and there but should be used less as a motivation for a study on the basin. Also the period examined should be briefly mentioned in the objectives paragraph.

Given the entire paper needs restructuring (see general comments to the editor), the writers should more clearly define the objective of their paper and then write an introduction providing information relevant for the paper and not just general characteristics of Mediterranean SST.
Alternative structure of the Introduction could be:
1. Brief description of Mediterranean SST, trends, decadal variability, future trends & its importance
2.Briew review of MHWs in the Med.Sea
3. Brief introduction of EMS idea and how this differs from MHWs
4. Motivation of the study/objectives.

Methods:
The definition of the EMS has fundamental problems:
1) "*We then 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*". If the four summer periods (JAS) where the SST exceeds the 95th percentile are selected then it is likely these 4 summers will be towards the end of the 1950-2020(This is the first impression the readers get when they start reading the manuscript since the de-trending of the SST comes as an explanation much later in the Methods, which is confusing). In addition why the four warmest summers and not 5 or less summers? Or why not all the summers which are above the 95th percentile in the 1950-2020 period?Is there a specific reason for this choice? Have any sensitivity tests been performed for this choice? The choice of the 4 summers has to be well justified. Is it based on any physical quantity or previous study?

2) "*By splitting the summer rank days into different parts of the SST distribution (e.g., the coldest and the warmest half), we quantify the relative contribution of each part to the observed EMS RDA*." I do not see that kind of processing anywhere in Figure 1 so I am not really sure I understand what do the writers mean here. What do you mean by "*contribution of the observed EMS RDA*"? I only see Rank Day anomalies in Fig.1c.

Lines 170-173: "*The trend value of each summer season is also removed from all days belonging to this summer to obtain a detrended summer dataset of daily SST values as well*". On top of the de-trending of the multi-decadal SST you also remove the summer trend: Only from the already detrended dataset or from the non-detrended dataset as well? Why is this? What does that summer detrending offer in the study? Also, which months represent the summer detrending here? Jun-August or the Jul-Sep months used to define the EMS? How do you calculate the summer trend? Is it based on the daily summer SST of each year or the trend of the yearly-averaged SST?It is not clear. And what are the days that "*belonging to this (which) summer*" the de-trending is happening on? The choice of the summer detrending has to be justified and also the method by which the summer detrending happened has to be mentioned somewhere. At the moment this step of the method is not clearly explained.

- The term "SST substructure" could be better replaced by "SST distribution" since the ranking of daily SST resembles an SST distribution.

Alternative structure of Methods section is
2.1 Datasets
2.2 Definition and Detection of Extreme Marine Summers (Currently named subsections 2.2.1 and 2.2.2 belong to the same section). This is also because the detrending of the SST timeseries should be explained along with the definition of the EMSs. Otherwise the readers get the impression that the EMS are the four warm summers towards the end of the period. These two subsections need to be re-written and merged in the right order so that the reader undestands the SST warming trend does not affect the EMS definition. Currently the definition of the EMS is scattered a bit here and there in these paragraphs. Also, justification of the choice of definition/examples on the SST substructures & SST de-trending
2.3 If the idea of the EMS idea is retained then it has to be clearly differentiated from the MHW definition and then it would merit giving a brief descriptions of the MHWs.
2.4 Description of EMS drivers a) Air-sea heat fluxes, b) Upper ocean preconditioning (Sections currretly named as 2.2.4 and 2.2.5 could be merged)
 - Description of the P_{EMS} metric.

Results:

The Results section should be differentiated from the Discussion section (that needs to be created) by providing solely indicative numbers on the variables described. There has to be a quantification of the results and not a qualititative description based on visuals.
Re-structuring of Results section could be:
1. Characteristics of EMSs (in original and detrended dataset). Not sure what is the essential difference between Sections 3.1.1 and 3.1.2 but they could be merged and shortened into one clear message.
2. Role of MHWs in EMS (?) (if EMS are clearly differentiated from MHWs)
- MHW detection
- MHW properties (original & detrended datasets)
3. Drivers of EMS. Merging of Sections 3.3.1, 3.3.2,3.3.3 where one quantifiable method of air-sea heat fluxes contribution should be chosen and described with indicative numbers.
 - original dataset
 - detrended dataset
 - Case study the MHW/EMS 2015.

Section 3.1.1:
- I understand the rational behind the focus on the contribution of the warmest and coolest half summer days in Figures 3, however, what about the days that are ranked in the middle of the summer? What about their contribution? At the moment the method sort of examines the SST distribution by only looking at its extreme statistical moments (the warmest and coolest quantiles of the distribution). But how about the mean of the SST distribution? Does this change too apart from the warmest days getting warmer? It is well known that there is a mean SST shift which drivers warm days to become warmer in the basin.
- More importantly, how is the contribution calculated? The writers simply refer to a contribution to an EMS. How is the EMS represented here as a quantity whose contribution is split into warmest and coldest days? Not clear.

Section 3.2.2
Line 389:"…*Supporting the latter scenario, MHWs in EMSs appear longer lasting as well as more frequent..*".Although increased frequency and duration of MHWs is a very

commonly highlighted result in most studies,i am first curious about a) the increased frequency. In the selected examples of Fig.4 the EMS years were only 4 out of the entire period. How can MHWs be more frequent when the EMS years are only 4 out of 70 years (1950-2020)?Or how can you compare the MHWs during EMS years which (drawn from the example of Fig.4) are less in number compared to the entire period of 70 years? And how can this comparison yield an increased frequency? How can there be an increased frequency of events that are happening in a limited number of years? Unless the writers mean something else here. The only way for a particular point to present more MHWs during EMS period compared to the normal summer periods is if these events are very short and somehow they happen very often, much more often than the normal summer. This however, contradicts the increased duration found for most points in the Med.Sea. In general this comparison of MHW characteristics during EMS and non-EMS years does not seem to have a solid foundation in a sense that the sample of EMS years is very small compared to the average summer period examined. To this end, I do not see the point of comparing the frequencies of MHWs. I am not sure what kind of quantities are compared here.

Section 3.3.1

 - Any kind of attempt to correlate the spatial distributions (maps) between the different air-sea heat fluxes should be accompanied by at least a pattern correlation coefficient. At the moment the comparison are described solely as visuals, without any indicative numbers or coefficients denoting correlation between the different variables. Otherwise the claims made by the writers are not substantiated. I suggest the writers re-write the entire section, at least performing pattern correlations coefficient between the maps of this section. I feel this section belongs better to a Discussion section where the writers need to be careful not to repeat the results from section 3.3.3, where the contributions od the different air-sea heat fluxes are actually quantified. Section 3.3.3 belongs better to the Results section.

