The section addresses a relevant and very pertinent topic: the study of marine heatwaves in the Baltic Sea is extremely necessary. However, the section still requires more organization. The writing is often unclear. Important decisions were made for this section, although they still require support with the inclusion of more incisive justifications in the text. The section requires MINOR REVISIONS. Its current version cannot be approved for publication. To improve the manuscript, some questions are raised in the comments below, which can guide the authors.

We thank the reviewer for their thoughtful and thorough review and helpful suggestions. We addressed the reviewer comments below in blue.

General comments

1) Why did the validation use only moorings (LT Kiel and Northern Baltic)? Are the analyses at two close stations (diamond and circle symbols in Fig. 2) representative for the entire Baltic Sea?
   In general, the model is validated with data from various stations distributed across the Baltic Sea. For the revised version of the section we will provide a more comprehensive model validation using k-means statistics, but move this to the supplementary material in agreement with the editor. Furthermore we will provide a map with all stations and place names mentioned in the text.

   In our Figure 2, actually three stations were marked - admittedly not too obviously. Two of those stations, 'Lighthouse Kiel' in the western Baltic Sea and 'Northern Baltic' in the north-east corner of the Baltic Proper, are used to validate the MHW parameters computed from the model (see Figure 4), since these two stations have relatively long records at the surface. Usually, the underlying climatology to assess MHW occurrences is computed from about 30 years worth of data or as close as possible to that.
   Please note that L82 currently states that the data at Northern Baltic is available down to depths of 103.8 m, which in fact is not the case. There is only surface data available. This will be corrected in the revised manuscript. The third station 'BMPH2' was therefore originally used to show the model's temperature skill at a deep station at 150m depth and close to Northern Baltic (see Figure 1c). While this monitoring station has data in greater depths, the availability is irregular in time and depth. So it is not possible to calculate a MHW climatology with this dataset.
   In the revised manuscript, we plan to drop the additional validation at BMPH2, as we do not actually use this station in the manuscript and the newly included k-means statistics will provide sufficient insight into the quality of the model over the entire Baltic Sea.

2) SST anomalies are computed with respect to a climatological SST considering the period 1997-2021, whereas MHWs are detected concerning the climatological data from 1993 to 2021. Why the choice of a dataset restricted to 1997-2021 for satellite SST? I suggest that the authors take a look at the link: https://data.marine.copernicus.eu/product/SST_BAL_SST_L4_REP_OBSERVATIONS_010_016/description

   Here, we wanted to use a dataset that is independent of the model data, i.e., our BSH operational routine, which - as mentioned in Sect. 2.1- is available after 1997. The

Copernicus Marine Baltic Sea SST reprocessed L3S product, which is what the L4 product is based on, is assimilated into the Baltic Sea reanalysis model and would therefore not represent an independent dataset.

3) In section 2.4 “Model validation”, the author should describe the metrics and formulas applied to perform the analyses/study, and move Fig.1 and its description to a topic in section 3.
We will include Figs 1a and b as a supplement and drop BMPH2 (Fig 1c). As mentioned (see (1)), we will instead include a k-means model validation. The corresponding metrics and formulas will be included wherever appropriate.

4) OSR8 guideline suggests a maximum of 5 figures. Confirm it!
We will move the current Fig. 1 to the supplement. If permission is granted by the editor, we will instead add a map of the regions of the Baltic Sea as a new Fig. 1, which would still leave us with a maximum of 5 figures.

5) Improve the writing in general. For instance, Section 3.2.2 shows interesting results that can increase the impact of the section, but are not described carefully and clearly. Review it!
We will discuss Fig. 5 in more detail, especially regarding the subsurface MHWs and their potential drivers. We will also aim to improve the wording in the entire document.

6) English writing requires proofreading.
The revised manuscript will be proofread by an English native speaker.

Specific comments

(line 28) confirm the average period 1850-1900. Confirmed, the WMO State of the Global Climate 2022 provides this global mean anomaly relative to the 1850-1900 preindustrial reference period.

(line 29) fifth or sixth? Replace with “among the 6 warmest years” We will rephrase this accordingly in the revised manuscript.

(line 35) 1.35 K / 0.54ºC; use same unit We will adjust to °C throughout the manuscript

(lines 51-70) Reduce the details to explain the satellite observations. Put their more relevant characteristics, then cite their main references in the text. - We will compress the section explaining the satellite observations, concentrating on the most relevant details.

(line 72) Remove “which are” will be removed

(line 75) replace the first “and” with “to”. will be replaced

(line 96) replace “chapter” with “section” will be replaced

(lines 99-107) This part describes results that should not be present in the section 2 “Data and Methods” As mentioned, Sect. 2.4 will be extensively modified

(lines 112-114) These lines describe results again. Organize the text and move them to the appropriate place. See above

(line 116) The acronym MHWs is already mentioned in line 25. will be adapted accordingly

(lines 118-120) If the tools produce identical results, it is not relevant to mention the use of python and Matlab. Mention only "We use open-source tools to detect MHW (Oliver, 2016; Zhao and Marin, 2019)" Will be changed

(line 126) "the climatological data of 1993 to 2022", replace "of" with "from". Figure 5 says 1993-2021.
The paragraph will be slightly rephrased. The climatological period will be homogenized to cover the years 1993 to 2021 for all MHW analyses based on the model data (Sect. 3.2, Sect. 3.2.2, Figs. 3 and 5).

