Revision of "MESMAR v1: A new regional coupled climate model for downscaling, predictability, and data assimilation studies in the Mediterranean region" by Storto et al.(2023)

REVIEWER 1

We thank the reviewer for the encouraging comments and for the suggestions to improve the quality of the manuscript. Below, we provide a point-by-point reply (reviewer in bold, our answer in light font). All coauthors concur with the proposed changes. We refer to the revised version of the manuscript in answering the questions.

Abstract

Line 17: I would report the horizontal resolution of the models described rather than using the adjective "moderate" or "coarser"
Corrected, thanks
Line 21: Are you planning to use the model to produce climate projections? Please specify: The Authors use several times the word "predictions" but never "projections"
We have no plan to use this system for downscaling projections, so we prefer not to add any reference to it. We added "long-range" to predictions, so that readers can understand we don't mention projections.
Line 23: "Intense Mediterranean cyclones "rather than "intense mid-latitude cyclones"
Corrected
Introduction
Line 38: Usually RCM refers to Regional climate models and thus atmosphere not to coupled systems (RCSMs for example, Reale et al., 2022). Please correct that in the text to avoid confusion
We add a sentence to avoid confusion, thanks. We now mention that RCMs are usually intended to be atmospheric models with physics suites targeted to long-term studies, but recently tend to include an interactive ocean model component.
Line 38: Please update Giorgi, (1990) that is a bit old as reference
Updated, thanks
Line 51: "Coupled…unexplored"..see Sevault et al., 2014. The coupled CNRM model uses the spectral nudging.
What we mean by "coupled data assimilation" is the simultaneous use of both ocean and atmosphere observations. The Sevault et al. setup includes spectral nudging in the atmosphere, only, but not in the ocean, so we do not consider it as an example of regional coupled data assimilation. These, to our knowledge, are very few, and anyway missing in any Mediterranean Sea system. We prefer to keep the sentence as it is.
The Authors use several times the word "predictions" but never "projections". Are you planning to use your model also to produce projections? Please specify
Please see the answer above. We prefer not to specify what we do not want to do (no plan to use it for projections, but you never know).
Line 72: Be cautious since hurricanes are (from some points of view) very different from medicanes. See Flaounas et al., 2022. Please reformulate this sentence
Modified, thanks.
In the introduction it is missing a clear description of the present state concerning the coupled models in the MED region and at which extent the new modeling system described by the Authors is different or eventually represent an advancement with respect the preexisting modeling tools.
We have added a sentence on the specificities and advancement of our system. We also expanded the discussion about this important point in the Summary and Discussion section, also to respond to

the last comment of the reviewer. In the Intro, we added a sentence about the uniqueness of our modeling system.

2.Earth system model configuration

Line 94: What do you mean with stationary geophysical fields? Is the topography a geophysical field? Please explain
Yes, plus many more. We use now the term "Geographical Static Data" as given by WRF (https://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html) and detail what they are.

Line 101: Do you refer to the width of the sponge layer? Please specify
Yes, this is now specified.

Line 110-115: As far as I understand you are using shortwave radiation as forcing that attenuates along the water column according to a water attenuation coefficient plus the chl-a concentration: since you are using the satellite chl-a (first 10 m as far as I remember) how do you quantify the attenuation effect led by chl-a below 10 m?
This is detailed in the paper cited: 2d chlorophyll fields are combined with a depth-dependent function to provide 3D attenuation factors. We modified the text to make this more clear.

Line 117-120: Are you using an open boundary in the Atlantic? Are you applying also the sea surface height at boundary?
Yes, the description refers to an open boundary. The Flather scheme corrects simultaneously the inner barotropic velocities and ssh according to the external gravity waves. We modified the sentence accordingly.

Line 126-128: Is the Nile missing from the numerical settings? It is not mentioned in the text.
Not sure to understand the question: the HD model implements, by construction, a European setup that includes all river basins for the European seas (see fig1), including the Nile. We only mention the Dardanelles because they are a strait and not a river, so HD cannot resolve the freshwater incoming from there

Line 138: Is "25" resolution dependent? Please explain
This has been tested only in this configuration, so we cannot answer if the choice of 5x5=25 gridpoints depends on the resolution of MESMAR or not. So, we prefer to keep the sentence unchanged (we anyway corrected 25 to 5x5 to make it clear that it is a rectangle centered on a gridpoint).

Line 143: Do you mean that all the model components exchange field every 30 min? please specify
Yes. We changed it to "for all exchanged fields" to make it more clear.

Line 148-150: Not very clear. Please reformulate. As far as I understand WRF transfers surface and subsurface runoff to HD that is remapped on the ocean grid and passed to NEMO after HD runs. Does HD pass only the river discharge to NEMO? Do you need to locate the river mouths on the ocean model domain?
HD implements the remapping onto NEMO directly (aware of the NEMO mask, of course), so it passes to NEMO the river runoff directly on the actual NEMO grid. We modified the sentence to make clear HD remaps the runoff onto the NEMO river mouths (OASIS does not perform any interpolation in this case).