Lines 495-496: "*Negative MLD anomalies in the western part of the Gulf of Lions also imply wind-induced mixing reduction in the vertical (Fig. 7f,g)*". First of all, I am not sure whether the MLD anomalies shown in Fig(7g) can be directly compared and linked with the anomalies of the rest of the Fig.7 figures. According to the caption of Fig.7 all the anomalies were calculated relative to the 1950-2020 period, whereas the MLD ones were based on the 1987-2019 period. Does that mean that the MLD anomalies were also computed for the EMS summers with respect to the non-EMS summers of 1987-2019 or that the writers simply created a climatology of the MLD for 1987-2019 and then computed the anomalies? Clarify. In any case, the comparison of anomalies between 1987-2020 and 1950-2020 is not advisable since the period before the 1987 encompasses a lot of the non-climate change signal. The writers should compute the anomalies of all the variable for the same period in order to be able to substantiate any claims on the link between the spatial distribution of wind and the rest of the air-sea heat flux variables. At the moment, there is no point in connecting these fields between them as their anomalies depict different things here.

Discussion:

- Currently there is no Discussion section in the paper and Results section contains comments which could belong to the discussion section of the paper. The entire paper needs restructuring as well as re-writing with the help from a native speaker.

<u>Minor Comments</u>:

Abstract:
Line 6: Suggestion to re-write:"*The Mediterranean Sea (MS) has ~~been~~ experienc~~ing~~ significant surface warming, ~~particularly pronounced~~ during* **the** *summers when? Which years?]* ~~and~~ *,* **which was** *associated with devastating impacts*" to *(on what?where? Ecosystems? Society?)*"
Line 9: What does "*SST substructure*" mean here? For someone that reads this term the first time in the abstract it is not clear. Maybe write it differently?
Line 10: "*...identified in most of the basin...*". Do the writers mean that the EMSs appear more often covering a large area of the Mediterranean basin at the same time or that EMSs usually occur in many of the sub-basins of the Med.Sea (but not simultaneously necessarily)?. Better to rephrase this sentence.
Line 11: When the writer refers to "*ranked daily SST*", ranking can be anything and relative to different things. Here it is not explained what kind of ranking is used. It is a bit of a vague statement.
Line 13: "*..and more frequent than usual..*". Does "usual" mean relative to the climatology of the summer MHWs or the climatology of the summer period used to define the EMS (which by the way it is not mentioned and it need to be given as an information)? Also "*....mainly in* **the** *northern MS regions.*"
Line 14: "*However, the relative contribution of MHWs in EMSs is more pronounced in the central and eastern basin*". In the previous sentence the writers mention the characteristics of MHWs, which are more pronounced in the northern part of the basin during an EMS period. This sentence however, refers to a contribution (what kind of contribution is this?) of MHWs in the EMS, which is a bit confusing that is more pronounced in a different part of the basin. Also, by definition of "Extreme Marine Summers" one can understand that EMS is a different type of MHW and this sentence refers to a contribution of MHW to the EMS? It is a bit confusing. Consider rephrasing it.
Line 17: "*Upper ocean preconditioning is* **also important for** *the EMS formation..*"
Line 20: "*...regardless of* **the** long-term trends".

Introduction:
Line 22: "**The** *global ocean..*". The end of this sentence needs a citation.
Line 23-24: "*..Environmental and societal implications –current and projected in the future– underline the need for continuous ocean monitoring and deeper understanding of the ocean climate, in terms of natural variability and anthropogenic climate change*" could be re-written as:
"The current e*nvironmental and societal implications highlight the need to improve our understanding of the anthropogenic forcing influence on the ocean climate through continuous ocean monitoring*".
Lines 25-30: *These lines could be re-written and re-arranged as shown :*
"**The** *Sea Surface Temperature (SST) is a fundamental climate variable and global climate change indicator. A rapidly growing literature has shown its key role in the intensification of atmospheric/oceanic events and processes, e.g., heavy precipitation events (Pastor et al., 2015), surface air temperature variations (Xu et al., 2019), MHWs under global warming (Frölicher et al., 2018), increase in global wave power (Kaur et al., 2021).* **More** *importantly, SST is the oceanic parameter that regulates air-sea energy exchanges, reflecting the role of the ocean's thermal inertia (Deser at al., 2010)."*
Lines 31-44: I am not sure if such a long overview of the SST trends is needed in this introduction. I would better save the information for comparison with the paper's result in the discussion section. The writers could very briefly give a range of the documented SST trends from all the literature and continue with the paragraph of lines 45-49.

Lines 47-49: *The sentence could be re-written as: "The high-emission scenarios (SSP5-8.5) of the CMIP6 multi-model projections suggest an SST increase of 0.8° C to 3.5° C in the near- (2021-2040) and long-term (2081-2100) 21ˢᵗ century, respectively, relative to 1995-2014 (Iturbide et al., 2021).*

Line 57-59: *".. By that time, AMO has entered a declining phase, in agreement with the observed upper ocean cooling in the Atlantic that reversed the previous warming trends (Robson et al., 2016)".* Previous warming trends of the Atlantic or the Med.Sea?Not clear.

Line 65: Instead of "*...warm extreme oceanic events such as Marine Heatwaves (MHW), have attracted great research interest...*" better to write: "..extreme warm ocean temperature events, such as Marine heatwaves have gained great research interest.."

Line 66-71: These sentences could be re-written as: "***Over the past decades***, ***an** increased MHW intensity and frequency have been documented based on observational and modelled SST datasets, (Oliver et al., 2018; Holbrook et al., 2019; Darmaraki et al., 2019a; Juza et al., 2022; Pastor and Khodayar, 2023; Dayan et al., 2023). Further **increase of MHWs trends** is expected at **a** global and Mediterranean scale **over the 21ˢᵗ century** (Oliver et al., 2019; Darmaraki et al., 2019b; Plecha and Soares, 2019; Hayashida et al., 2020), due to anthropogenic forcing and especially under high-emission scenarios (Oliver et al., 2019).*"

Line 73: "*...warming trends and warm extremes is ~~highly~~ motivated by their detrimental impacts...*"

Line 79: "*...Climate-related local extinctions..*" of what exactly?

Line 84: *"Summer periods are of particular interest as they are associated with greater surface warming both in present and future climate studies".* The greater surface warming is a given on summers of any time period (past or future). The reason behind the interest in summers is that this elevated warming has profound effects on marine ecosystems/communities and not because it is associated with elevated warming in current and future climate studies.

Line 85: "*...~~Actually~~, ~~t~~The Mediterranean warming rates calculated...*"

Line 86: ""*..~~Additionally~~, Indeed model projections suggest that that maximum SST increase is expected in summers*". Are you referring to the Med.Sea or to the ocean in general? Better to remain consistent and talk about the Med. basin.

Line 87: "*..Moreover, summer MHWs present the highest SST anomalies (Gupta et al., 2020) and are associated with a stronger ecological footprint (Oliver et al., 2019)*". Is this true on average for the global ocean or the Med. Sea? Better to build the argument behind the choice to investigate summer SST anomalies based on similar facts/studies about the Med.Sea.

