

## **Response to reviewer #1**

This study provides valuable insight by demonstrating, through various experiments in the Mediterranean Sea, that the choice of vertical mixing scheme is an important factor in reproducing oceanic physical variables. In this regard, the manuscript has sufficient merit and relevance for readers.

However, several aspects require more thorough discussion. If these points are adequately addressed, the study could serve as a useful reference for ocean numerical modeling.

We thank the reviewer for the constructive and thoughtful comments, which significantly helped improving the manuscript. We have addressed the points raised and implemented the requested revisions throughout the paper. During the revision process, we have also identified and corrected a minor bug in our simulations, which does not affect the results or conclusions of the manuscript; the figures have been updated accordingly.

A detailed response and explanation are provided below (reviewer's questions are in blue, our answers in black). Line numbers refer to the revised manuscript with track changes enabled, in No Markup view.

### **[Major Comments]**

1. This study compares differences in the vertical structure of the Mediterranean Sea associated with different vertical mixing schemes and convective adjustment, primarily focusing on basin-averaged scales. However, the current discussion remains somewhat superficial, and a more in-depth analysis is needed.
2. In particular, more detailed explanations are required regarding how the schemes and EVD operate in key regions where major water masses are formed, such as deep water and intermediate water formation areas in each experiment. It would also be valuable to describe the physical processes through which these differences develop over time.
3. Rather than simply describing differences among experimental results, the discussion should further address the physical mechanisms responsible for these differences and elaborate on their scientific implications.

We thank the reviewer for these insightful comments. As all these three comments concern the need for a more in-depth and process-oriented discussion of the differences among the experiments, we address them together here. We have substantially expanded the discussion to provide a more in-depth analysis of the differences among the experiments, with a stronger focus on the underlying physical mechanisms.

In particular, we have:

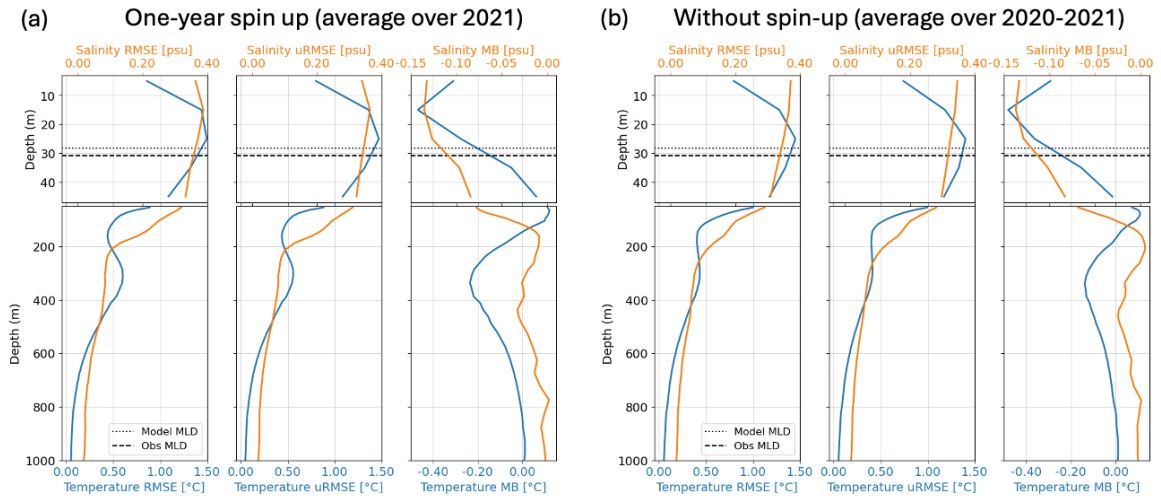
- Expanded the description on **how the different mixing schemes handle convection** (section 2.2.1, lines 205-213, 225-229, 256-263) and explicitly linked the results to the physical description of the mixing schemes, ultimately moving from a descriptive analysis to a mechanistic one;
  - added a **seasonal analysis of Argo profiles in key deep and intermediate water formation regions**, namely the South Adriatic Pit and the Rhodes Gyre region through the addition of two figures (Fig. 1d and 1e) and a discussion in the manuscript (lines 129-142);
  - **expanded the discussion of deep-water formation in the South Adriatic**, both in the text (Section 3.3.1) and through the addition of a new figure (Fig. 10) showing the temporal evolution of the salinity field, highlighting how the different vertical mixing schemes and convective adjustment affect local water mass properties and their temporal evolution. We also further extended this analysis to **include the role of Levantine Intermediate Water intrusion** and its **evolution over time**, providing additional insight into the mechanisms driving the differences among the experiments (lines 564-597);
  - **reshaped the entire paragraph on the Rhodes Gyre** (Section 3.3.2) by: a) slightly enlarging the region (by 1 degree east and south) to capture not only LIW formation but also the beginning of its spreading (lines 604-610 and Figure 1a); b) moving Fig 11 from a representation of the mean bias to the absolute value of salinity, which allows us to discuss the physical processes rather than just comparing the model output to observations; and c) discussing the results by first defining the physical processes observed in the region and then explicitly linking the discrepancies in the results to the differences in the physics of each mixing scheme (lines 614-641).
  - **expanded the discussion of how summer convection is represented** by the different mixing schemes and the role of convective adjustment, through the addition of a new figure (Fig. 12) showing the **temporal evolution** of vertical mixing and summer convection in terms of eddy diffusivity coefficient; a detailed discussion of the underlying mixing processes in this context has been added in lines 665-695.
  - **Reshaped the entire conclusive section** (Section 4) to summarize our findings on the connection between the internal logic of the schemes and the resulting physical response of the basin.
4. In Section 3.1, before discussing the differences between the Control experiment and observations (OBS), it would be necessary to first present the differences between the initial conditions used in this study and the observations. This information would provide an important basis for interpreting how the model results shown in Figs. 4 and 5 are actually influenced by the vertical mixing schemes.

