Reviewer #1:

This manuscript uses a cyclone tracking algorithm to track the most intense Mediterranean cyclones in atmosphere-only and coupled atmosphere-ocean simulations. It is shown that both STD and CPL simulations represent the climatology of storms in the Mediterranean with no notable advantage to CPL. Also, it is shown that CPL has an SST bias relative to STD, which affects various fields in the PBL. CPL can be used to understand dynamical mechanisms in the ocean mixed layer in the presence of atmospheric cyclones.

Overall, the work performed for this study is impressive – a combination of coupled model simulations with cyclone tracking algorithms and the subsequent analysis. The presentation of the results and the discussions are interesting and well-structured. Yet, I struggle to see the innovative part of this paper. Instead, I see a nice comparison between two simulations – coupled and uncoupled. Ultimately, the primary distinction between simulations is the SST bias in CPL, which imprints on various fields in the PBL. The relevancy of the coupling vs. non-coupling is only demonstrated in Figure 10, Panels a and b. There, the advantage of coupling is clear.

I think it would be helpful to use more careful language that does not attribute the difference of various fields to the coupling (or the explicitly resolved SST). These suggested changes perhaps mean that the conclusions are more relaxed. Still, considering the large effort made by the authors, after addressing the above critics and the more specific comments below, I would recommend the paper for publication.

We thank the Reviewer for dedicating time to review our manuscript and for the detailed observations that have raised the quality of the paper.

The valuable feedback has helped us improve the clarity of the work and better emphasise its results and novelty. Specifically, we have significantly revised the Results section to provide a more detailed explanation of the physical mechanisms behind the differences in the atmospheric processes between CPL and STD during extreme Mediterranean cyclones and to address the specific questions raised by the reviewer.

In addition, we have revised the conclusion to emphasize the novelty of the study. Specifically, we modified lines 432-434, as follow:

"*This study investigates for the first time (to the best of our knowledge) how extreme Mediterranean cyclones affects simultaneously the atmosphere and the ocean at different vertical levels, comparing two high-resolution RCM simulations, one atmosphere-ocean coupled (CPL) and one atmosphere stand-alone (STD), over the period 1982-2014*.",

and lines 454-456, as follow:

"*This research highlights the ability of the coupled model to coherently simulate the entire atmosphere-ocean system, thus providing new insights into how sea surface energy is redistributed between the atmospheric boundary layer and the ocean mixed layer, and how this impacts the precipitation and the wind speed during extreme cyclone events*.".

Please note that figures 4, 7 and 9 have been revised following the reviewers' comments. The figures now show latent and sensible heat fluxes, 10 m wind speed, potential temperature lapse rate, convective precipitation and total (large scale + convective) precipitation.

Please also note that the text-line references mentioned in our responses correspond to the revised manuscript.

Below, the Reviewer will find our detailed, point-by-point, answers.

Specific comments:

Lines 103-104: The authors try to answer the question, "To which extent in the vertical column, and through which physical mechanisms, the explicitly resolved SST distribution and sea surface fluxes impact the precipitation, and the wind speed during extreme cyclones?", but the designed simulations can't really separate the effect of explicitly resolved SST when the SST in the Western Mediterranean is about 1.5 degrees warmer. In that case, I think the only thing that can be done is to downgrade the question to something that fits the analysis in the paper.

We thank the Reviewer for highlighting this. We have revised the second research question to better align with the scope of the analysis. The updated question is (L103-105):

"*How do differences in SST distribution shape atmospheric processes within the planetary boundary layer (PBL) during extreme cyclone events, and how do these mechanisms, in turn, impact the cyclone-related precipitation and wind speed?*"

Line 167-168: "Two cyclones are considered the same event if their minimum of SLP is within a 500 km distance and within a time range of 12 hours." – this stage removes from the analysis all cases in which CPL is different from STD (~30%). This difference by itself sounds very large, suggesting that many cyclones are represented very differently in the CPL simulation. The authors do not compare their tracks with ERA5-based tracks (e.g., by calculating the RMSE of the distance between observed and simulated cyclone location at maximum intensity), and it is hard to tell which simulation is better. Therefore, it may lead to an unverified conclusion that CPL represents the cyclones well (although it has a large SST bias) and that CPL does not have an added value.

We thank the Reviewer for the valuable feedback. First, we would like to point out that there was a typo error in the manuscript: the same extreme cyclones between CPL and STD are 312 and not 341. This corresponds to approximately 62 % cyclones of the 500 most intense cyclones.

