Dear Reviewer #2, thank you for reading and suggesting modifications to our manuscript entitled "Advances in Monitoring Black Sea Dynamics: A New Multidecadal High-Resolution Reanalysis".

We believe that your review has helped to substantially improve the revised manuscript. The changes in the manuscript have been highlighted in red. Additionally, please find below a list with our point-by-point answers (*in italic*) to your comments and suggestions.

This manuscript presents a new high-resolution multidecadal reanalysis for the Black Sea (BLK-REA), incorporating key advancements in model resolution, lateral boundary conditions, and data assimilation techniques. The study delivers an extensive validation of the reanalysis and discusses its application to ocean monitoring indicators (OMIs) such as ocean heat content, the Cold Intermediate Layer (CIL), Rim Current variability, and meridional overturning circulation (MOC). These contributions are highly relevant and timely, particularly in the context of climate-related changes in regional semi-enclosed basins.

However, despite its technical strengths, the manuscript requires significant revision before it can be considered for publication. Several methodological and interpretative aspects need clarification, refinement, or deeper discussion. I recommend Major Revisions to address the concerns outlined in the detailed comments. My detailed comments are followed below:

1. The manuscript states in line 178 that a "quasi-independent validation" was performed. However, for an objective and rigorous validation, the use of fully independent observational datasets is typically recommended. In particular, the use of observations that were excluded from data assimilation due to background quality control raises concern about possible contamination by erroneous data. For instance, in Figure 2, there are conspicuous, vertically extensive anomalies in temperature bias around 1998 and salinity bias in 1996, which could be indicative of problematic observational data rather than model performance. A discussion on the possibility of observation-induced artifacts and how such risks were mitigated is strongly recommended.

   We thank the reviewer for this valuable comment and fully agree that, for an objective and rigorous validation, the use of fully independent observational datasets is typically recommended. In our study, we used the term quasi-independent validation to indicate that the validation includes both observations assimilated in the system, which does not constitute a fully independent validation, but also those excluded by the background quality control, thereby encompassing all available observations. We note that all datasets, whether assimilated or used only in validation, had already passed standard quality control procedures prior to assimilation. Nevertheless, some high errors observed in the T/S Hovmöller during the validation stage (Figure 3), around 1996 for salinity and 1998 for temperature, may be associated with observations that were not assimilated but are included in the validation dataset. These errors are likely due to the presence of such data, but also to sparse observational coverage, and local model biases. These points have been clarified in the revised manuscript.

   Revised manuscript text (Lines 162–167):

   > *"Similar to the previous version, a background quality check is implemented in the data assimilation system to reject observations that deviate significantly from the model prior solution. Rejection by*

background quality control does not necessarily indicate erroneous data, but often reflects large innovations that would otherwise introduce undesirable shocks in the model state. The quality control procedures and data rejection thresholds are applied as described in Lima et al. (2021), with no changes introduced in BLK-REA."

and Lines 280-282:

During the validation stage, the elevated T/S errors around 1996 (salinity) and 1998 (temperature) in the Hovmöller diagram (Figure 3) may be linked to observations that were not assimilated but included in the quasi-independent validation. These peaks are likely due to such observations, but also sparse coverage, and local model biases.

2. In Figures 3 and 4, temperature and salinity biases in the Eastern region show a sudden increase below 700 m, which is unexpected. Generally, variability and errors tend to decrease with depth. The authors should examine whether these anomalies arise from insufficient data qulity control, issues in model initialization, or perhaps unresolved physical processes. An explanation or hypothesis for these unusual patterns would improve credibility.

We note the unexpected increase in temperature and salinity biases below 700 m in the Eastern region. This behavior likely results from a combination of factors: sparse observational coverage in deeper layers before 2003 (as already shown in the Hovmöller diagrams, Figure 3), limitations in model initialization, and unresolved physical processes. Even with strict quality control, some measurement errors or inconsistencies may have persisted in the assimilated data. The introduction of Argo profiling floats from 2003 improved deep-layer observations, revealing biases that were previously undetected and not corrected. Additionally, some observations excluded from assimilation were still included in validation, potentially contributing to the apparent increase in errors.