For example, it should be clarified whether the statistical metrics presented in Figs. 4 and 5 primarily reflect characteristics of the initial conditions, or whether they result from physical processes associated with the vertical mixing schemes that amplify or reduce model errors.

We thank the reviewer for giving us the opportunity to clarify this important point, which was indeed not well explained in our original manuscript. In our original manuscript, we

only mentioned that the same initial conditions were used across all experiments, taken from the Mediterranean Forecasting System (section 2.3). We have also clarified in the manuscript that these initial conditions are analyses released by the Mediterranean Forecasting System of the Copernicus Marine Service that incorporate data assimilation, thus providing an optimal starting point very close to the observed state of the Mediterranean Sea. We have added this clarification in the manuscript (lines 323-325).

Because of that, the conclusions drawn from the statistical metrics in Fig. 5 (former Figs. 4 and 5) do not reflect the initial condition, but rather the changes in vertical mixing scheme and convective adjustment. In the following figure, we compare the misfits with and without a spin-up period. The figure shows the basin-average temperature and salinity Root Mean Square Error (RMSE), the unbiased RMSE (uRMSE) and the Mean Bias (MB) for the control run, considering: (a) averaging over 2021 with a one-year spin up, and (b) averaging over 2020-2021 without a spin-up period. We observe very minor differences and the main features described in section 3.2 and Fig. 5 (former Figs. 4 and 5) remain the same and the overall conclusions of the paper remain unchanged. We added a sentence concerning this evidence in the manuscript (lines 325-327 and lines 380-381).



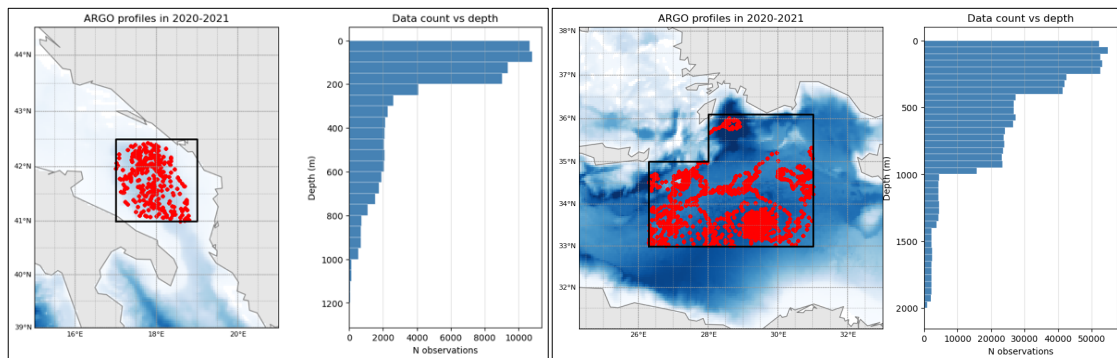
- In addition to Argo data, incorporating ship-based observational datasets (e.g., repeated hydrographic sections or long-term monitoring stations), if available, could provide a more comprehensive observational basis and allow for a more detailed discussion of experiment-specific characteristics across different regions of the Mediterranean.

We thank the reviewer for this valuable suggestion. While we agree that ship-based observational datasets (e.g., repeated hydrographic sections or long-term monitoring stations) can provide important complementary information, such datasets are not part of our observational dataset. If data from other sources were incorporated, such datasets may follow different quality control procedures and therefore would not be fully homogeneous. Such datasets would also be significantly more limited in spatial and temporal coverage compared to Argo observations.