We agree with the Reviewer's observation that the extreme cyclones in common between STD and CPL can appear substantially different. However, similar differences were previously found also in Flaounas et al. 2018a (around 60 % of common cyclones). In fact, differences in SST fields influence the distribution of the SLP, even if it is forced by the same large-scale atmospheric condition (ERA5 in this case). This does not imply that the cyclones themselves are fundamentally different but rather that a different SST may affect the timing and location of the cyclones SLP minima. To further test this statement, we changed the 500 km distance and the 12-hour time window criterion, that qualifies two cyclones as being "the same" in the two simulations, with 1000 km (i.e. the maximum area of influence of Mediterranean cyclones, Flaounas et al. 2016) and 48-hour. In this case, we found 458 common events (92%) between CPL and STD. For these 458 cases, we repeated the same analysis (shown in Figure A1 and attached below). The results are very similar to those obtained for the original 312 common cases (Fig. 7), indicating that our main results are not sensitive to the chosen constraints or to the number of selected "common" cyclones.

Regarding the track comparison between RCMs and ERA5, our objective is not to determine whether the CPL or STD tracks better match ERA5-based tracks but to better understand their development mechanisms regards to air-sea interactions. Nevertheless, we calculated the RMSE and spatial correlation (R, Pearson correlation coefficient) for the locations of the minima of the 500 most intense cyclones in CPL and STD against ERA5. The metrics are similar for both models, with CPL showing slightly lower RMSE (2.16 vs. 2.17 in STD) and slightly higher R (0.73 vs. 0.72 in STD). This is attributed to CPL's higher frequency of cyclones in the Tyrrhenian Sea, which aligns more closely with the ERA5 distribution (see Fig. 3b vs. 3c).

We have clarified in section 2.2.2 (model comparison) the methodology applied to identify common cyclone events, as follow (L176-184):

"*To compare CPL with STD in terms of sub-daily fields associated to the cyclones, the same events between the two simulations are selected. Two cyclones are considered the same event if their minimum of SLP is within a 500 km distance and within a time range of 12 hours. With these criteria, a total of 312 cyclones from the 500 most intense (around 62 %) are found in common between CPL and STD, of which 129 occurring in winter (DJF), 110 in spring (MAM), 17 in summer (JJA) and 56 in autumn (SON). This well align with results from Flaounas et al. (2018a), who also found that approximately 60% of the 500 most intense cyclone tracks were consistent between the coupled and standalone RCMs, using similar identification criteria. Extending the distance criterion to 1000 km (i.e. the maximum area of influence of Mediterranean cyclones, Flaounas et al. 2016) and the time window to 48-hour, the percentage of detected cyclones in common between STD and CPL increases to 92%, but the outcomes of this study do not change (not shown).*"