Line 152-157: Why do not use CMEMS MYOCEAN reanalysis specifically tuned for the MED sea instead of GLORYS? Moreover, there are no information about the spin up of your model for pot temperature and salinity. Did you perform the spin up at least for the long run? I would also include additional information concerning the computational performances of the model (number of cores, computational times etc.). It would be useful having a table summarizing the main settings of the model at least in the most important experiments.

GLORYS12 is preferred to the CMEMS MED reanalysis because the resolution is closer to our model implementation (~1/12 for both systems). We added this and other information about the computational resources in the text.

3.Sensitivity experiments

3.1 Impact of the interactive river discharge

I do understand the idea of the Authors to show the importance of river online on the simulate salinity. However, I think that 2 years of runs are too small to assess the importance of river inflow on salinity (in particular along the water column where the signal need sometimes to penetrate) in absence of information about the spin up. I would change Fig.3 and Fig 4 adding the comparison with observations (instead of the comparison between the two configurations) as You did in Figure 6. Moreover, why did you use EN4 instead CMEMS or JRA55 instead of Ludwig et al., 2009.
Sorry, but the suggestions are not clear: Figure 4 already shows the comparison with observations as suggested by the Reviewer. Not clear what is meant for CMEMS here (reanalyses, observations, objective analyses?), in any case, EN4 is a state-of-the-art profile and objective analysis product that we use for validation (in Fig 3 as SSS analyses, in Fig 4 as raw observation data). There is no reason why we should use CMEMS (and why not the NOAA dataset then? Or the JMA one?).
We clarified this point in the text. JRA55-do is a daily reprocessed dataset, while Ludwig 2009 is a climatology, so it is more appropriate as verifying dataset for our 2-year experiment. We modified the text to stress that the experimental period is short, so the results are only indicative of the potential of HD coupling.

3.2 Nemo vertical physics

There are no information about the length of tests. Moreover, there is no quantification of biases (smaller etc is too generic). Please quantify the biases
Thanks, we added this information in section 3.2

3.3 WRF configuration

Why have your tests run only for the period 1993-2021 (also in the reference simulation) instead of covering the entire ERA5-ORAS5 period? Please specify
This is a standard period that includes altimetry, and it is used, for instance, for CMEMS reanalysis dissemination, C3S seasonal prediction forecasts, and more. Of course, the choice is arbitrary but consistent with other datasets. We do not think we need to justify the choice of the experimental period, so we leave the text as it is.

What do not you investigate also the behavior of wind and precipitation on the ocean domain since their importance for E-P and mixing?
The impact was found not significant on the ocean bias, we added a comment on this at the end of the section.

4 Reference simulation

Line 235: I would add also Soto-Navarro et al., (2020) and Reale et al., (2022) that addressed this issue.
**Soto-Navarro, J., Jordá, G., Amores, A., Cabos, W., Somot, S., Sevault, F., et al. (2020). Evolution of Mediterranean Sea water properties under climate change scenarios in the Med-CORDEX ensemble. Clim. Dyn. 54, 2135–2165. doi: 10.1007/s00382-019-05105-4**

Reale, M., Cossarini, G., Lazzari, P., Lovato, T., Bolzon, G., Masina, S., ... & Salon, S. (2022). Acidification, deoxygenation, and nutrient and biomass declines in a warming Mediterranean Sea.Biogeosciences, 19(17), 4035-4065.
Added, thank you.

