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:

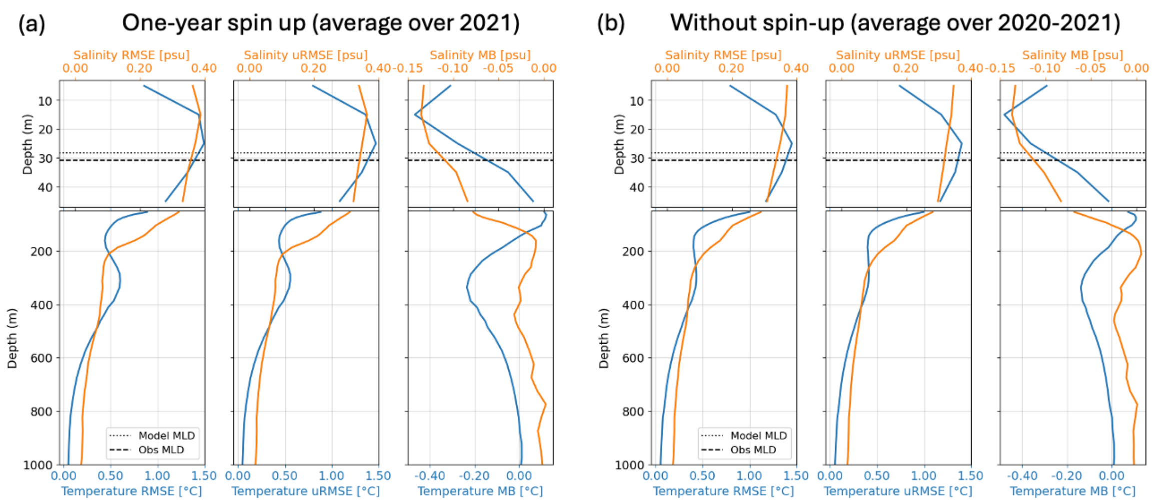
- Expanded the description on **how the different mixing schemes handle convection** (section 2.2.1, lines 205-213, 225-229, 256-263) and explicitly linked the results to the physical description of the mixing schemes, ultimately moving from a descriptive analysis to a mechanistic one;
- added a **seasonal analysis of Argo profiles in key deep and intermediate water formation regions**, namely the South Adratic 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 Main 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).



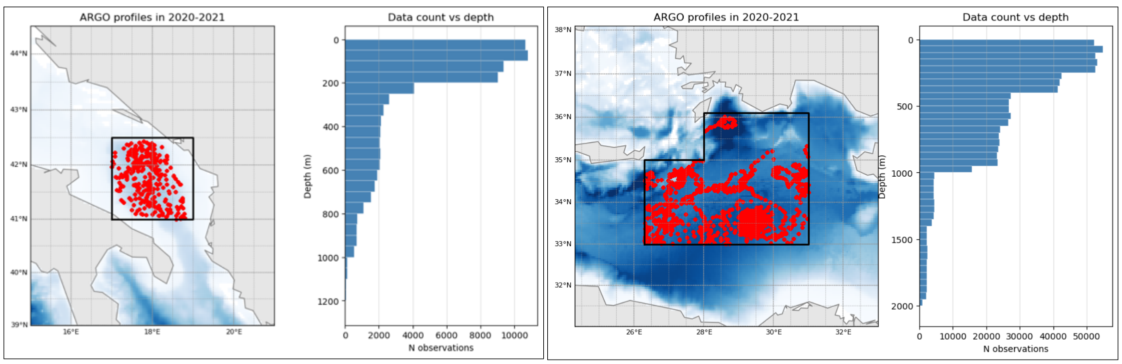
5. In addition to Argo data, incorporating ship-based observational datasets (e.g., repeated hydrographic sections or long-term monitoring stations), if available, could provide a more comprehensive observational basis and allow for a more detailed discussion of experiment-specific characteristics across 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 Marin 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 err_{diff} in Equation (1)), for whatever the metrics, we now have added a suffix like in equation (1) (that is: MB_{diff}, $uRMSE_{diff}$, $RMSE_{diff}$, CC_{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 Richarson 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.