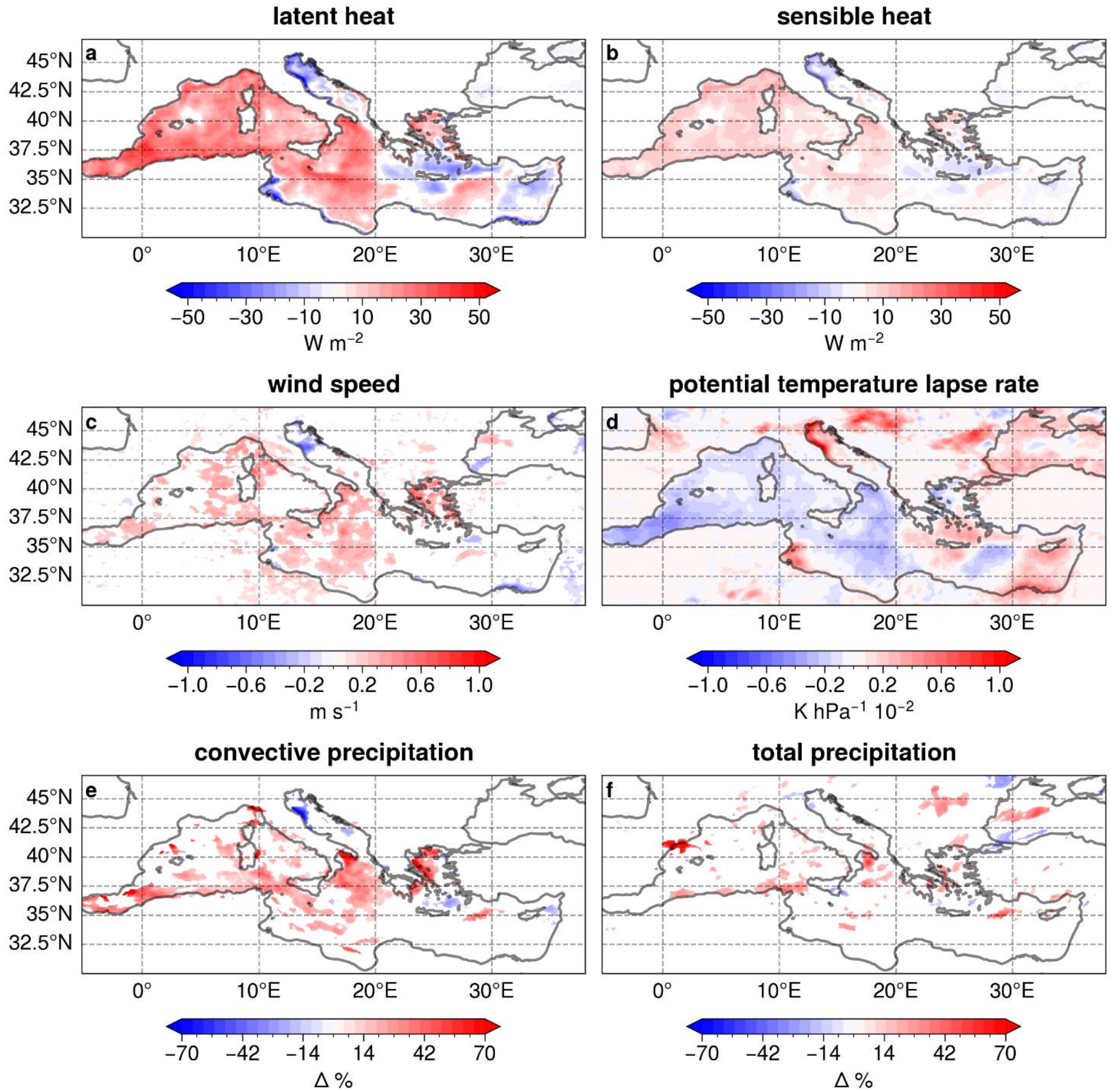

Figure A1: Same as figure 7 but for the 458 common cyclones between CPL and STD.

Line 172-173: does this mean that out of 341, 199 cyclones occur in DJF and SON? Please clarify this.

We have revised lines 172-173 to clarify the seasonal distribution of the same cyclone events as follows (L178-180):

"*With these criteria, a total of 312 cyclones from the 500 most intense (around 62 %) is found in common between CPL and STD, of which 129 occurring in winter (DJF), 110 in spring (MAM), 17 in summer (JJA) and 56 in autumn (SON).*"

Line 176: can you explain or provide a reference for why this field at the specific level was chosen?

In the revised manuscript we have included in the analysis only the temperature and specific humidity fields at both 950 hPa and 850 hPa levels (Fig. 8). These two levels were selected to look at two different atmospheric layers within the PBL, allowing us to assess the impact of the SST differences not only near the surface but throughout the entire PBL. The 950 hPa level (approximately 500 m altitude) is widely used to analyse atmospheric low-level properties (e.g. Raveh-Rubin and Wernli, 2016), while 850 hPa level (approximately 1500 m altitude) is close to the top of the PBL, providing insight into upper part conditions of the PBL (e.g. Fosser et al. 2015).

Fosser, G., S. Khodayar, and P. Berg, 2015: Benefit of convection permitting climate model simulations in the representation of convective precipitation. Clim. Dyn., 44, 45–60, doi:10.1007/s00382-014-2242-1.

Figure 3: It is not clear how this figure was made. Is the percentage calculated from all days in the specific grid cell or is it a percentage from all cyclones in the region? Could you please clarify this? Also, while the spatial variability is well represented in the models, there is still a very large difference in the percentage relative to observations. The authors should discuss this, at least by providing some information about the source of this large difference.

We thank the Reviewer for pointing out the need for clarification. We have revised both figure 3 and section 3.1 to better clarify the results on the comparison between RCMs and ERA5 500 most intense cyclones. We have also provided references on the methodology applied and during the discussion of the results.

Figure 3 is now computed as follow (L247-251):

"*Figure 3 instead shows the maps of cyclone centre densities (CCD; Neu et al. 2013, Flaounas et al. 2018a) for ERA5 and the differences in CDD between ERA5 and RCMs. The CCD is defined as the absolute number of occurrences of the 500 most intense cyclone centres. To highlight the cyclones' area of influence, each centre is represented by a circular area with radius of 1.5 degrees around the tracked minimum SLP point*."

