

Reviewer #2's comment:

review of "The added value of Med-CORDEX Coupled High-Resolution Regional Climate Models in representing Sea Surface Temperature and Marine Heatwaves in the Mediterranean Sea" by De Rovere et al

This study applies a series of diagnostics to the daily SST field from ~10 MedCordex simulations and their parent GCM, so as to diagnose when and where RCSMs adds value. The key result is that except for MHW duration and certain aspects of the PDF of SST variability where RCSMs clearly adds value, there is little evidence for systematic improvements across models away from coastlines or regions with topography. Further, rather than horizontal resolution adding substantial value, the authors seem to highlight vertical resolution in the mixed layer as a potentially more important model setting, though in an ensemble of opportunity as used here it is difficult to isolate processes.

This is a solid study that is nearly ready to be accepted. It also highlights the limitations of dynamical downscaling- this reviewer would have hoped that all of this CPU time and effort would have helped matters more. But there is one major issue: I expect a more careful assessment of the pros and cons of the diagnostics used. Specifically, they aren't using process-based diagnostics that attempt to follow the processes that drive SST variability, rather just examine SST itself. For example, we know that MHWs (and SST variability more generally) are driven by anomalies in evaporation, which in turn is driven by weak wind speed and enhanced atmospheric near-surface humidity, and also by enhanced surface downwelling radiation associated with fewer clouds (Bonino et al, already cited; Garfinkel et al 2026; GRL). I think the assessment of SST variability and MHWs might be more informative if the authors analyzed how these processes are represented in GCM vs. RCSMs. I would expect that evaporation would be better handled with enhanced horizontal resolution as fronts are now resolved, and the representation of clouds will also be very different in a GCM vs. RCSMs especially if the RCM turns off convection. But this needs to be examined closely. The paper currently discussed how oceanic processes might differ in a GCM vs. a RCSMs, but atmospheric processes matter too (Bonino et al and Garfinkel et al 2026) and they are completely ignored.

To be specific, I don't think the current study needs to include the analysis I'm suggesting above. Rather, a new paragraph in the Discussion would suffice. This is mainly because at least four additional fields would be needed - near-surface wind speed, surface downwelling shortwave, cloud fraction, specific humidity in the atmospheric surface layer - and I don't know if this data is even available. But I certainly think that more could be done with the GCM and RCSMs used here to better understand why RCSMs don't help as much as this reviewer would have hoped. If this data is not available, then this could be a recommendation for the data request in future phases of the Cordex project. More generally, I suggest adding a sentence to the Discussion that notes the limitations of dynamical downscaling; this reviewer would have hoped that all of this CPU time and effort would help matters more.

specific comments:

1. Abstract: The role of vertical resolution in the upper ocean seems to be more important than that of horizontal resolution (based on the Discussion). This would seem to me to be an important point worth highlighting in the abstract.

2 line 82: "We focus on such period to maximizes overlap with" needs to be rewritten

3. Figure 1 and similar: the red and green shading are not colorblind friendly. Please consider changing these colors to ones with more contrast. Also, why doesn't the orange overcompensation shading extend all the way to the green line for perfect error correction?

line 179, 202, 204, 229,230,272: My first language isn't English either, but I think "impoverished" is used incorrectly. Do you mean degraded?

Answer:

We sincerely thank the reviewer for the careful reading of our manuscript and for the constructive comments, which we believe have improved the quality of the paper. We are pleased that the overall quality of the study was found to be solid. Below we address each point raised by the reviewer.

In the following, the points and questions raised by the reviewer are reproduced in bold, text from the first version of the manuscript is shown in red, and planned changes and additions are shown in blue. Planned changes and additions will be implemented in the revised version of the manuscript, which will be submitted after the closing of the public discussion. All lines refer to the first version of the manuscript.

Main point: "The authors currently discussed how oceanic processes might differ in a GCM vs. a RCMs, but atmospheric processes matter too (Bonino et al., 2025; and Garfinkel et al. 2026) and they are completely ignored."

We thank the reviewer for this insightful comment and for the suggested reference. We agree that a process-based assessment of how atmospheric drivers of MHWs, in particular wind speed, evaporation, cloud fraction, and near-surface humidity, are represented across the Med-CORDEX ensemble would provide a more complete picture of where and why RCMs add value. We acknowledge this as a limitation of the current study. We plan to add a new paragraph in the Discussion (in between lines 399 and 400) explicitly acknowledging this limitation and recommending it as a priority for future work:

"The present study focuses exclusively on SST and derived MHW properties, without examining the atmospheric drivers that govern SST variability and MHW onset. Recent work has shown that Mediterranean MHWs are primarily driven by anomalies in surface heat fluxes, in particular reduced latent heat loss associated with weak winds, enhanced near-surface humidity, and reduced cloud cover associated with enhanced surface downwelling radiation (Bonino et al., 2025; Garfinkel et al., 2026). RCSMs improve the representation of atmospheric fronts and mesoscale wind systems (Herrmann et al., 2011; Obermann et al., 2018), which is expected to translate into better representation of surface latent heat fluxes and evaporation (Hamon et al., 2016; Herrman et al., 2016), key drivers of MHW onset and decay. However, a process-based assessment of how such atmospheric processes influence MHWs across the Med-CORDEX ensemble was beyond the scope of this study but represents an important direction for future work."

