general comments

This study uses an air-sea coupled model and a coupled data assimilation method to produce an air-sea coupled reanalysis dataset for the Mediterranean region from 1993 to 2024, and conducts data analysis and comparisons among the results from coupled data assimilation, uncoupled data assimilation, and model simulations without assimilation. The air-sea coupled reanalysis dataset generated by this project is of great significance and value for studying changes in the oceanic and atmospheric fields of the Mediterranean region.

I particularly appreciate Figure 4, Figure 8, and the related discussions in the manuscript. I believe that the discussions on the responses of ocean stratification obtained under different assimilation-scheme designs provide a substantial amount of novel information and contain potential content worthy of further in-depth exploration. They also point to future research directions for air-sea coupled data assimilation in this region.

The methodology of the manuscript is sound and the data are detailed. In my view, only minor revisions are needed for the paper to reach a publishable standard.

specific comments

1）Since this study is primarily based on an air-sea coupled data assimilation method, there is not much methodological innovation. Nevertheless, I still think that Section 2.2 should provide a more detailed description of the assimilation method used, rather than merely citing previously published literature.

For example, L109 states that the ocean component uses a 3DVAR method and cites Storto et al. (2018). However, the work by Storto et al. (2018) is clearly a hybrid four-dimensional variational assimilation method, which makes this statement rather confusing. Therefore, I think it is necessary to provide an explicit equation for the cost function in the manuscript. The mathematical description can be extremely concise, but it should not be entirely absent.

In addition, the paper by Storto et al. (2013) cited in L124 is missing from the reference list.

2) L160: Regarding the MEDREA24 reanalysis dataset, I checked the paper by Escudier et al. (2021). According to Table 1 in that paper, its underlying model also uses NEMO, and its assimilation system also follows Storto et al. (2016). What are the differences between MEDREA24 and the UOA experiment in this study? Clarifying this point would help readers understand the relationship between the differences among reanalysis datasets discussed in Section 3 and the differences among the different experiments discussed in Section 4.

In addition, the reanalysis dataset in Escudier et al. (2021) covers 1987-2019, whereas the time period used in this study is 2000-2020. How exactly was the mismatch between these time periods handled?

3) L193, Table 2: The RMSE values of SST and SSS for MEDREA24 are both 0.331. I believe this is a typographical error, because although I could not find the sea-surface values in the original paper, Figures 5 and 6 of that study show that the temperature and salinity differences averaged over 0-10 m are different.

4) L206: The difference near the Po River mouth is indeed evident. According to the description of the MESMAR system, it includes a hydrology model. Is this difference caused by the hydrology model, or is it an effect of the coupled data assimilation?

5) L237: In the evaluation of the atmospheric component of MESMAR-R, the authors use e-OBS gridded near-surface atmospheric observations. Since the spatial resolution of this gridded observational product is closer to that of ERA reanalysis, the larger RMSE of MESMAR-R does not necessarily imply poorer performance.

MESMAR-R’s 15 km atmospheric model can be viewed as a downscaling of ERA, but the effects introduced by the higher atmospheric resolution under the coupled simulation and assimilation framework are not demonstrated clearly enough.

In addition, according to the cited Haylock et al. reference, the original e-OBS dataset covers 1950-2006, which differs substantially from the temporal coverage of the three datasets evaluated in this study. The manuscript does not explain how E-OBS is used in this evaluation, nor whether the data processing method or spatial resolution after 2006 has changed. These details should be clarified.

6) Around L245, most of the discussion only describes the spatial distribution of differences between MESMAR-R and ERA5/CERRA. However, the atmospheric-field differences themselves are quite interesting and deserve deeper analysis. The RMSEs of T2M, RH2M, and WS10M show large values at very fine spatial scales near the southern European coastlines. In light of the later analysis, these local atmospheric differences also appear to be related to differences in the upper ocean over the northern Mediterranean.

However, the statement at L247 that these results “show MESMAR-R advantages” is not well supported. Except for T2M, the small-scale RMSE differences between MESMAR-R and ERA show both positive and negative values. Moreover, the ERA nudging used in WRF may partly counteract the potential benefits of increased model resolution. The detailed analysis of these fine-scale differences should be valuable, but is currently insufficient. This also affects my interpretation of the later results on coastal air-sea heat fluxes and freshwater fluxes.

It may not be possible to fully identify the causes of these differences without a pure atmospheric experiment for comparison. Nevertheless, the authors should still provide, as far as possible, an explanation of the relevant model-configuration factors or physical mechanisms that may lead to these differences.