Lines 84-92: Too many sentences with too many connecting words (e.g. Additionally, moreover, In addition etc). It is like similar arguments are being "thrown" one after the other but without forming a clear argument on the importance of summer SST. Re-write this paragraph, maybe starting with the reasons of summer SST importance step by step.

Lines 94-95: "*Considering the above, we propose the concept of extreme marine summer (EMS), from **an** ocean perspective, and we ~~put our effort to~~ explore **here** Mediterranean EMSs in a climatological framework.*"What do you mean by the climatological framework?

Lines 96-105: The entire paragraph here needs re-writing (preferably with help from a native speaker). Also it does not belong to the introduction, rather it belongs to Methods as it describes in detail the concept of EMs. Also when the writers talk about the *mean summer SST*, which years have examinated? It is not stated.

Line 100: "*..due to uniformly increased SST values throughout its duration..*". What do you mean by *uniformly increased SST* ?If we are in a EMS then it is a given that the SST values are going to be elevated throughout its duration,no?So what is the uniformly about?

Line 101: "*…due to warmer SSTs of extreme events alone. This relates to the fact that marine species have..*". What exactly is related with the marine species here? Not clear.

Line 104: "*..For instance, animals able to migrate to avoid anomalously warm water conditions lasting for several days may not be able to cope with longer lasting heat stress (Alexander et al., 2018)*". How much longer is the "longer lasting heat stress" from "a warm event whose duration is many days"? I do not understand this sentence. Is the duration of the EMS that differentiates it from the MHW definition and makes it more pertinent to the marine ecosystems? Or is it the change of the uniform SST change or the change in the SST substructures?

Line 104-105: "*Therefore, even in* **the** *absence of extreme SST values or extreme warm events, a summer season may present extreme mean conditions thereby strongly affecting marine life.*". How can a summer present extreme mean condition if the SST is not extreme and there is no extreme warm event?Confusing.

Line 106: "*An EMS may arise due to elevated SST anomalies* ~~taking place~~ *over the summer* ~~duration..~~ *period*". What do the writers consider here as summer duration? June-August? May- September? Not specified. Also elevated SST anomalies..don't already mean.."favoring of initial thermal conditions", since there is already a positive temperature anomaly?Otherwise the writers need to explain what they mean by "initial thermal conditions".Also the period examined is nowhere to be found.

Lines 107-110: "*The former may result from the interplay of atmospheric and oceanic factors, being the air-sea heat exchanges (turbulent and radiative fluxes), horizontal (Ekman and geostrophic currents) and vertical (entrainment, Ekman pumping) advection (Deser et al., 2010 and standard oceanography references therein)*". There is always an interplay between the atmosphere and the ocean, whether there is an EMS or not. Perhaps the writers meant the EMS emerged due to specific atmospheric and oceanographic conditions in the basin?Also if the former (SST anomalies) may be due to air-sea interactions, the latter (favoring initial thermal conditions) may be due to.. what?

Lines 111-112: "*~~Several studies focusing on the Mediterranean SST ariability have shown t~~The crucial role of* **the** *air-sea heat fluxes in* **the** *Mediterranean SST variability and observed warming trends in particular* **has been shown in several studies** *(e.g., Skliris et al., 2012; Shaltout and Omstedt, 2014)*". Is this paragraph devoted to the drivers of an EMS or to the motivation behind the investigation of the air-sea heat fluxes? This paragraph needs to be re-written and its content has to be more clear.

Lines 113*:"In this context, ~~an extra focus is put within~~ this study* **aims to** ~~on~~ *understand~~ing what is~~ the ~~driving~~ role of air-sea heat fluxes in the formation of EMS".*

Methods:

Line 129: "T*his* **is a** ~~dataset corresponds to foundation SST (SST~~ *free* ~~of~~ *from diurnal variations~~)~~ SST product* ~~and~~ *derived~~s~~ from a combination of* **the** *HadISST2 and OSTIA datasets. Atmosphere ~~ic~~ variables from the ERA5 product are also…*"

Line 129-130:"*… 10-meter wind speed, net short****wave, longwave*** ~~and long-wave~~ *radiation as well as* ~~at the sea surface~~*, latent and sensible surface heat fluxes, total cloud cover~~,~~* **and** *specific humidity*".

Lines 133: "*All fields have a grid spacing of 0.25°x0.25° in longitude and latitude and are provided in hourly time intervals.*". Are you using hourly time interval in the study or not? If not, better to talk about the timestep you are using the datasets in the study. Also better to re-write the sentence and merge it with a previous sentence like this:" *This is a free from diurnal variations SST product derived from a combination of the HadISST2 and OSTIA datasets, with a 0.25°x0.25° spatial resolution, processed in a daily/monthly/hourly timestep*".

Lines 134-136: "*~~In addition to ERA5 (reference SST dataset), we use the CMEMS L4 satellite SST product (EU Copernicus Marine Service Product, 2022c) for the period 1982-~~*

*2019, at 0.05°x0.05° grid spacing, t*To cross check the quality of the reference dataset against a high-resolution observational SST dataset, **we also use the CMEMS L4 satellite SST product (EU Copernicus Marine Service Product, 2022c) for the period 1982-2019, at 0.05°x0.05° spatial resolution"**.

Lines 135-138: Is the CMEMS MLD different from the mixed layer thickness production extracted from the CMEMS reanalysis?Why are they mentioned twice?Confusing.

Line 145: *"After identifying EMSs, we apply the methods of Röthlisberger et al. (2020) to assess the associated SST substructures at each..".* What is an SST substructure? Clarify.

Line 154: *"The example of Fig. 1a shows for a certain grid point that the summer of 2003 has been entirely warmer than..".* When the writers say *"entirely warmer than"* do they mean that the entire SST distribution is anomalously warm (and not only the extreme percentiles)? Then they should define somewhere in the methods that they are looking at the SST distribution moments and their anomalies instead of using words like "entirely".

Line 155-156: *"RDAs in this example (Fig. 1c) are much higher for the higher rank days, suggesting that this summer is identified as extreme primarily due to the warmer summer days being warmer than usual"*. By looking at Fig.1c I see all the days of 2003 being anomalously higher than the average ranking between 1950-2020. All the days have ranking anomalies ranging from 1-3 C. Is there any criterion based on which you decide that the 3C ranking anomaly is the one primarily linked to the 2003 warming? I see all days having an anomalous contribution. So I am not convinced that the warming of 2003 was due to only the warming of the warmest days. The warmest days have the highest anomalies but does that mean that the anomalies of the rest of the days did not play a role in the 2003 warming? I do not understand.

Section 2.2.2: This section should be merged with Lines 143-145 since it is part of the EMS definition. Without this part the readers get confused with the EMS definition which is currently scattered here and there.