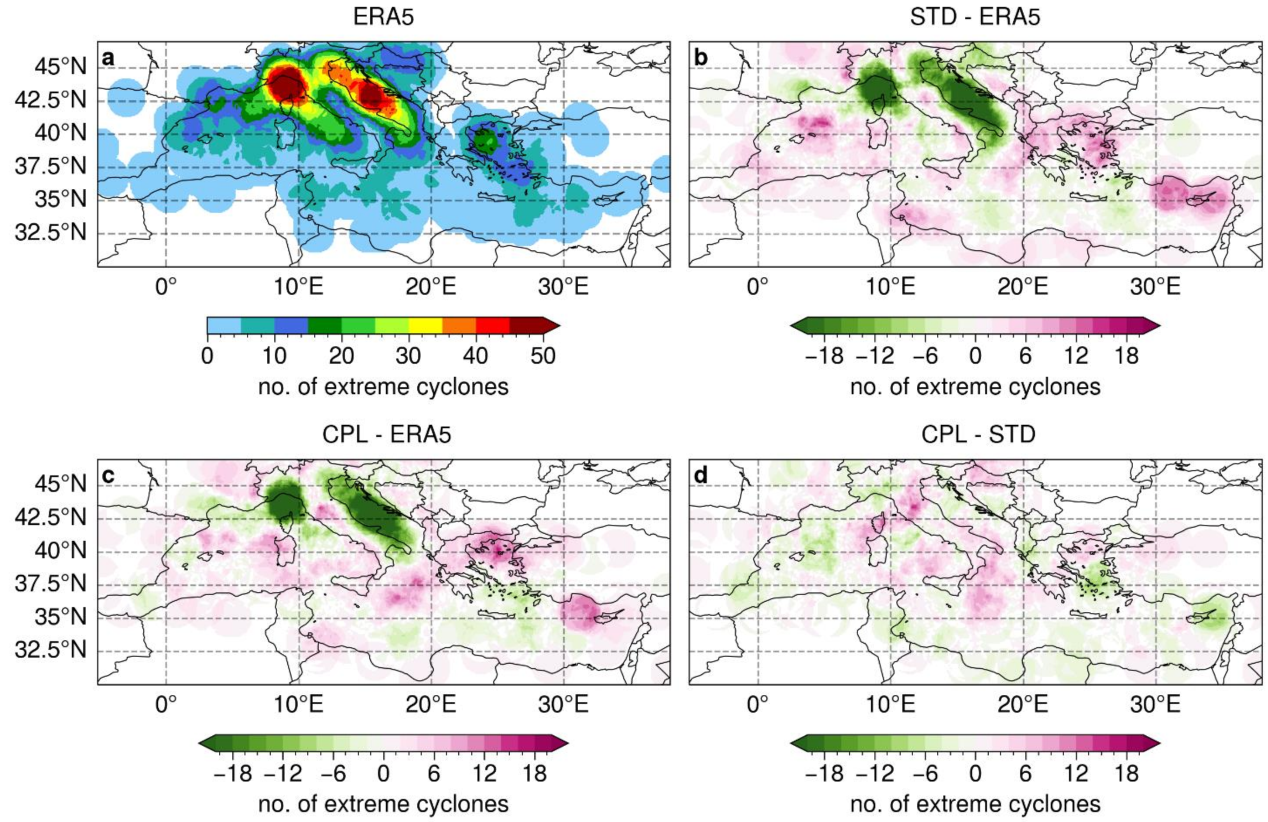

Figure 3: Number of occurrences of cyclone centre densities (CCD) for the 500 most intense cyclones in ERA5 (a), along with CCD differences between STD and ERA5 (b), CPL and ERA5 (c), and CPL and STD (d). To highlight the cyclones' area of influence, each centre is represented by a circular area with radius of 1.5 degrees around the tracked minimum SLP point.