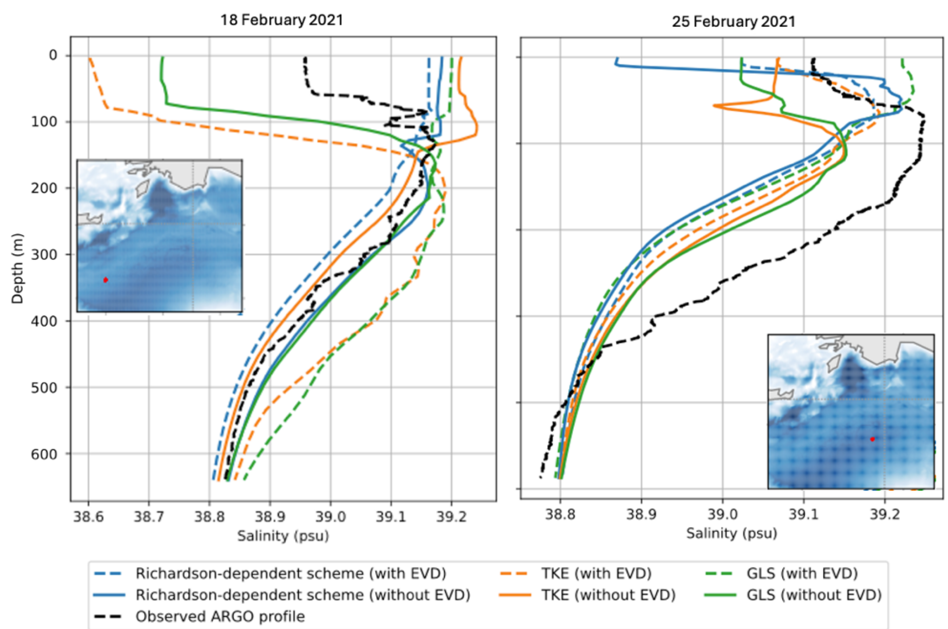
11. The Rhodes Gyre region should be defined and its location presented when it is first mentioned in the manuscript.

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° 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 for18 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 process 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).



13. 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.

Response to reviewer #2

General comment

This manuscript investigates the impact of different vertical mixing schemes and convective adjustment in a regional Mediterranean Sea model. The topic is relevant for ocean modeling, and the experimental design comparing several turbulence closure schemes with and without convective adjustment is useful. However, the manuscript mainly presents a sensitivity comparison between different parameterizations, while the mechanisms driving the simulated differences are not sufficiently analyzed. A deeper analysis of these mechanisms would significantly improve the manuscript. In particular, the role of convective adjustment in relation to turbulence closure schemes is not clearly justified. Overall, the study is potentially valuable, but major revisions are required before it can be considered for publication.

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 answer in black). Line numbers refer to the revised manuscript with track changes enabled, in No Markup view.

Major concerns

1. The title *"***Untangling the effects of vertical mixing schemes and convective adjustment in the Mediterranean Sea***"* suggests that the manuscript aims to clearly separate the roles of these two processes. However, the current presentation appears closer to a sensitivity comparison between parameterization choices and combination, and the extent to which the individual effects are truly disentangled remains somewhat limited. If the title is to remain unchanged, the analysis may need to be strengthened with clearer process-oriented diagnostics and attribution analysis, in order to better identify the respective roles of vertical mixing and convective adjustment.

   Thank you for this comment. In the revised version of the manuscript, we have substantially strengthened the analysis and expanded the discussion to provide a more in-depth and process-oriented comparison among the experiments, moving beyond a primarily descriptive assessment toward a more mechanistic interpretation.

   The revised manuscript now places stronger emphasis on the underlying physical mechanisms, highlighting how the different vertical mixing schemes and convective adjustment impact local water mass properties and temporal evolution. We now first define the key physical processes occurring in the region and then explicitly relate the discrepancies among the simulations to differences in the physics represented by each mixing scheme in the presence and absence of convective adjustment. This improvement has been achieved through a through revision of large parts of the Results (especially for what concerns the deep and intermediate water formation regions) and the conclusive section, as well as with the addition of several new figures (Figs. 1d, 1f, 2, 10, 11, and 12) which support a cleared attribution of the observed differences.

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 Adratic 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.