Figure 2 caption: Better to give the lat/lon coordinates used to split the Med. Into the different sub-basins instead of saying "Strait of Sicily up to 22E". The science must be reproducible, therefore, the exact coordinates of the subregions should be mentioned in the captions or in the description of the Methods. Also, the legend for the detrended SST timeseries in panel (b) says *"Med SST anomaly"* which is confusing, as usually the SST anomaly refers to the SST difference from the climatology. Better to just refer to the timeseries as "detrended summer SST". The zero line does not need a legend. It is evident.

Line 171-172: Better to decide to use either past or present tense in the study. Currently some sentences are in past and other in present which is confusing.

Lines 175-179: This paragraph may well belong to the introduction where the writers should give a brief overview of the EMSs and their added value. However, if the writers removed the signal of the multi-decadal SST oscillations how did they also remove *"any long-term variability of the Mediterranean SST"* apart from the multi-decadal oscillation one? Did you do an additional de-trending of the SST timeseries or does this refer to the multi-decadal signal? Clarify. This is confusing for the reader.

Line 178: *"...into the ~~actually~~ warmest EMSs and..."*

Line 191: *"...Therefore, such discrepancies resulting from different climatological periods are not considered significant in the scope of the present study"*. What kind of discrepancies do you mean here?Do you test different climatology periods for the detection of MHWs?Why is this information relevant for the study?

Lines 192-195: *"...In addition, the availability of satellite SST data during the selected climatological period....to detrended SST data"*.All these sentences need to be re-phrased with help from a native English speaker. Also, what is the point of mentioning the comparison between the MHWs detected using ERA5 and those using satellite if a) the

comparison is not shown anywhere and b) the results from their comparison are not used anywhere else in the study?

Line 204: "*To investigate the role of surface heat fluxes during EMSs, we first compute the mean anomaly of EMS Qnet and its components with respect to their mean summer value over the study period...*". What do the writers mean here by computing the "*mean anomaly of EMS Qnet with respect to the mean summer value?* When one computes an anomaly and takes its average in time, then this process goes back to the mean value from where the anomalies were created. Clarify.

Line 217: "*Apart from the contribution of positive SST anomalies during the season, warmer than usual initial conditions may also favor...*". It is better to avoid using "initial conditions" since the writers are talking about upper ocean preconditioning in this section. The phrase "initial conditions" usually refers to initial conditions e.g. of a climate model etc.The entire section here, in fact, talks about the contribution of the ocean heat content to the EMS development whereas the previous is simply describing the atmospheric contribution. This is a common method to investigate drivers of extreme warm temperatures in the ocean, therefore the terminology should just refer to atmosphere and ocean heat contributions (or upper ocean preconditioning).

Line 222:"*... and the corresponding EMS anomalies (June before an EMS – mean June for all years) at each grid point..*". Better to write this as an equation instead of using words.

Line 228: "*To overcome any ~~unfair~~ OHC inter-comparisons due to different integration depths..*" Rephrase.

Results:

Line 239: What is the reason behind the choice of these particular 3 points whose SST substructures are investigated? Why not other points of the MS ?Is it random?Justify if not. Also, what is the period examined and represented by the grey lines? Currently we are only aware of the EMS years from the Figure legends. Caption needs to mention the entire period examined.

Line 241: "*Regardless **of** the variability..*"

Figures 3 & 4: What is the (K) in the titles and y axis label of Figures 3a) and 4a) respectively?Also panel 4 somehow is described as a whole before panel 3 of figures. I would suggest moving Figures 4 as Figures 3 and then describe the maps currently named as Figure 3, since anyway their description is now interrupted from the description of Figures 4.

Figure 4. The caption of the figure is missing the reference to plots d,e,f. They are only referred to as bottom row. They should be referenced together with the top row. e.g. a,d) southern of the Gulf of Lions (…) etc.

Lines 266-267: "*To gain insight into the SST substructures over the basin, Figs 3b and 3c show the fractional contributions of the warmest and the coldest half of the RDA to the formation of EMSs*". What do you mean by "*warmest and coldest half*"? Which ranks are the warmest and coldest ones out of the distributions shown here? Not clear.This is important in order to also understand the contribution maps of Figure 3. Also, what is the contribution of the middle ranks? The ones which are not too cold or warm?And how is the contribution calculated?Relative to what quantity? This is not clear.

- The description of Figures 3d-f comes later in Section 3.1.3 and after the description of Figures 4 & 5 which is confusing for the reader. It seems ad-hoc and the reader has to scroll up and down the manuscript. The description of the figures should be done sequentially. So section 3.1.2 should come after section 3.1.3

Line 275: If the variance of the RDA is calculated following the standard mathematical equation, Var = (RDA-mean RDA)^2/(number of summers). Are the writers sure that this quantity applied to the anomalous filed does not bring the Variance back to the mean? And probably this is why the maps of Fig5 resemble a lot the mas of Figure 3?Why did the writers choose the variance and not the standard deviation?

Lines 279-280: "*Results come in agreement with Shaltout and Omstedt (2014) ~~who~~ that examined the SST variability...*"

Lines 279-281: "*Results come in agreement with Shaltout and Omstedt (2014) who examined the SST variability in the basin showing that the maximum and minimum seasonal stability are found close to the southern Levantine sub-basin and the Gulf of Lions, respectively.*". This type of comments belong to the discussion and not in the result section.

Lines 282-283: "*We then calculate the fractional contribution of the coldest and the warmest half of the ranked summer days to the RDA variance for all 71 summers (Figs 5b,c)*". Again which ones are the warmest and coldest half of ranked summer days, you need to name then. Otherwise it is vague. Also how do you calculate this contribution? Not clear.

Lines 286-287: "*In other words, the locations where the warmest (coldest) part of the rank day distribution has higher spread, experience EMSs primarily due to the contribution of the warmest (coldest) part of the EMS rank day distribution*". In Figure 5a) I only see higher spread in northwest Med. Sea where there is also a high contribution from the warmest and coldest half of rank days (from Figure3). However there is also a high contribution from the warmest and coldest half or rank days in e.g. the north Aegean, where the spread (Fig5a) does not seem to be high. How do the writers explain this? Is their statement here true for all the areas of the Med. Sea?

Lines 290: "*At the location at 41.5° N-5° E (Fig. 4a), SST substructures appear particularly different among the EMSs.*".Although this is true I can also discern other years (grey lines) with the "warmest" part of the SST distribution contributing more to the summer temperatures (following the criteria of high values just like the writers). So what is the difference in this contribution with the contribution of the EMS RDA?

Lines 292: "*This is the part presenting the largest RDA spread in the rest of the years as well (6.7° C and 5.4° C for the warmest and the coldest part, respectively) (Fig. 4a)*". Which years? The non EMS or the 3 other EMSs?Not clear. Also, what is the spread? Can you give the variance value?