The Argo dataset used here has undergone careful quality control as part of the Global Ocean-In-Situ Near-Real-Time Observations dataset (<https://doi.org/10.48670/moi-00036>), and it is operationally employed by the Mediterranean Forecasting System to produce the physical component of the Copernicus Marine Service product. As such, it provides a robust and consistent observational basis across the entire Mediterranean basin.

In addition, the dataset used here is already very extensive, comprising 11,053 Argo profiles across the Mediterranean Sea collected over a two-year period. This large number of observations allows for a statistically robust analysis of the processes under investigation. The two deep water formation regions analyzed in this study are characterized by particularly high data coverage, as illustrated by the maps and by the histograms showing the vertical distribution of observations included below.

We have added these considerations at lines 113-115.



### [Specific Comments]

1. (line 84) Although statistical differences in deep water and intermediate water formation regions are compared across experiments, it is unclear whether a two-year model integration is sufficient to evaluate these differences. In particular, it should be examined whether spin-up effects may still influence the results.

Following our response to Major Comment 4, we moved the description of the initial conditions to Section 2.3 (lines 323-327). Moreover, the formation of deep and intermediate waters occurs on seasonal timescales, with interannual variability. Therefore, we clarified that a two-year model integration is sufficient to capture the main differences between experiments arising from changes in vertical mixing and convective adjustment at lines 96-98.

2. (line 94) The phrase “our model domain” in the description of Argo float data sounds somewhat awkward.

We thank the reviewer for pointing out this inconsistency. We agree and we have moved the paragraph describing the geographical limits of our model domain from Section 2.1 (“Argo float data in the Mediterranean Sea”) to Section 2.2 (“Model setup”) at lines 152-159.

3. (Fig. 1) It would be helpful to label major sub-basins of the Mediterranean on the map and include a schematic of major currents and regions where convection frequently occurs. In particular, key regions repeatedly mentioned in the manuscript should be clearly indicated to help readers better understand the regional characteristics.

We added labels to Fig 1a indicating the major sub-basins, deep and intermediate water formation regions, and the two major straits connecting the Mediterranean Sea to the Atlantic Ocean and the Black Sea. This highlights the key regions discussed throughout the manuscript and provides readers with a clearer view of the regional characteristics; we thank the reviewer for this suggestion.

4. (Section 2.1) The study presents basin-averaged vertical profiles of temperature and salinity using Argo float data. However, presenting seasonal vertical profile characteristics specifically in deep and intermediate water formation regions would more clearly illustrate the effects of different vertical mixing schemes and convective adjustment.

We added two panels to Fig. 1, showing Argo profiles averaged by season in the South Adriatic Pit (Fig. 1d) and the Rhodes Gyre region (Fig. 1e). We also expanded the manuscript (lines 129-142) with a discussion of the seasonal variability of these profiles in relation to deep and intermediate water formation and spreading.

5. (line 130) The longitude and latitude ranges should be clearly specified.

We added the longitude and latitude ranges of our model domain at lines 154-155.

6. (line 156) In the sentence “The latter rely on ~~,” it should be clarified what “the latter” refers to.

We modified the sentence (now at lines 193-194) to make it explicit that we are referring to physics-based parameterizations mentioned in the previous sentence.

7. (Fig. 4) Would it be possible to present basin-averaged mixed layer depth by showing both model results and observations together for comparison?

The basin-average mixed layer depth is presented in Fig. 8 and discussed in Section 3.2.2 (“Representation of the mixed layer depth”). However, we agree with the reviewer that providing the time- and space-averaged values of both observations and the control simulation in Section 3.2.1 (“Impact of tracers”) would help the discussion. Therefore, we have modified Fig. 5 (previously Figs. 4 and 5) by adding dashed lines indicating the observed and modeled MLD values. We have also included this information in the figure caption and in the text at line 392.

8. (line 341) Clarification is needed regarding the meaning of negative RMSE values. How can RMSE take negative values?

We thank the reviewer for pointing out this apparent issue, which arises from the inconsistency between the nomenclature in Equation (1) and Figs. 6 and 7, and the one used in the text. Indeed, we were referring to the RMSE difference, as described in equation (1), which can be positive or negative. To avoid confusion and to be constant with equation (1) and Figs. 6 and 7, we have modified the nomenclature in the text between lines 432-

456: when referring to the error difference (called  $\text{err}_{\text{diff}}$  in Equation (1)), for whatever the metrics, we now have added a suffix like in equation (1) (that is:  $\text{MB}_{\text{diff}}$ ,  $\text{uRMSE}_{\text{diff}}$ ,  $\text{RMSE}_{\text{diff}}$ ,  $\text{CC}_{\text{diff}}$ ).

9. (line 351–353) A more detailed explanation is needed regarding why the result described in “The Richardson-dependent scheme without EVD ~~~” was expected.