(line 131) This line is not clear.
This paragraph has been reworked for clarification:

> *"Then, in order to evaluate the development of those MHW metrics over time, block averages (using a block length of one year) for each MHW metric are computed for both the observations (product ref. no 2 in Table 1) and the model data (product ref. no 3 in Table 1) at two stations: Lighthouse Kiel and Northern Baltic. The yearly MHW metrics from observations and the model are compared and the linear trends (95% significance) are calculated for each of those annual MHW metrics. Finally, the correlation of the annual MHW metrics to the annual mean temperature based on model data was assessed using a linear least-squares regression and a two-sided t-test for significance."*

(lines 138-140) "Bothnian Sea", "Bothnian Bay", "Baltic Proper", "Gulf of Finland" and "Gulf of Riga". Find a way to specify/highlight these regions in one of the maps in Fig. 2.
We will add an additional with a map of the regions of the Baltic Sea

(line 144) describe the metric "SST anomaly rank" in section 2.4 "Model validation"
The anomaly ranks shown in Fig. 2 are a very simple way to provide climatological context, providing information about how extreme an anomaly of a given magnitude is. For every grid point the anomalies of all years for a given month - e.g. August (calculated against the climatological mean of the respective month) - are sorted according to their magnitude. The anomaly rank depicts simply the rank in this sorting. For example, a plotted rank 1 of August 2022 in Fig. 2 therefore means that 2022 featured the hottest August in the entire dataset for the respective grid point. In Fig. 2 we show color shadings for the hottest eight and coldest eight ranks, respectively. We will provide a more detailed explanation in the revised manuscript.

(line 146) It would be better to avoid the mention of "obvious" here; replace it. Will be replaced

(lines 158-170) Specify the lat/lon points that delimiter the regions in discussion. The reading is difficult now. The regions will be depicted on an additional map of the Baltic Sea. Furthermore, the paragraph will be adjusted to improve the readability. We will also add a table (to the supplement?), which will contain the values of all regional MHW statistics. Furthermore, the climatology will be updated to cover only the years 1993 to 2021. Hence, the values of the statistics will be updated accordingly in the revised manuscript.

(lines 168-169) Would the Gulf of Bothnia include both the Bothnian Sea and Bothnian bay? Looking at Fig. 2, the highest SST anomalies are present in June and July in these regions. Why does the highest cumulative intensity derive from the temperature anomaly in November?
Yes, the Gulf of Bothnia includes the Bothnian Sea and the Bothnian Bay. Unfortunately, the statement in line 168-169 was incorrect. The high MHW metric values in the western Gulf of Bothnia stem mostly from a short, but very intense MHW that started at the end of June and ended at the beginning of July. This will be corrected in the revised manuscript.
The figures and metrics in Fig 2 and Fig 3 are not 100% comparable. The SST anomalies and the ranks are based on monthly mean values of weekly produced satellite-derived SST data within the data period 1997-2022. We use this independent dataset to complement the model and station data and to highlight the seasonal development of temperatures in the entire Baltic Sea.
The MHW analysis presented in Figure 3 is based on daily modeled SST data from the MYP covering the data period 1993-2022.

(lines 169-170) Sentence is also not clear. see above

(lines 176-177) replace “after 2020” with “in the evaluated period”. will be replaced

(line 181) Replace “rising mean temperatures”, my suggestion is “warming temperature trend”? Replaced with:

> *“The maximum (Fig. 4c) and cumulative intensities (Fig. 4e) of observed MHWs do not show a clear trend and are not correlated to the warming annual mean temperatures (Fig. 4d and Fig. 4f).”*

(lines 190-191) suggestion: replace “2018 was exceptional” with “In terms of total MHW days, the highest number is viewed in 2018” will be rephrased

(lines 203-204) “has a significant lower temperature of -0.3°C to 4.5°C as the climatological mean”, this sentence is not clear, since Fig. 3b shows only 2022. Improve the writing Fig. 5b (which is what is referenced here) in fact shows the deviation between 2022 and the climatological mean. This will be clarified in the figure caption. We will also modify the sentence to make it clearer: “has a significantly lower temperature than the climatological mean (up to -7 °C deviation; Fig. 5b)”

(lines 208-209) “A few weeks prior to this MHW”, is the author referring to the extreme or severe MHW? Specify the time period. Improve the writing Will be clarified and the paragraph extended to provide a more in-depth analysis of the (subsurface) MHWs.

(line 385) replace “chapter” with “section”. will be replaced

(line 402) subplots 5g and 5h are not present will be removed from the caption

(Figure 3) change the color pallet The color pallet will be updated to be consistent with Fig. 2

(Figure 5) line 402 states “The period used for the climatology is 1993-2021”, while line 126 says 1993-2022. In the end, what climatological period is used? The climatological period will be homogenized to 1993-2021 for all MHW analyses.

Improve the description of the caption in Fig. 5. will be done