Overall, the manuscript now more explicitly addresses the "why" behind the results, linking them to the physical behavior of the parametrizations. Considering these substantial changes, and in order to better reflect both the scope and the methodological framework of this study, while specifying that it is based on sensitivity experiments, as the reviewer suggested, we propose the following revised title: **Untangling the effects of vertical mixing schemes and convective adjustment in the Mediterranean Sea: insights from a sensitivity study.**

2. The manuscript discusses the combined use of turbulence closure schemes and convective adjustment. However, turbulence closure schemes such as TKE and GLS already represent convective mixing through buoyancy production when stratification becomes unstable**.** In this

context, introducing an additional convective adjustment does not represent a separate physical process, but rather an alternative term of the same mechanism. Therefore, the manuscript should clarify the physical rationale for applying a convective adjustment together with these turbulence closures. In particular, it should explain whether the closure schemes alone fail to adequately represent convective mixing, and why an additional adjustment is considered necessary.

We thank the reviewer for raising this important point, which was not sufficiently explained in the original manuscript and indeed deserves further clarification. We agree that turbulence closure schemes such as TKE or GLS include buoyancy production terms and therefore, in principle, are able to represent convective mixing when the stratification becomes unstable. However, in practice, these schemes may not always remove static stabilities rapidly enough, particularly at the temporal and vertical resolution typically used in regional ocean models. As a consequence, convective mixing may be underestimated, and this potential issue remains an open question.

For this reason, many ocean circulation models combine turbulence closure schemes and a convective adjustment parameterization (e.g. the physical component of the Mediterranean Forecasting System of the Copernicus Marine Service). The convective adjustment acts by stabilizing the water column, allowing for a rapid vertical redistribution of tracers and a prompt restoration of stabile stratification. In this sense, it does not represent an independent physical process, but rather a numerical parameterization designed to mimic the fast overturning associated with unresolved convective events.

In this regard, the literature still provides limited assessment of the combined effects of turbulence closure schemes and convective adjustment, and whether a mixing scheme alone (i.e., without convective adjustment) is sufficient to represent the vertical structure and variability of the upper ocean remains unclear. Our results suggest that this capability may depend on the specific closure scheme used, as we find that GLS is able to reproduce these features without convective adjustment, while TKE benefits from its inclusion. To clarify this aspect, we have revised the manuscript by adding a short explanation of the physical rationale for applying a convective adjustment together with physics-based vertical mixing schemes in the Introduction (lines 51-67 and 78-80) and in the conclusive section (lines 719-722).

Minor comments

1. In several places, the manuscript refers to the "**interaction**" between turbulence closures and convective adjustment. This terminology may be misleading and could be clarified.
   We thank the reviewer for this comment. We agree that the term *interaction* could be misleading in this context and have replaced it throughout the manuscript with a more appropriate wording to emphasize that we are studying here the *combined effect* of these two parameterizations.

2. The introduction identifies the combined use of turbulence closure schemes and convective adjustment as a research gap. However, the physical motivation for this question is not entirely clear, since turbulence closures such as TKE and GLS already represent convective mixing through buoyancy production.

As explained above (point 2 of the Major Comments), to clarify the physical motivation for combining turbulence closure schemes with convective adjustment, we have revised the manuscript by adding new paragraphs in the Introduction (lines 51-67 and 78-80) and in the Conclusions (lines 79-722).

3. The manuscript does not clearly explain why convection is parameterized through enhanced vertical diffusivity. Convection is a non-hydrostatic overturning process that differs from shear-driven turbulent mixing, and the rationale for representing it simply by increasing the vertical eddy diffusivity should be clarified.
We thank the reviewer for pointing out the need to clarify the rationale behind using the enhanced vertical diffusivity to represent convection. In the revised manuscript, we have added the following paragraph to Section 2.2.2 to clarify this point:

Since convection has the effect of effectively homogenizing the upper ocean boundary layer, it can be parameterized by applying a high eddy diffusivity which has a similar effect. Although this procedure does not accurately represent the physics of convection with its narrow and dense downward plumes compensated by a broader upwelling in between, leading to non-local fluxes, it effectively mimics the net impact of convective processes, including the rapid vertical redistribution of properties associated with convective overturning. The increased eddy diffusivity coefficient stabilizes the water column, allowing the model to maintain realistic stratification while enabling the development of large-scale circulation patterns.

4. The description of the model configuration and parameterization schemes in Lines 75–80 would be more appropriate in Sections 2.2–2.3.
We agree that the description of the model configuration is more appropriate in the “Data and methods” section rather than in the Introduction. Accordingly, we have moved the description of the numerical models (formerly lines 75-80) to Section 2.2 and the description of the parameterizations to Section 2.3.