Revised manuscript addition (Lines 326-332):

"In the Eastern region, a slight increase in temperature and salinity biases occurs below 700 m (Figures 4 and 5). This pattern likely arises from sparse deep-layer observations prior to 2003 (as shown in the Hovmöller diagrams, Figure 3), limitations in model initialization, and unresolved physical processes. The introduction of Argo profiling floats from 2003 increased deep data coverage, revealing biases that were previously undetected and never corrected; below 700 m, LSBC toward WOA2018 climatology is not sufficient to constrain the model. In addition, some observations excluded from assimilation were still included in validation, further contributing to the apparent increase in errors. Even with strict quality control, some measurement errors or inconsistencies may have persisted.

3. Although the Eastern region has the highest number of salinity observations (Figure 4, right), the RMSD and bias near the surface are larger than in other regions. The manuscript does not address this, and a discussion is warranted. Potential causes could include persistent local biases in atmospheric forcing, river discharge representation, or misrepresentation of mixing processes.

We thank the reviewer for this comment. Although the comment refers to the Eastern region, the elevated near-surface salinity RMSD and bias actually occur in the Western region (Figure 5). These errors are mainly due to limitations in freshwater inputs from major rivers: only the Danube runoff is represented with monthly varying discharge, while other rivers follow climatologies without intra-annual variability, causing persistent local biases. Also, complex hydrodynamics near the Bosphorus Strait are included through prescribed boundary conditions, but uncertainties in these inputs can further increase local errors. Data assimilation can partially improve the model representation under these circumstances by incorporating observations with strict quality control..

Revised manuscript addition (Lines 319-325):

> *"The same analysis for salinity is shown in Figure 5. The Western region exhibits the highest near-surface RMSD, with values reaching over 0.8 psu in the upper 100 m, despite high observational coverage. In contrast, RMSD values are lower in the central and Eastern regions, remaining below 0.2 psu in the upper layers. These elevated values in the Western region are mainly due to limitations in freshwater inputs from major rivers: only the Danube uses monthly varying discharge, while other rivers follow climatologies without intra-annual variability, producing persistent local biases. Prescribed boundary conditions near the Bosphorus Strait improve the physical representation but also introduce uncertainties. Data assimilation can partially improve model representation under these circumstances by incorporating observations with strict quality control, but some errors can still persist".*

4. The sentence "Bias values are in general small, and in particular in the western region" seems inconsistent with the plotted salinity bias in Figure 4, where large positive biases are evident near the surface in the western region. This discrepancy should be corrected or clarified.

   We thank the reviewer for this observation and we agree that the original wording was inconsistent with the plotted salinity bias in Figure 5. In the revised version, we have corrected the text to better reflect the results, emphasizing that salinity biases are large in the western region near the surface (Lines 302-312).

5. The description of how SLA is assimilated (e.g., through the dynamic height operator and use of MDT) could be elaborated further. It is unclear how the model handles vertical mapping and whether SLA assimilation is constrained only over deep waters. Furthermore, in Figure 5, the RMSD of SLA decreases after 2000-likely due to increased availability of satellite altimetry data. This hypothesis should be explicitly stated and supported with quantitative or qualitative evidence. The temporary increase in SLA RMSD around 2016 is notable but unexplained. Likewise, elevated RMSD near the coastline (Figure 6) may stem from the fact that SLA assimilation is performed only in areas deeper than 1000 m, thus leaving high-variability coastal regions unconstrained. A discussion of these spatial and temporal variations in SLA errors is necessary to clarify the model's performance boundaries.

We thank the reviewer for these very valuable comments regarding the assimilation and validation of SLA. We have expanded the description of the SLA assimilation methodology and clarified the interpretation of the SLA validation results, following the reviewer's suggestions:

In the revised manuscript, we now explain that SLA increments are applied through a balance model (dynamic height operator) that imposes local hydrostatic and geostrophic balance among SLA, temperature, and salinity increments. A "level of no motion" is assumed at 1000 m depth, below which horizontal velocities are considered negligible. Following Storto et al. (2011), we further specified that SLA assimilation is performed via a local hydrostatic adjustment scheme: SLA increments, proportional to the vertically integrated density increments, are partitioned into thermo- and halosteric contributions and then distributed along vertical T/S profiles according to the background-error vertical covariances, thereby restricting SLA assimilation to deep-water regions where the balance assumptions are valid.