Line 293: "*Similarly, at the location at 35°N-28°E (Fig. 4b), the highest rank days which contribute the most to the mean EMS RDA are also the most varying from summer to summer within the study period.*". There is no location 35° N-28° E indicated in Fig4b. The point 35° N-28° E is located in the central Levantine basin, which according to Fig5b) is characterised by an increased contribution from the high rank days. However, by looking at Fig5a) I do not see a particularly large spread in the area. In fact,the spread (variance) is much smaller than the one shown for the northwest basin.

Line 295: "*The RDA range of the warmest part here is 3.3° C (2.7° C for the coldest part)*". When "range" is used then two numbers should be given that represent a *range* of what? What do the writers mean something else here. It is confusing.

Lines 296-297: "*In contrast, at the location 34.5° N-13° E (Fig. 4c), which exhibits greater RDA variability in the first rank days of the RDA distribution, RDA values of the coldest half display a range of 4.6° C (3.5° C for the warmest part)*". Looking at Fig 4c) I see the RDA ranging from 1K – 3K (is this Kelvin or something else?) for the coldest half of the rank days. So where does the "range of 4.6 ° C" come from?Unless the writers mean something else.

Line 299: "*...the part of the RDA distribution that contributes the most to the EMS RDA is the one presenting the largest spread climatologically*". Where is this climatological spread shown exactly? Not clear.

Line 300: "*When the multi-decadal trend is removed, smaller SST anomalies in EMSs are found in the entire MS. (Figs 3a,d)*". The anomalies are smaller relative to the climatology but also in comparison with the non-detrended dataset.

Line 306: "*In both fields, RDA peaks are found in the Gulf of Lions and the Ligurian Sea, followed by the Adriatic Sea.*" The Results section is where specific numbers should be

given. Consider mentioning the the RDA peak numbers you are referring to here. In general a range of values that represent big or small differences in the RDA between the left and right column of Fig3 should be given in the Result section.

Lines 313:"*In particular, in the detrended dataset, locations where the coldest rank days contribute to the RDA variability more than 50% are not restricted in the southern-central area north of the African coasts as in the original dataset. It is also the central Adriatic, part of the southern Levantine and the north Aegean Seas that exhibit a slightly enhanced contribution of the cool summer days to the RDA variability (Figs 5e,f).*". Rephrase the sentence and give more representative numbers. The result section is where values should be given, whereas the discussion section is used for commenting on the differences between the numbers of the results section.

Lines 320: "To better illustrate this, the SST substructures in the three example cases discussed above are very similar for the original and the detrended dataset, despite the actual EMSs being warmer (Fig. 4a-c and Fig. 4d-f, respectively)". The actual EMSs being warmer...than what? Not clear.

Lines 320-324: Re-phrase with help from a native speaker.

Lines 326-329: "In addition, SST substructures in EMSs seem to be independent of the selected study period. Sensitivity tests performed for different sub-periods show a consistent statistical behavior of the detrended SST dataset compared to the original one (not shown).". What kind of consistent statistical behavior do the writers mean here? Not clear.

Line 331: "...EMSs are formed based on a "background: SST substructure field, largely depending on the climatological ranked daily SST variability in the MS". I thought the SST substructures where represented through the RDA and not the climatological ranking of the daily SST. This sentence is confusing.

Line 334: "*Analyzing the EMS SST substructures in the previous section revealed that EMSs in the basin are commonly formed due to the warmer summer days being warmer than normal*". This is not true always, since the example case of the Central Med. Sea (Fig.4c) shows that the EMS develop mostly due to the..coldest rank of days becoming warmer.

Line 336: "*To complement these findings, this section investigates the role of MHW events during EMSs.*". Why does the role of MHWs acts as complementary to the EMS? Are the writers sure that there are EMSs without a MHW occurring?

Line 336-337: Re-phrase the sentences.

Line 337-338: "*Then, we examine the relative role of MHWs in EMSs by means of changes in MHW properties during EMSs (Sect. 3.2.2) with respect to mean MHW conditions*". Do the writers mean the non-EMS conditions when referring to "mean MHW conditions"?. Not clear.

Lines 359 – 371: These paragraphs belongs to a discussion section and not to the Results.

Lines 269-371: "*This difference is mostly related to our choice to focus on summer MHWs that begin and decay within the JAS summer period. Taking into account all events throughout the year, similar MHW duration values are reproduced (not shown)*".Do the writers mean that duration of the MHWs would change if the examined period would be between June – August or any other season of the year? What are the similarities of the MHW duration when taking into account all the events of the year? Similarities with the MHW duration of the current study or with previous studies?Not clear.

Line 374-376: "*Consequently, they represent the mean seasonal conditions of i) all summer MHWs in the study period and ii) the MHWs occurring in EMS, respectively, both not expected to be indicative of the most extreme MHW conditions.*".What do the writers mean here by "*both not expected to be indicative of..*"? The mean seasonal conditions are not expected to be indicative of MHW conditions. However the EMS are summers with

extreme warm conditions. So why are they not expected to be indicative of MHW conditions?

Lines 377-378: "*MHWs may promote the formation of an EMS either by being more frequent, of greater duration or of greater intensity than usual, thus, contributing in different ways –not necessarily bearing extreme characteristics– with positive SST anomalies to the seasonal SST being the focus of this study*". I am not sure the difference between the EMS and the MHWs is properly explained so far in order to understand why the MHWs *may promote EMS formation* instead of being part of/always present when an EMS is happening or how the EMS conditions are differentiated from MHW conditions, or why an EMS cannot promote a MHW instead.

Section 3.2.2

Line 380: "*To understand the role of MHWs in EMSs, this section presents the mean EMS anomalies of MHW properties with respect to their mean summer values.*". The way this sentence is written is confusing for the reader. Does that mean that this section presents the EMS anomalies of MHW properties relative to the mean EMS summer values or relative to the mean summer MHW values?

Figure 6: Does the computation of the mean MHW properties during 1950-2020 include the EMS years in the average or not? Because if the EMS years are included in this average what is the point in creating the anomalies of MHW properties only during EMS years (Figure 6d,e,f)? Clarify.

Line 385: "*In terms of mean basin values, mean summer MHW intensity equals 1.56° C while when isolating events during EMSs, it drops to 1.45° C (Table 1). This suggests either the suppression of MHWs in EMSs or the existence of alternative mechanisms potentially enhancing the MHW conditions in EMSs.*". Table 1 shows a basin-mean Imean of 1.41 ° C instead of 1.45° C. I am not sure I understand how the reduction of Imean during EMS-MHW conditions can mean either a suppresion or enhancement of MHW conditions. What do the writers mean here?What kind of mechanism are they referring to and how this differentiates when there is MHW suppresion from MHW enhancement?