The discussion about the differences between RCMs and ERA5 has been modified as follow (L251-271):

"*Compared to ERA5, both RCMs tend to capture the main regions of frequent cyclogenesis (over the gulf of Genoa, over the Adriatic and Aegean Seas and the marine areas close to Cyprus). This can be expected since the most intense Mediterranean cyclones are formed due to large scale forcing, i.e. the intrusion of upper tropospheric systems as a result of Rossby wave breaking over the Atlantic Ocean (Flaounas et al., 2022). This upper tropospheric forcing is identically introduced to the two simulations through the boundary conditions. However, cyclones seasonality and location also depend on diabatic forcing due to convection within the cyclone systems, as well as on the basin's orography. Both RCMs show a higher occurrence of cyclones in summer and spring (Fig. 2d and Fig. S1 in supplementary) and compared to ERA5, they tend to underestimate the CCD over the Mediterranean Sea while overestimating it over land and over the Aegean and Levantine Sea (Fig. 3b and c). Differences between the two RCMs and ERA5 arise primarily from the different resolution, dynamics and physical parameterisation. These factors influence how the models reproduce key processes, such as. the impact of orography on cyclone dynamics and the role of convection in deepening the cyclones, resulting in local deeper minima of SLP over Mediterranean areas with complex land-sea distribution. Therefore, differences from ERA5 should not be taken purely as a weakness of RCMs, but rather as a result of differences when reproducing atmospheric processes. Indeed, the magnitude of these differences is comparable to the one found in previous studies (Flaounas et al., 2018a; Reale et al., 2022) and thus RCMs should be expected to deviate from reanalysis. In contrast, changes in the SST distribution have a minor impact on the dynamics of the cyclones, leading to small differences between STD and CPL, primarily in the location of cyclone minima over the sea (Fig. 3d). In conclusion, cyclone systems arise from a combination of large-scale processes (external to the cyclone) and small-scale processes (internal to the cyclone). In this context, atmosphere-ocean coupling is expected to have a stronger influence on the physical processes within the cyclone systems, and a rather weaker effect on their formation, distribution, and track characteristics*."

Line 307: "This is explained by the higher Θ gradient of the CPL (Fig. 7d), that makes the PBL less stratified and higher" – what is exactly explained by the higher Θ gradient, and does this gradient is the reason why PBL is less stratified and higher? I would say that this is because of the higher SST, as mentioned in the previous sentence. This reasoning is not clear to me. I would say that, in general, the PBL should be well mixed, and differences between STD and CPL should be pretty small in terms of the temperature gradients inside the PBL. I would attribute the difference only to the SST difference.

We agree with the Reviewer's comment. It is indeed the warmer SST in the CPL model that causes the PBL to be higher and less stratified (stable). We demonstrated this by showing the differences in the potential temperature lapse rate between CPL and STD (Fig. 7d). To note that we have updated the terminology from "Θ gradient" to "potential temperature lapse rate", according to Reviewer 3's comment.

To better clarify the atmospheric processes occurring during the extreme cyclones we have revised the manuscript (L348-352) as follow:

"*In regions with warmer sea, the higher sensible and latent heat fluxes in the CPL model affect, not only surface atmospheric properties, but also modify atmospheric characteristics throughout the entire PBL. In fact, the CPL remains warmer and moister at both 950 hPa and 850 hPa (Fig. 8), and the vertical transport processes of energy are intensified, destabilising the PBL. This is proved by the lower potential temperature lapse rate in the PBL of the CPL model (Fig. 7d), indicating reduced stratification and stability*."

Figure 10: It is unclear which region is considered when calculating the SST difference. Is it one grid cell where maximum cyclone intensity occurred, or is it a regional average?

We thank the Reviewer for highlighting the need for further clarification. The method used to compute Figure 10 was originally explained only in the methodology section. To ensure clarity in Section 3.5 ("Ocean response to extreme cyclones") as well, we have added the following sentence on lines 398-400:

"*For each cyclone the ocean temperature is averaged over a circular area with 1.5° radius, around the minimum SLP tracking point and then averaged over the cyclones considered*."

In addition, we have provided this information in the caption of Figure 10 as follow:

"*SST evolution compared with the SST on the day of the cyclone from five days before to five days after the event for CPL (blue line), STD (green line) and CMEMS MED-Currents reanalysis (orange line), averaged over the same cyclones in DJF (a) and in SON (b). The vertical profiles of the ocean temperature computed as difference between 2 days before and the day of the cyclones (similarly for 2 days after the event) for CPL (blue and red lines) and CMEMS MED-Currents (light blue and orange lines), averaged over the same cyclones in DJF (c) and in SON (d). In each figures the temperature values represent the average over a circular area with 1.5° radius, around the minimum SLP tracking point, and over the cyclones considered. The colour bands represent the confidence interval between +- 1 standard deviation of the mean of the temperature differences*."

Figures: can you explain what the deltas at the top of the panels mean? is it a simple domain average?

The deltas represent the domain-averaged differences only where the values are statistically significant. In the revised manuscript, deltas are shown exclusively for the SST maps (Fig. 5 and 6), and we have updated the figure captions to include an explanation of how the deltas are computed.