We have modified the text (now lines 447-450) to clarify that this behavior is expected because this parameterization estimates vertical mixing solely from the Richardson number, i.e., from the balance between shear and stratification, and it does not include convective mixing.

10. The location of the Rhodes Gyre region should be indicated in Fig. 1.

We added the Rhode Gyre region in Fig 1a.

11. The Rhodes Gyre region should be defined and its location presented when it is first mentioned in the manuscript.

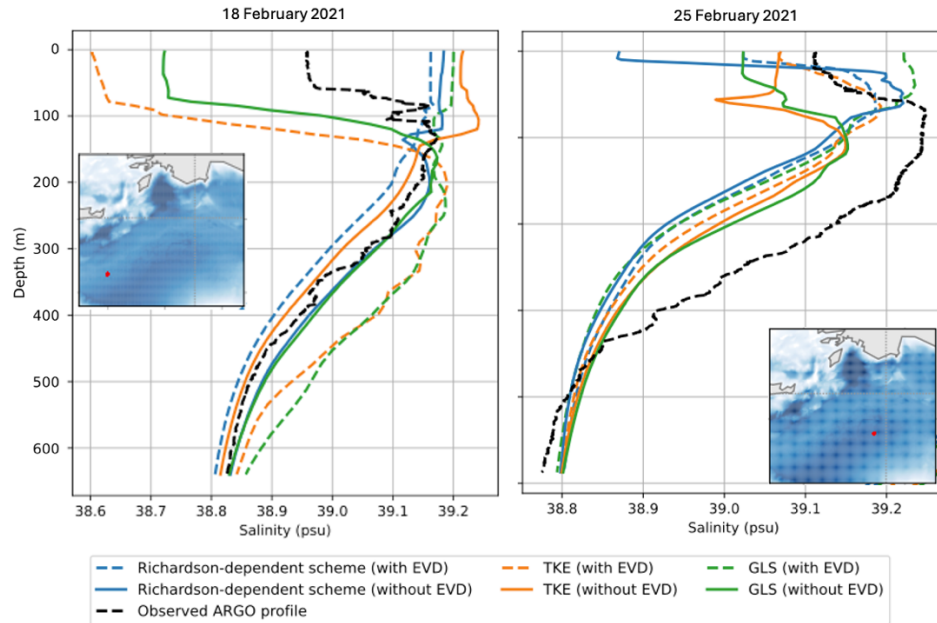
We moved the definition of the location of the Rhodes Gyre region to the beginning of section 3.3.2 (lines 605-606), where the region is described. Moreover, the location of the Rhodes Gyre region, together with the other deep and intermediate water formation regions, has been added to Fig. 1a. and the description of all the deep and intermediate water formation regions have been also added in Section 2.1 (lines 139-142).

12. (Figure 11) The manuscript states that the largest errors occur between 200–500 m depth across all experiments. However, the figure appears to show that intermediate-depth errors increase at specific times after model integration begins. A discussion is needed to explain the physical processes responsible for the formation of these errors. For example, it should be considered whether differences in intermediate water formation associated with different schemes lead to the development of these errors during periods of water mass formation.

In response to the reviewer’s requests to better explain the physical mechanisms underlying the differences among the schemes (Major requests 1-3), we have enlarged the Rhodes Gyre region by  $1^\circ$  east and south and replaced the former Fig. 11 (mean bias) with a Hovmöller diagram of absolute salinity. The new figure shows observations and model data from all available Argo floats between January and April 2021 (i.e., around the LIW formation period), arranged over time regardless of their location within the region. Each vertical line represents a full vertical profile as a colored column, so the observed increases in intermediate-depth errors at specific times correspond simply to errors of individual Argo profiles rather than systematic model biases.

An example of two salinity profiles is shown below in this letter for 18 February (left) and 25 February (right) at two locations within the region. We note that on 25 February all models exhibit a bias relative to the observations, corresponding to the intermediate-depth deviations see at specific times in the former Fig. 11, whereas on 18 February no bias is present, consistent with the times when deviations were absent. Importantly, the profile shapes are well captured by the TKE and GLS models without EVD (solid orange and

green lines), indicating that the underlying physical processes are represented correctly. The new Fig. 11, by showing these profiles directly as absolute salinity instead of the mean bias, allows us to better illustrate this point and address the reviewer's question (lines 612-641).



- It would also be beneficial to extend the discussion to regions with characteristics similar to the Mediterranean, and to elaborate on how the findings of this study could be considered or applied in a broader regional context.

We thank the reviewer for this suggestion. To address this point, we have extended the discussion in Section 4 (lines 789-794) by commenting on the potential relevance of our results for regional ocean modeling in other basins with characteristics similar to the Mediterranean Sea.