Another puzzling point is why PREC includes part of the marine area over the central Mediterranean, whereas the other fields do not. Including T2M and RH2M over the ocean, as much as possible, would also help readers understand the later analysis of net air-sea heat flux and freshwater flux.

Finally, when comparing Figure 3 and Figure 5, the coastal details in both figures are difficult to discern at the current figure size. Figure 5 is barely acceptable, but Figure 3 is much too small. This is especially important because precipitation over densely populated coastal regions is a key variable.

7) L297-300: The differences between COS and COA are quite intriguing. These differences could be analyzed together with the issues raised in Comment 6. At present,

the manuscript only quantifies the COS-COA differences but does not show their spatial distribution, which makes the interpretation unsatisfactory.

8) L335 mentions that the reanalysis products show clear differences in regions with strong tides. This raises a related question: the manuscript does not clearly describe how tides are treated in the production of the reanalysis products. For example, does the ocean model include tides or not? Are all assimilated observations filtered to remove tidal-period signals before assimilation? How are tide-induced residual circulation, tide-induced mixing, and their constraints, if any, represented in the model?

9) L378: The difference between COS and UOS should not be interpreted simply as the difference between coupled and uncoupled simulations, because the atmospheric forcing used in UOS is constrained by atmospheric observations. A more appropriate comparison for isolating the effect of coupling would be to use the atmospheric fields simulated by COS to force the ocean model again in an uncoupled configuration.

Figure 8 also shows that the ocean heat content differs substantially between 2000-2005 and 2015-2020. This large difference likely originates from the reanalysis atmospheric forcing, but may not be present in COS. This issue should be discussed more carefully.

10) L478: The boundary conditions used in MESMAR-R should be described more explicitly. Otherwise, the conclusion here is not very convincing. The manuscript should explain through what pathway the boundary conditions affect the inner domain. For example, does the model enforce basin-wide volume conservation at every integration step? If not, are there any other indirect constraints on the total volume of the basin?

11) In Figure 15, part of the title text is cut off. In addition, the DJF linear precipitation trends over the Mediterranean coastal land areas appear to resemble the spatial pattern of precipitation RMSE in MESMAR-R shown in Figure 3. It would be useful to clarify whether these stronger precipitation trends also exist in low-resolution reanalysis products or observations, or whether they are unique to the high-resolution downscaled simulation.

12) L531, Table 5: This table is potentially misleading, because Figure 15 shows that

the long-term trends in at least precipitation and wind speed have strong regional heterogeneity. For example, the winter precipitation trend over the mountainous areas of northern Italy appears to be much stronger than the summer trend. Therefore, the purpose of presenting this table here is unclear.

It would be more useful to compare how the linear trends differ across subregions between the downscaled coupled simulation and low-resolution reanalysis products or previous studies, rather than emphasizing similarities in domain-mean values.

13) The reanalysis data and simulations in this article cover a relatively long time period, from 1993 to 2024. During this period, the observation system should have made significant progress. In the assessment of interdecadal reanalysis data, the distribution density of observational data should be a key consideration. Over these 30 years, has there been any change in the distribution density of the observations used for assimilation, and have there been any changes in the types and proportions of observation equipment? Because all of these may have an impact on the assimilation results.

14) Similar to 7), one of the points worth studying in the actual sea-air coupled reanalysis data is the differences in the strong coupling objects under unconstrained simulation, semi-constrained simulation, and sea-air constraint conditions. The main factors are the coupling feedbacks between strong winds and the inter-cellular layers of the upper ocean. For example, the formation of deep water caused by strong winds in winter, the sea-air flux events of Mistral/Bora strong winds, and the Mediterranean heat wave events. These representative coupling events are missing in the assessment of this paper.

technical corrections

1) At L134, the abbreviation “MESMAR-R” appears in parentheses for the first time, but the preceding text does not provide its full form. Moreover, the abbreviation MESMAR-R has already appeared earlier in the manuscript, for example at Line 66. In principle, the full form of MESMAR-R should be clearly provided when it first appears in the manuscript.

MESMAR-R is the name of the core dataset produced in this study. Even if the name MESMAR has been defined in previous literature, it is still necessary to define its full form again in this manuscript.

2) What does “REA” mean in Figure 4? Please define this abbreviation clearly in the figure caption or in the main text where it is first used.

3) L321: The explanation of “downward” and “upward” is not sufficiently clear. Does “downward” mean that the domain-integrated heat transport is downward overall, or does it mean that positive values represent downward heat transport, i.e., heat gain by the ocean? This description should be clarified more explicitly.