Being a long term simulation, it would be interested having the analysis of long term timeseries of temperature and salinity with respect to observational datasets to have an idea about the behavior of the thermohaline properties at different depths.
Thanks for the suggestion. We report below a figure showing the Mean, Bias, and RMSE of temperature in observation space (profile data). We include a comment on this in the text, but we think the figure is not very informative to be included.
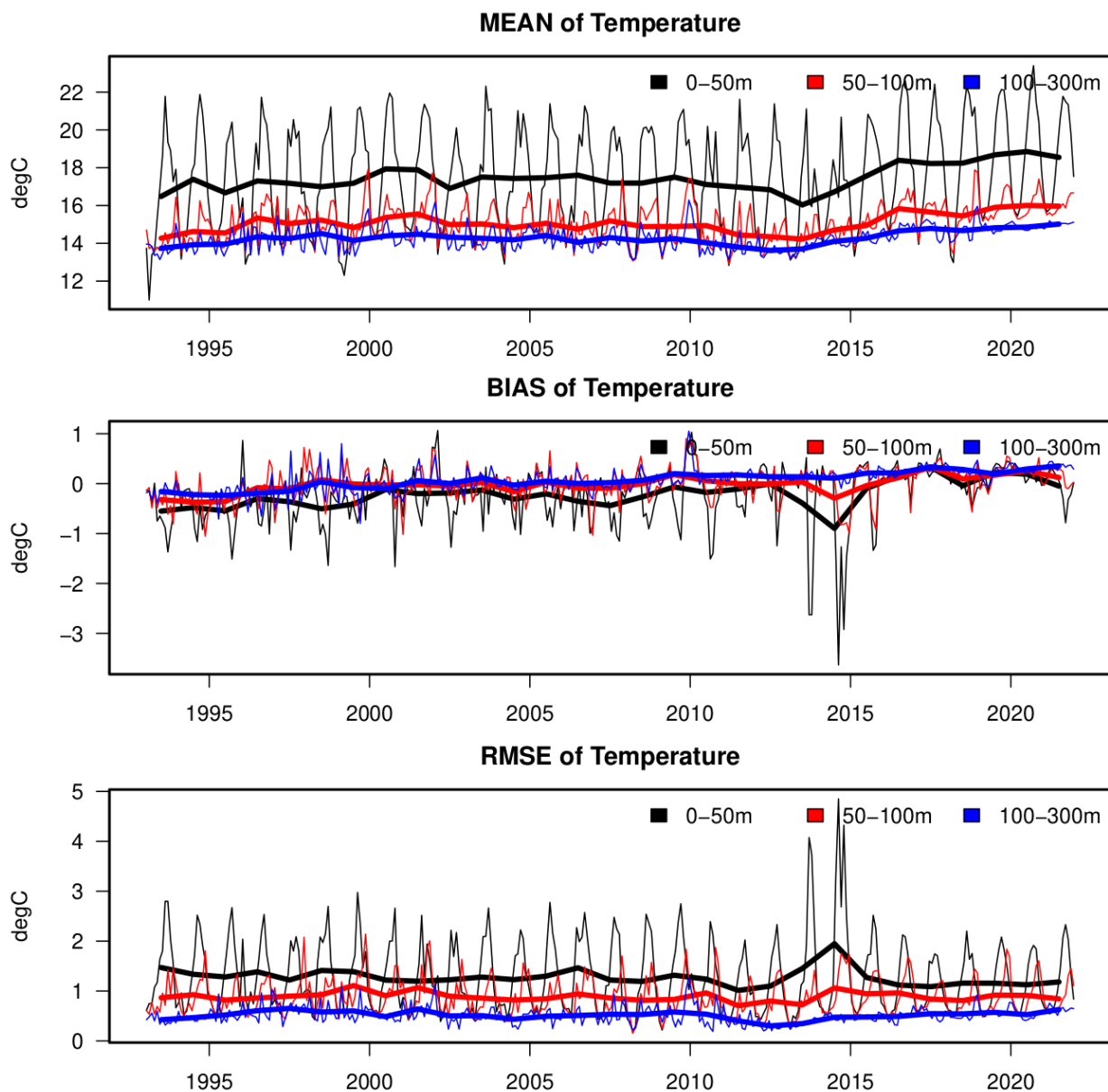
Mean, bias and RMSE of T in observation space for the entire basin

It would be important also to assess the source of the error in the total net heat flux. Is it an underestimation of the net shortwave? Latent or sensible heat fluxes? Please think to include this analysis

This was already stated, at the end of section 4, where we discussed the difference of the individual air-sea heat flux components.

Figure 11 Why is the area outside Gibraltar colored in the upper panels and not in the bottom panels?
We have now masked out areas outside Gibraltar in all panels, for consistency between panels, thanks.

5 Data assimilation

I found the data assimilation an interesting new advancement in the coupled model. However, I do not understand the reason for relaxing at the surface when already both atmosphere and ocean assimilate at high frequency data to correct errors in the simulated field. Could you please explain better that?
As this system is intended also for long runs and reanalyses, SST relaxation provides a temporally consistent way to ingest sea surface data without spurious variations linked to the observational sampling. This is a very standard approach used in most state-of-the-art ocean reanalyses (ORAS5, CGLORS, etc.). We shortly explained this in the text. Note, there is no redundancy of data, as the variational DA assimilates profiles, and the nudging ingests sea surface temperature data (from satellite, referenced to drifter SST data).

Moreover, I would ask the Authors to better discuss why with fully data assimilation active (and high frequency of assimilation) the improvements in the biases is relatively small. Low quality of input data? Please infer on that.
We do not agree with the reviewer. The impact of DA is significant. Looking at Figure 13 bottom middle panel, for instance, we see a large impact of DA. Table 3 quantifies the impact. This is reflected in all ocean skill scores when ocean DA is switched on.

5 Mediterranean hurricanes How do you track and reconstruct the temporal evolution of the systems analyzed in the manuscript? Please infer on that.
Tracks correspond to the surface pressure minimum. We added this in the main text.

6 Discussion and conclusions There is a lack of the comparison with respect to previous modeling systems. Why should a potential user use your modeling tool instead of another one? Do you expect that increasing the horizontal resolution should improve the performances of your model or should make it slower?
Thanks for pointing this issue out. We have expanded the first paragraph of the Summary section to address this point, and the one mentioned above by Reviewer 1 in the Introduction, and pointed out that the resolution increase is intended for short experiments (e.g. mimicking short-range forecasts).