We thank the reviewer for this observation. We agree that the decrease in SLA RMSD after 2000 is closely linked to the increased availability of satellite altimetry data, as indicated by the rise in the number of observations (Figure 6). At that time, in situ T/S profiles were still scarce (see number of T/S observations in Figure 3), so the model constraints were essentially driven by SLA assimilation. On the other hand, the temporary increase in SLA RMSD around 2016 coincides with the larger availability of Argo profiles (see number of T/S observations in Figure 3). Their assimilation, in combination with SLA, appears to have degraded the SLA skill, contributing to the observed increase in RMSD. This effect likely reflects the multivariate nature of the system. Nevertheless, the SLA RMSD remained below the instrumental error level of 0.04 m used in the assimilation. These points have been clarified in the revised manuscript.

The revised text now contains sentences to detail the SLA assimilation, which follows Lima et al. (2021) and Storto et al. (2011 (Lines 200-203):

> *The dynamic height operator in V_η imposes local hydrostatic and geostrophic balance among SLA, temperature, and salinity increments, following Storto et al. (2011), with a level of no motion assumed at 1000 m, where this balance is valid. This restricts SLA assimilation to deep-water regions.*

and in Lines 333–339, we have refined the description of SLA validation and model skill:

> *The time series of spatially averaged SLA RMSD (Figure 6) shows strong skill after 2000, with errors around 0.02 m, below the instrumental error of 0.04 m used in assimilation. This improvement reflects the increased availability of satellite altimetry, while in situ T/S profiles remained scarce, leaving SLA as the main constraint on the model. Instead, the slight RMSD increase around 2016 coincides with the larger availability of Argo profiles. Their assimilation, together with SLA, may have slightly degraded SLA skill due to the multivariate nature of the system. Nevertheless, the RMSD remains well within acceptable limits, and these points are intended to provide an overview of SLA*

performance rather than a detailed attribution of small temporal fluctuations.

6. The recovery of the Black Sea Meridional Overturning Circulation (MOC) after 2010 is intriguing, especially given the concurrent decline or near-disappearance of the Cold Intermediate Layer (CIL). Although the authors note this decoupling, a deeper discussion on alternative dynamics that might drive the MOC increase (e.g., changes in wind-driven circulation, lateral advection, or water mass transformation) would be beneficial. The potential use of age-tracer or Lagrangian particle tracking experiments to explore such processes could be suggested for future work.

   We thank the reviewer for the suggestion. An ideal age tracer release would be a huge asset, not just for understanding the Meridional Overturning Circulation (MOC) but also for revealing the Black Sea's ventilation processes and how its deep basin changes over time. Additionally, we could use spiciness to investigate if the temperature increase is being compensated for by the increase in salinity which might be due to E-P budget changes. We also add the following part in the discussion (Lines 466-472):

   > *"Different water mass transformations could be the potential mechanisms behind an increase in the MOC. Specifically, an increase in salinity could compensate for the decrease in the formation of cold, dense water, which would otherwise weaken the circulation. Reanalysis model results show that there is an upward trend of SSS in the Black Sea. In addition, running multiple cycles of decadal reanalysis simulations is likely necessary to achieve a more accurate spin-up of the deep ocean. However, to investigate the detailed dynamics of the meridional overturning circulation (MOC) in the Black Sea, further research is required. This should involve using multi-cycle reanalysis model simulations combined with passive tracers, such as ideal age, to better understand the circulation patterns and timescales."*

7. Much of the text after line 419 in the "Conclusions" section reads more like future planning than summarizing key findings. It may be clearer and more structured to rename this section "Discussion and Outlook", where the manuscript first summarizes the conclusions, then outlines future improvements and data needs.

   We have renamed Section 4 to "Discussion and Outlook" and revised the text in Lines 103–104 in accordance with the reviewer's suggestion.

References

Lima, L., Ciliberti, S. A., Aydoğdu, A., Masina, S., Escudier, R., Cipollone, A., Azevedo, D., Causio, S.; Peneva, E., Lecci, R.; Clementi, E., Jansen, E., Ilicak, M.; Cretì S., Stefanizzi, L., Palermo, F., & Coppini, G. (2021) Climate Signals in the Black Sea From a Multidecadal Eddy-Resolving Reanalysis. Front. Mar. Sci. 8:710973. doi: 10.3389/fmars.2021.710973.

Storto, A., Dobricic, S., Masina, S., and Di Pietro, P. (2011). Assimilating along-track altimetric observations through local hydrostatic adjustment in a global ocean variational assimilation system. Mon. Weather Rev. 139, 738–754. doi: 10.1175/2010mwr3350.1