

## Authors' Comments to Reviewer 1

*Reviewers comments formatted in italic dark red font.*

Our response is formatted in thin black font.

*Summary: The manuscript presents an analysis of a data set from the Baltic Sea reanalysis, focusing on temperature, salinity, ocean heat content and freshwater content. This data set covers the period 1993-2023. The manuscript investigates long-term trends and the connection between these variables. The conclusions of the analysis are not totally surprising. They identify a warming trend and a strong connection between water temperature and ocean heat content and between salinity and freshwater content.*

*Recommendation: I have concerns about the manuscript's framing, title, and objectives. I do not see problems with the data analysis. My recommendation is that the manuscript requires substantial revisions.*

*Main points:*

*1) The title is misleading. I could not identify any 'new conceptual framework framework' nor find a definition for 'state of the Baltic Sea'. The manuscript is perhaps a valuable presentation of the ocean reanalysis, but it does not present any 'framework#' and leaves many concepts in the title and in the introduction undefined. I think this manuscript is indeed a presentation of this data set, which is fine, but it tries to present it as a more substantial advance than it really is.*

*If the authors believe the manuscript presents a new conceptual framework, they should explain it clearly in the introduction. I failed to see it.*

We will explain the framework in more detail in the Introduction. We have changed the title by adding the term “physical”, which means that we are dealing with the physical state, not the biogeochemical or ecological state. On lines 51-52 we say explicitly that “We propose the following conceptual model, which merges the analysis of temperature and salinity with their integral counterparts OHC and FWC”

*2) The introduction often presents the manuscript in a too-bright light. For instance, the text states (line 45) that it presents a 'model'. I cannot see any model. Again, the manuscript analyzes the connections between different ocean and atmospheric variables, but it does not contain any model that would allow predictions or that includes any physical mechanisms, equations, etc.*

We agree that we have not introduced any model. Therefore we have changed the term “model” to the term “framework”.

*3) I have problems understanding the 'state of the Baltic Sea'. First, it appears that the manuscript deals only with physical variables and leaves geochemical or biological variables out of the analysis. The use of*

*'state of the Baltic Sea' on this account alone seems exaggerated. But more importantly, what does 'state' mean here? Is it a snapshot of the ocean at a particular time? Does it mean a more value-centered assessment of the situation in the Baltic Sea (good, bad, etc.)? Readers curious about the title may be vastly disappointed when reading the manuscript.*

We mean physical state and have changed the title accordingly. On lines 45-56, in a more detailed description of the framework we stated what we mean with the “term” physical state.

*4) I do not see the need for a Random Forest method. The manuscript applies this algorithm to conclude that the driver for the heat content is the temperature at all layers and that the driver for freshwater content is salinity at all layers. Do we need an RF algorithm? Any reader would be stunned if the results would have been different. Ocean heat content can be directly calculated from temperature, and freshwater can be directly calculated from salinity. I do not see the need or the advantage of using a machine-learning algorithm to identify those connections. They are obvious.*

A Random Forest method captures complex, non-linear relationships between variables. In this study we use four different RF models: RF1 for Ocean Heat Content (OHC) with meteorological variables, RF2 for Freshwater Content (FWC) with meteorological variables, RF3 for OHC with temperature profiles  $T(z)$ , and RF4 for FWC with salinity profiles  $S(z)$ .

We completely agree that ocean heat content is determined by spatially averaged temperature profile and freshwater content by its spatially averaged salinity profile. Using the RF model in the current context is to understand at which depth layers temperature contribute the most significantly to the overall heat content and salinity to the freshwater content.

*5) On the other hand, the research question is unclear. What knowledge gap would the manuscript fill? What is unknown in the variability of the Baltic Sea that this data set may help to clarify? The introduction is silent about this.*

We will elaborate the research question further in the Introduction.

*Summarizing these previous points, it seems to me that the manicurist tries to push up a correct and useful analysis of the ocean data by using exaggerated terms (which often do not have a clear meaning)*

*Particular comments:*

*6) line 33 exceptional increase in global sea surface temperature*

*Exceptional in which sense? At which time scale? Th Earth's temperatures have been warmer than now in the geological past.*

We mean exceptional warming in recent history. This is the estimation for the period 1979-2024 (McGrath et al., 2024). We will clarify it in the revised manuscript.

*7) Figure 1 : Conceptual Scheme of the Baltic Sea State parameters.*

*Again, what is the conceptual scheme shown in this figure? This figure shows the obvious links between those physical variables and only physical variables.*

Yes, these are obvious links, but rarely if at all combined into a single framework for physical assessment of the water basin either closed one or with open boundaries.

*8) Some methodological aspects are not clearly explained. For instance, I struggled to find the time scales of analysis. I think it is only mentioned in one figure caption or line 171, which alludes to annual means. I kept wondering for a long time if the authors were analyzing daily means, monthly means, seasonal means or annual means*

All analyses are prepared in a common timescale covering period from 1993 to 2023. We will improve the Figure 4 and its caption, where the annual timeseries are shown. That is an essential part of the proposed framework. Everything is clarified in the revised manuscript.

*9) Likewise, it is not all clear whether these variables are considered at the grid-cell scale, water column scale or averaged over the Baltic Sea*

This is explained in the manuscript. In the conceptual framework we involve spatially averaged values, either it is averaged over the whole Baltic Sea or at each vertical level. We will specify it more clearly in the revised manuscript.

*10) Often, physical units are missing, for instance, when stating trends. Trends must have units of variable per unit of time. It is not clear if the trends refer to changes per year, for instance, or over the whole period*

These trends are calculated for z-scored values. Therefore they do not have units for physical variables. All trends have been calculated per year. Will be specified in the revised manuscript.

*11) "Surface net solar radiation has a weaker but still significant positive trend of  $0.058 \pm 0.035$ , and the evaporation time series shows a negative trend of  $-0.041 \pm 0.039$ "*

*Units missing*

These trends are calculated for z-scored values. Therefore they do not have units for physical variables. All trends have been calculated per year. Will be specified in the revised manuscript.