Lines 400-401: "*MHW analysis of this part further suggests that enhanced MHW conditions commonly contribute to the observed SST anomalies in EMSs*".. I am not sure that the comparisons of Fig.6 shows this result (see comment of this Section in Major comments). But even if this is true, I do not understand why the MHWs contribute to the EMS conditions instead of being the reason why an EMS is formed in the first place?Also in Lines 407-408 the writers mention:"This suggests that in this location, the higher rank days which are the most responsible for the EMS formation (Fig. 3b), face more difficulty in producing distinct MHWs". Does the EMS create MHWs or the other way around in the end? These two sentences are a bit contradictory.

Lines 399-415: These two paragraphs are discussing the same thing and should be merged in one. Either more numbers of the indicative duration and intensities should be given here or move this sentences as part of the discussion.

Section 3.2.3:

Although this is a Results section the description of MHW properties gives no numbers. Rather it is limited in comparisons of "greater" or "smaller" magnitude. Either move these paragraphs to the discussion section offering possible mechanisms behind the results described or simply refer to the differences seen in MHW properties with the de-trended datasets using numbers.

Lines 424: Apart from the multi-decadal signal being removed is the trend value of the summer also removed here like in lines 170 – 173? Not clear.This information was never mentioned again in the manuscript after the beginning of the paper.

Lines 426 – 427:"*This suggests the presence of similar event-driving mechanisms and reveals the positive contribution of the multi-decadal signal to the observed sea surface warming, this time specifically via the SST anomalies caused by MHWs.*". How exactly has the positive contribution of the multi-decadal signal to the observed SST warming and

(especially through the MHWs) been revealed here if the mechanisms have not changed at all in this and the original dataset? I do not understand. Clarify

Lines 432-437: These are arguments that belong to a discussion section and not to the results as they are commenting on literature and possible explanations.

Line 432-433: "During EMSs, MHW intensification is greater when using the original dataset. This is more pronounced in the eastern basin and particularly in the Aegean Sea". Apart from Table 1, Fig. 6 should be cited here as well.

Table 1: The categorization in "Orig" and "EMS" is not a good choice of names. Better to find some other names to distinguish between the two datasets used here.

Line 439: "*Moreover, we observe a similar behavior of MHW properties in the detrended compared to the original dataset*" I thought the writers were already talking about the comparison between the detrended and original dataset. So why repeat the argument here?

Lines 440: "'*…longer lasting events in EMSs (compared to the mean state) counterbalance the lower event intensity values in EMS (Table 1).*"This sentence is really unclear. I am not sure where to look in Table 1. Are you comparing the MHW days between the mean summer MHWs and the EMSs MHWs on the Orig dataset or the days between mean summer MHWs and the EMSs MHWs or the duration of events between the Orig and EMD datasets? Clarify. Better to rename the datasets of Table 1 and everytime you compare between them, refer to what datasets you compare. And also, how do you know that longer lasting events are "counterbalancing" the lower intensity? What do you mean by this? Not clear.

Lines 442-446: This paragraph refers to which Figure exactly? Not really sure where to look for these results.Is it if Figure 6? Very confusing.

Line 442:" *…in absence of long term trends*….". Why do you need to refer to the absent trends and what do you mean by that? That there are some trends that you have not shown or that there are no trends in the particular areas you are referring to? Not clear.

Lines 444 – 445: "*Among these regions, the Aegean Sea stands out as the one presenting a pronounced MHW contribution to EMSs both in the original and the detrended data…*". Again where am I supposed to look for these results? Is it figure 6 or not? Figure 6 does not include comparison between the detrended and the original dataset, no? These statemens here look like they come from nowhere since there is no reference to any figure inside the paper or any other result.

Lines 446-447: "*…Given that, the pronounced role of MHWs in EMSs in the eastern basin can be only partly attributed to the sea surface warming trend in the MS during the past few decades.*". How do the writers know that the pronounced role of MHWs in the MES in the eastern basin is only "partly" attributed to SST warming? Where do they show that it is "only partly" attributed? Have they compared any other variable such as trends of the MLD, winds?(e.g. trends of lower winds in the area or trends of shallower than normal MLD). I find these arguments a bit unfounded since there are no results to look at and compare at this point.

Section 3.3

The entire section here describes results on air-sea heat fluxes without giving any indicative numbers on the increased or decreased contributions presented in Fig.7. The description here is something between a Results section and a Discussion. The entire section should be re-written where some indicative numbers of the spatial distribution of the air-sea heat fluxes should be given, separately as Results and then in the Discussion section there should be a comparison with current knowledge and literature.

Lines 449-458: Is this analysis only for the original dataset and the detrended one is only shortly discussed in Section 3.3.5?Clarify in this paragraph.

Lines 465 – 467:" *Moreover, positive Qnet anomalies appear mainly in areas where enhanced MHW conditions in EMSs were detected (northern Mediterranean regions and particularly the Aegean and the Adriatic Seas; Fig. 7A vs Figs 4d-f)*". I am not sure what is

the point behind the connection between Fig.7a maps with the results from Fig.4d-f that refer to points. After all, the points that were (randomly?) selected from 3 regions (Fig.4) cannot represent the EMS profile of the entire area (from where they came for). In addition, we do not know if the EMSs of other points, e,g in the south-central Med. Sea exhibit the same behaviour as those in Fig.4,but are in areas characterised by negative Qnet. So I am not sure what is the argument behind this sentence. Is a cause-effect result implied here?

Lines 468-469: "This similarity is at least partly attributable to the widely explored driving role of air-sea heat fluxes in the formation of MHWs". How can a similarity between a map and 4 points (?) be attributed to "a widely explored role or air-sea heat fluxes"? What exactly is the similarity here and how is it related to the fact that we know a lot about air-sea heat fluxes? I do not understand the argument here. Clarify. Also, the studies cited here are global studies examining drivers of MHWs at a local scale. Better to cite studies that refer to the role of air-sea heat fluxes in Mediterranean MHWs and not on global events.

Line 475: Either 'above-mentioned' or 'aforementioned'.Replace.

Lines 485 – 487: This is a far-fetched claim to be made. Although there seems, indeed,to be a slight reduction of MLD across the Turkish coasts, I am not convinced that the reduced upwelling can be inferred solely from an MLD indicator. Especially because the MLD refers to extreme EMS summer which, at the moment, I am not quite sure which period exactly they refer to and if this period coincides with the exact period of the upwelling in the area. Also, the anomalies of the MLD apparently are computed with respect to a different baseline period, which cannot perhaps be directly compared with the anomalies of the rest of the variables of Fig.7, let alone to be linked with the impacts of EMS. In addition, is this reduction of MLD significant enough to suggest there is going to be reduced upwelling? Either compute trends of the upwelling index in the area or the authors should entirely remove this claim from the paper as it is not substantiated.