5. Line 130, Please provide more details on the vertical grid distribution (e.g., layer thickness near the surface and within the mixed layer), as this may influence the representation of mixed-layer processes.
We added a new figure to the manuscript (Fig. 2) showing the vertical distribution of the layer thickness, with a zoom in the upper 50 m. The discretization is particularly dense in this region (2 m near the surface to 4.6 m at 50 m depth) to accurately capture processes within the mixed layer. A brief description of the figure and comment on this dense upper-ocean discretization have been added (lines 173-176).

6. The manuscript states that the model is nested into the “**daily analysis and forecast data of the Copernicus Marine Global product**”. It is unclear whether the boundary conditions are taken from the analysis fields, the forecast fields, or a combination of both. Please clarify this point. In addition, it would be helpful to explain why this product is used instead of the available reanalysis product, which is commonly used for ocean model boundary forcing
We thank the reviewer for pointing this out and we apologize for the confusion. We had included the name of the Copernicus product we used, but in this study, we only use analysis

data. Therefore, we have removed the word *forecast* from this sentence and clarified that *Global Ocean Physics Analysis and Forecast* is the name of the Copernicus product as listed in the online catalogue.
The reason we use this product at the Atlantic boundary is that our reference model for this study is the physical component of the Mediterranean Forecasting System of the Copernicus Marine Service, and we therefore use the same input datasets. This has also been specified in the manuscript (lines 296-299).

7. The authors should justify the choice of the typical value of 10 $m^2 s^{-1}$ for the convective adjustment used in the simulations.
   Typical values recommended in the NEMO manual range from 1 to 100 $m^2 s^{-1}$. A value of 10 $m^2 s^{-1}$ was chosen based on previous sensitivity tests and is the value currently used in the physical component of the Mediterranean Forecasting System of the Copernicus Marine Service. We added a sentence in the manuscript (lines 275-276) to include this information.

8. Figures 4 and 5 could be combined into a single figure like figure 1b.
   Following this recommendation, Figures 4 and 5 have been combined into a single figure (similarly to Figure 1b), now Figure 5. Presenting the results together makes the differences between the temperature and salinity misfits more evident and allowed us to expand the related discussion in the manuscript. We thank the reviewer for this suggestion.

9. Figure 8 shows a basin-averaged MLD from the model together with an "**observed MLD**" derived from Argo observations. However, it is not clear how the comparison is performed. In particular, Argo profiles do not provide coverage over the Mediterranean basin, please clarify whether (1) the model MLD is averaged over the entire basin and compared with an average observation directly from the available Argo profiles, or (2) the model fields are first sampled at each Argo location, and then averaged all the profiles. A clear description of the methodology used to construct the basin-averaged time series would be helpful.
   We thank the reviewer and agree that this point was not clearly explained in the manuscript. The method used here corresponds to the second approach hypothesized by the reviewer: the model fields are first sampled at each Argo location, MLD is computed for each profile in both models and observations and then averaged over all the available profiles. We have improved the description of this methodology in the manuscript (lines 471-473).

10. The purpose of Figure 10 is not entirely clear. Since the water mass formation rate is computed only from model outputs without observational reference, it would be helpful to clarify what additional insight this diagnostic provides.
    We thank the reviewer for this comment and agree that this diagnostic does not provide substantial additional information beyond what is already presented in the manuscript. After careful consideration, we decided to remove the figure and the related text to improve the clarity and focus of the manuscript.

11. The analysis of "**summer diurnal convection (Section 3.4)"** mainly shows temperature profile evolution at a single location. It remains unclear how convection is diagnosed in this section. Additional diagnostics or clearer explanation would help support the interpretation.

We thank the reviewer for this comment. We agree that Section 3.4 benefits from additional diagnostics and explanation regarding how summer convection occurs and how it translates into differences in temperature. To address this, we have added a new figure (Fig. 12) showing Hovmöller diagrams of the hourly eddy diffusivity coefficient for the six simulations at the same location and over the same three-day period. We also added a new portion of the paragraph (lines 665-684) describing how summer convection is represented by the different schemes, both with and without EVD, and linking these patterns to the differences in temperature observed in the following figure (lines 691-695 and lines 710-713).