Specific comments:

1. Abstract: The role of vertical resolution in the upper ocean seems to be more important than that of horizontal resolution (based on the Discussion). This would seem to me to be an important point worth highlighting in the abstract.

Thank you for pointing this out. We plan to add the following sentence at the end of the abstract:

"In particular, within the Med-CORDEX ensemble, inter-model differences in MHW added value appear to be correlated with improvements in vertical resolution."

2. line 82: "We focus on such period to maximizes overlap with" needs to be rewritten

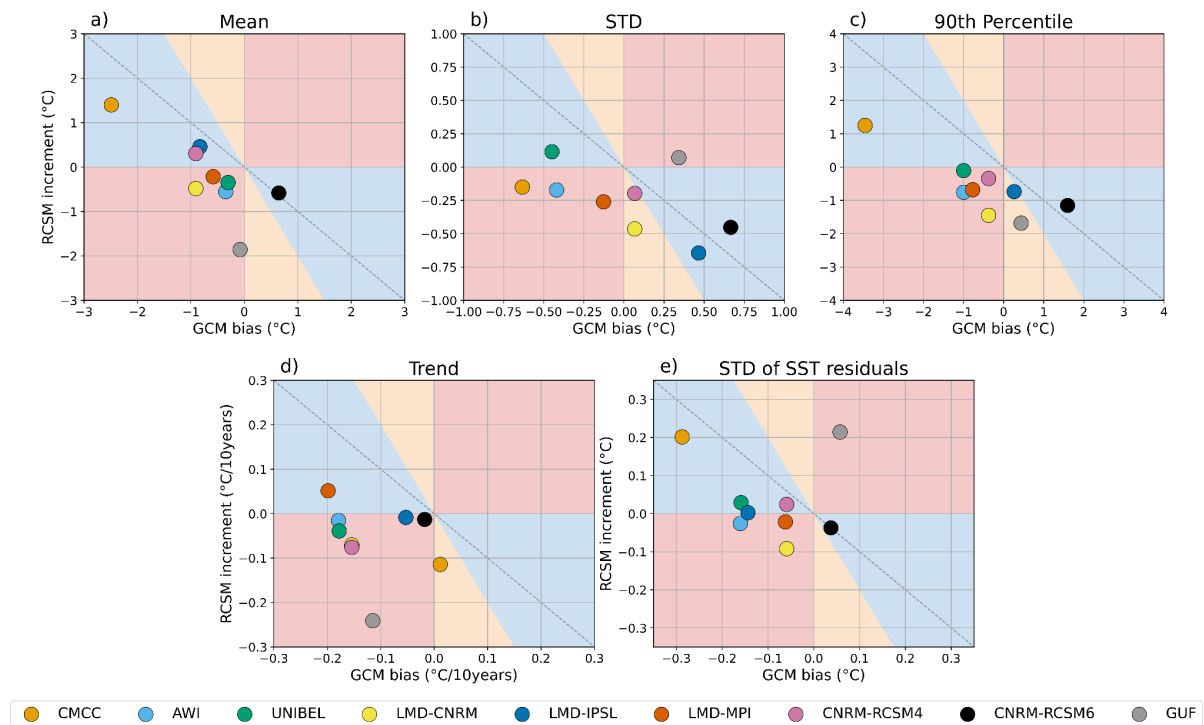
We will rewrite the sentence to:

"We select this period to maximize overlap with high-resolution OBS, improving the statistical robustness of the results while minimizing the use of data from future scenario simulations."

3. Figure 1 and similar: the red and green shading are not colorblind friendly. Please consider changing these colors to ones with more contrast. Also, why doesn't the orange overcompensation shading extend all the way to the green line for perfect error correction?

Thank you for raising this point. We have modified all scatterplots to use colorblind-friendly colors for both the background shading and the dots in Figures 1, 4, S4, and S12.

The following is the updated version of Figure 1 for reference.



We will update the caption as follows:

“Figure 1. RCSM increment versus GCM bias for area-averaged SST (a) mean, (b) STD, (c) 90th percentile, (d) trend, (e) STD of SST residuals over 1982-2020. The grey line indicates perfect error correction. Blue, red, and orange shading denote skillful error reduction, deterioration, and overcompensation (RCSM error opposite in sign but larger in magnitude than the GCM error), respectively. For the definition of the Mediterranean domain see Fig. 2.”

Regarding the orange overcompensation shading, this represents the area where the RCSM bias is larger in magnitude than the GCM bias but opposite in sign. This region is bounded by the lines $\text{RCSM increment} = -0.5 \times \text{GCM bias}$ and $\text{GCM bias} = 0$. In the new figures, the blue region extending from the orange shading to the line of perfect error correction identifies the area where the RCSM bias has opposite sign compared to the GCM bias but is smaller in magnitude.

line 179, 202, 204, 229,230,272: My first language isn't English either, but I think "impoverished" is used incorrectly. Do you mean degraded?

Yes, we will replace all “**impoverished/impoverishment**” entries with “**degraded/degradation**”. Thank you for the suggestion.