Lines 487 – 488; "SH flux (Fig. 7d) and net SWR anomalies (Fig. 7c) in the eastern Aegean and the central Levantine Seas are negative but of a much smaller magnitude.". How much is the magnitude here? Quantify this information wherever possible.

Figure 7. In the text shortwave radiation is referred to as SWR, whereas in Figure.7 as SW. Be consistent on the naming of variables everywhere.

Lines 502 – 506: How can the Qnet anomalies in the Adriatic Sea be "*almost fully determined*" by shortwave radiation (line 503) when there is also a reduction of LH losses according to Fig.7b? I think the LH also plays a significant role here and it is not just the SWR.

Line 504-505: "*The increased LH loss in EMSs here is exceeded by the gain in net SWR, leading to the observed net heat gain*". This sentence does not make sense at all. I do not understand what the writers want to say here. Also, I do not see anywhere an increase or LH losses. I rather observe a reduction of LH losses.

Lines 514 – 515: "*MLD and wind speed anomalies tend to present the same sign, especially where large anomalies appear*.". Although this is true for some regions of the basin, I can see also other areas (e.g. southeast Levantine) where the windspeed is lower than normal and MLD is higher than normal. Again, at the moment there cannot be a direct comparison between MLD and any other variable of Fig.7 given the fact that the anomalous fields are computed based on different reference periods.

Lines 509 – 520. This paragraph again is a mix between Results (without any indicative numbers) and a Discussion which does not reveal anything novel. The writers try to claim a relationship between wind and MLD (which is already well-known), albeit on anomalous fields that are computed based on different periods. Again, this paragraph should be re-written separately for the Results section, where indicative numbers are given and the comments should be addressed in the Discussion section.

Lines 531-532: "*These fields suggest that LWR and SWR (standing for net values from now on) in EMSs work complementarily in the basin.*" What do you mean by complementarily? I do not understand.

Lines 536: The Figures of the supplementary material better be named as e.g. Fig.**S**3a instead of Fig.**B**3a. What does "**B**" stand for? I have never seen this in the literature so far. Also from Fig.B3a, I can see positive anomalies of cloud coverage in the central Aegean but also negative cloud coverage to the southern-southeast Levantine, where the LWR anomalies are close to zero. So the statement "Indeed, positive EMS anomalies of total cloud cover derived from ERA5 are found in this area (Fig. B3a) leading to increased downwards LWR and the observed reduced net LWR loss from the sea surfaec (Fig.7e)". Is not entirely true for the Levantine basin.

Line 538-539: "Particularly in parts of the Adriatic and the northern Ionian Seas, the SWR surplus is responsible for the net heat gain observed in EMSs (Fig. 7a vs Figs 7b-e)". Since there is Section 3.3.3, where the contribution of each air-sea heat flux is quantified, better to combine these two section together. There is no point in describing qualitatively the air-sea heat contributions here, without giving any quantifiable information and then create a new section later (3.3.3) where the contributions are actually computed. Also, by looking the maps of Fig.7 I cannot deduce that it is only SWR surplus that is responsible for the net heat gain of the EMSs in the Adriatic and the Ionian Sea. I would argue that reduced LH losses contribute as well.

Line 541: "*Results suggest that Qnet anomalies (either positive or negative) in EMSs are primarily formed by LH fluxes in most areas, followed by SWR in the western and central basin and particularly in the Adriatic Sea.*". This contradicts the sentence just above, where SWR is responsible for the net heat gain!

Section 3.3.2:
As before, this section appears more as a Discussion section rather than a results section as no quantification is provided. This description (if needed) needs to be transferred to a discussion section.

Line 550-552:"*Differences in the behavior of surface heat fluxes are expected to exist among the locally detected EMSs and the mean EMS state presented in the previous Section. To shed light on this, Fig. 8 illustrates how common is a positive anomaly for each of the examined heat flux components during EMSs*". I do not see how the frequency of positive anomalies occurrence at each grid point, constitutes "*a mean EMS state*" while the analysis performed in the previous sections is on "*locally detected EMSs*". Since both analysis yield maps that means that both fields refer to locally-detected properties. I suggest to change the title of the entire section here.

Figure 8. I suggest changing the color pallete in this figure since red usually means higher and purple/blues lower. The red colour draws the attention a lot to the lower values of the frequency of occurrence.

Lines 556-557: "*The areas where positive Q net anomalies in EMSs are more frequently met are the same as where the largest positive mean EMS Q net anomalies were found (Fig. 7a vs Fig. 8a).*". I am not sure this is true. For example in Fig.7a, I can see higher Qnet anomalies in the Adriatic and the northeast Aegean, which do not exactly coincide with the highest frequency of Qnet anomalies of Fig.8a. For a better comparison between these two maps, it is better to use a pattern correlation coefficient instead of just a visual comparison. What is the point of showing both figures 7 & 8? I do not see any differences in the main message behind these two sections and the one coming after that (Section 3.3.3).

Lines 566-567: "*In the same locations, SWR radiation appears increased in most EMSs, reinforcing the mean EMS findings.*". What kind of EMS findings are reinforced here exactly?

Lines 569 – 571: "*Negative Qnet anomalies appear in every EMS in the Ligurian, the south-central as well as in a few specific spots in the WMED Sea, despite the*

systematically increased SWR. In these areas wind appears always reduced in EMSs (although not followed by decreased LH loss), while suppression of vertical mixing is observed only in half (or less) of the EMSs (Fig. 8c)". I am not sure where is the figure that shows negative Qnet anomalies as Fig.8 demonstrates maps on positive Qnet anomalies only. So I am not sure why Fig.8c is cited here.Confusing, Clarify.

Lines 578 – 580: "...*we expect that the responsible cool summer days (being warmer than usual) are mostly the early summer ones*.". The fact that the early summer days are...the cooler summer days with respect to the days during the heart of the summer .. is obvious. I do not understand the point behind this sentence.

Lines 580 – 583: "*In all EMSs in this area, the early summer (July) SST anomalies are the largest within the season. The increased mixed layer heat content in June is reflected in the positive SST anomalies of June while, also in May, SST in this area was found marginally larger than climatologically.*" Which figures shows all this? Confusing.

Lines 584: "*The temporal coverage of the modelled temperature and MLD used in the OHC calculation inevitably limits this task in using...*". Which task do you refer to? Not clear.

Lines 586 – 587: "*For this reason, a complementary task included the use of the detrended dataset only for locations that experience EMSs within the available period..*"? Which one exactly is the available period, not clear.

Lines 589 – 590:" *In the region discussed above (south-central basin), increased upper ocean heat content and spring SST values are observed before each EMS also in the detrended dataset, i.e., independently of the warming trend*". Isn't this something that one should expect given that as the summer approaches, the heat content of the ocean is ..bound to increase more and more? I am not sure why the writers expected that the de-trended dataset should show "no preconditioning" or no increase of the heat content of the ocean as we move to the summer. Also how much warming above average is the preconditioning considered as "preconditioning" and not just the normal increase of the period before the EMS? Not clear this.

Lines 594 – 596. "*Usefulness of this index is highlighted when examining single summers (e.g., the exemplary case of EMS 2015 in Sect. 3.3.4). Even if it is positive in most cases, its contribution is often enhanced when and where there is no causal link between the surface net heat balance and the observed surface warming thus revealing it constitutes an actually contributing EMS formation factor*". Does the "surface net heat balance" refer to Qnet? Keep consistent names throughout the paper. Generally it is not a very good practice to refer to a section and describe its results before the section itself is presented in the paper. It is disrupting for the reader. In addition, I have not seen anywhere so far in the paper **any causal link** between the surface Qnet (surface net heat balance?) and the SST warming. The only thing the authors have presented are visual comparisons between maps of physical variables, which are not supported by any good "statistical index" denoting the degree of similarity in their spatial distribution. So I would be careful not to use the words "causal" link, unless specific analysis has been done for specific events that prove the causal link.

Line 597: Apart from the fact that the entire paragraph does not belong to the Results section, as it mostly comments on the results, the writers suddenly go from Fig. 8d at the start of the paragraph, to referring Section 3.3.4 (that has not yet been presented) and then suddenly mention Fig.9d (also not presented yet) without having present figures Fig9.a-c before that. In general the Figure references should be more structured in the text and should come with an order.

Lines 598-599:"*Although horizontal advection is not investigated in this study, this finding potentially suggests that EMS preconditioning hardly takes place in areas of intensified circulation.*". Which areas do you refer to exactly?Clarify.

Section 3.3.3

I am not sure how the results from this section differ from those in Sections 3.3.1 – 3.3.2. I understand the writers used an extra way to quantify this time the contributions from the different air-sea heat fluxes (apart from the quantification of the contribution that was necessary to begin with). But what is the added value of another yet way to see the contributions of the different air-sea heat fluxes to the EMS? I do not understand, qualitatively, what is the new knowledge we gained in Section 3.3.3 compared to Section 3.3.2. and 3.3.1. For example, in line 608, where the contributions are discussed, the Fig.7g is also cited to support the hypothesis that was also evident from the quantified contributions themselves. Probably this quantification metric is more meaningful than the descriptions from Section 3.3.1 and 3.3.2. However, at the moment the metric is not very clear in its explanation in the Appendix. It is quite difficult to understand how the writers constructed it, e.g. what is the difference between SSTA cum >0 (which is stated as SST anomaly in the Appendix), with the $\Delta SST'obs$?Is the second data anomalies from observations only? It is very confusing.

In addition, before the writers start describing the P_{EMS} of Fig.9, they should give a qualitative definition at the start of the section to the readers as to what each positive or negative P_{EMS} means and then continue with the description of their spatial destribution. At the moment the P_{EMS} meaning is scattered here and there when the spatial distributions are analysed.

Line 606: "*In positive P EMS areas with 0 < P EMS <100%, other mechanisms are expected to work complementarily (i.e., towards greater SST anomalies)*". What does complementarily mean?What does it mean towards greater SST anomalies?

Line 612: "*..hence oceanic processes definitely drive EMSs in these areas*". I would avoid statements with "definitely" since the contributions from oceanic processes have not been actually proven here. They have just been inferred.

Figure 9: I am not sure what panels b), c) e) f) are showing exactly. What is the "warming phase" and "cooling phase" that are being subtracted here? How are they represented mathematically? Also I do not understand what is the use of an index where the warming and cooling phase are being subtracted. Clarify. In addition, in the citation of the figure, the writers should clarify the qualitative meaning of the P_{EMS} metric and not just refer to it like : " P_{EMS} *metric values (%) for the contribution....*". The citation is supposed to make the reader understand what are they looking at. At the moment I only understand that I am seeing some percentages of contribution.

Lines 627, 639: 1) "*These values indicate that oceanic processes are primarily responsible for the observed EMS SSTs in these areas*", "*Hence high SST anomalies leading to EMSs in these areas are formed despite the thermal energy deficit at the sea surface during the same periods.*"These sentences are confusing. You are saying that there are EMSs happening but..there is thermal energy deficit. How is this possible? Clarify.

Lines 631 – 633: "T*he present section complements this finding by revealing that a thermal energy deficit at the sea surface is consistently observed in these areas even when focusing solely on highly effective (for the EMS formation) warming phases*". Rephrase.

Line 650: "*Results also reveal a strong link between MHW properties and surface heat fluxes (Fig. 9a vs Fig. 6d)*". I think this is a proven statement, multiple times already??? Why do you need to prove it again?

Lines 654- 655: "*The spatial similarity of MHW intensity and P_{EMS} over the basin suggests that the crucial role of surface heat fluxes found for the northern MS is associated with their ability to drive high SST anomalies (thus intense MHWs)*". I still cannot understand the difference between EMS and MHWs in terms of high SST anomalies. They are both referring to high SST anomalies in general.

Section 3.3.4

The example of the Marine summer 2015 kind of suggests that there is not much of a difference between EMSs and MHWs. More specifically, Lines 569 – 661, are justifying the exploration of the EMS 2015, based on a citation about a MHW in the summer of 2015!. The writers need to either clarify what is the difference of EMS from the MHWs and thus, their usefulness compared to the MHWs or re-think the idea of EMS in total. In principle, these two definitions demonstrate two different ways to capture extreme warm temperatures in the ocean. Either, use one of them without comparison with the other or re-think the idea of the EMS.

Lines 679-681: "*On the other hand, the preconditioning index is clearly more pronounced here than in the positive part (Fig. 10d). Indeed, early summer days (July 2015) present...*". I thought the early summer days of the preconditioning phase was the month of June according to Section 2.2.5

Lines 685-686: "*Following 2015, other widely experienced EMSs such as 2003, 2012, 2018 (in descending order by means of spatial extent) are examined (not shown).*". These years also constitute MHW years in the Med. Sea. So I cannot see the usefulness/added value of the EMS definition in identifying extreme ocean periods when compared with the MHW definition.

Section 3.3.5

If the writers re-think the idea of the EMS and justify its usefulness as an index then the analysis of the contribution of air-sea heat fluxes in the detrended datasets deserves to be shown as a comparison to the equivalent results with the original dataset and its figures should be juxtaposed and commented in the main text.

Conclusions and Summary:

Lines 794- 796: "*Nevertheless, the multi-decadal signal contributes to the observed surface warming through more intense and slightly longer lasting MHWs, without affecting the MHW intensity spatial distribution, thus most probably also the event-driving mechanisms*". I do not understand how the multi-decadal signal contributes to more intense events without affecting the intensity spatial distribution and the event driving mechanisms. How do you infer that the event-driving mechanism is not affected? Clarify.