We thank Reviewer #1 for the constructive comments and suggestions, which greatly help to improve the quality of our manuscript. We have made revisions and replied to all the comments. Please find the point-by-point responses to the comments below. Our responses are shown in "Blue" and the changes in the manuscript are shown in "Red". The line numbers correspond to those in the clean version of our revised manuscript.

Response to the comments from Reviewer #1

General Comment:

The manuscript describes the implementation of a new weakly coupled ocean data assimilation in the E3SMv2. The authors show that using this assimilation method helps reduce temperature and salinity bias and RMSE in general in both surface and deeper ocean layers. The results of the manuscript are promising and will be valuable for the community. It is overall well-written and easy to follow. That being said I believe that more work must be done on this manuscript before being published. A lot of things need to be clarified or added to improve the clarity and robustness of the results.

As a general comment here, before going into details, the authors need to be more quantitative, in the text and in the abstract, when describing results in addition to saying increase/decrease or cold/warm biases please indicate some values. Also, units of temperature and salinity are missing on every figure, please add units in both figures and captions as otherwise, it makes things hard to understand, i.e. figure 10 see comment below.

One other main concern here is about section 3.6 on the influence of ocean data assimilation on the regional climate over land. Although this is an interesting topic, it is most likely a manuscript in itself. The results presented in this section are highly preliminary, no details on observations (definitely not enough to only briefly mention it in the Figure 11's caption) or methodology used are given. Very little can be said with certainty on the influence of ocean data assimilation on regional climate over land with only the figure and analysis presented (fig 11). This needs a much more rigorous analysis to be able to draw robust conclusions. In my opinion, this section should be removed and can't be published as is in this manuscript.

Response:

Thank you very much for taking the time to review our manuscript and providing us with very useful comments. We have revised the abstract (L28-31), results (L373-376, L379-380, and L381-383), and conclusion (L404-408) sections to include numerical values that quantify improvements in temperature and salinity biases. We have also added units to all relevant figures and captions (Figs 3-5, 8-9) and explicitly updated Fig. 10 to include missing units for better clarity.

After careful consideration, we have decided to remove section 3.6 and Figure 11 from the manuscript. We agree that the current results are preliminary and require further analysis in future work. The removal of this section 3.6 has been reflected in the revised manuscript, and all references to this section have been updated accordingly.

L28-31: For sea surface temperature, cold biases in the North Pacific and North Atlantic are diminished by about 1-2 °C, and warm biases in the Southern Ocean are corrected by approximately 1.5-2.5 °C. In terms of salinity, improvements are observed with bias reductions of about 0.5-1 psu in the North Atlantic and North Pacific and up to 1.5 psu in parts of the Southern Ocean.

L373-376: In contrast, these SST biases found in CTRL are substantially reduced by ASSIM (Fig. 10b), with cold biases in the North Pacific and North Atlantic diminished by approximately 1-2 °C, and warm biases in the Southern Ocean corrected by about 1.5-2.5 °C.

L379-380: Notably, in the Mediterranean Sea, CTRL exhibits a large positive salinity bias exceeding 2.5 psu.

L381-383: Notable improvements are observed in the North Atlantic and North Pacific, where salinity biases are reduced by approximately 0.5-1 psu, and in parts of the Southern Ocean, where reductions reach up to 1.5 psu.

L404-408: ASSIM substantially mitigates model biases for SST and SSS observed in CTRL, particularly reducing cold biases in the North Pacific and North Atlantic by approximately 1-2 °C, correcting warm biases in the Southern Ocean by about 1.5-2.5 °C, and significantly increasing salinity estimates to reduce the model fresh biases by approximately 0.5-1 psu in the North Atlantic and North Pacific, and up to 1.5 psu in parts of the Southern Ocean.

Comment#1:

Line 114-116: As the manuscript focuses on ocean, more info is needed on the ocean model set-up here. Does the ocean model has the same horizontal resolution as the atmosphere? how many vertical layers, vertical resolution? What is the vertical mixing scheme used?

Response:

Thank you for your constructive comment to clarify the ocean model configuration in our manuscript. The ocean model in E3SMv2 employs the Model for Prediction Across Scales-Ocean (MPAS-O), which has a horizontal resolution of ~60 km in the midlatitudes and ~30 km at the equator and poles, differing from the atmospheric model's resolution of 110 km. MPAS-O is configured with 60 vertical layers, with finer resolution (~10 m) near the surface and coarser resolution (~200 m) at depth. The vertical mixing scheme employed is the K-profile parameterization, as described by Van Roekel et al. (2018).

Based on your suggestion, we have revised the manuscript to provide a more detailed description of the ocean model configuration (L120-124).

L120-124: MPAS-O operates at a horizontal resolution of ~60 km in the midlatitudes and ~30 km at the equator and poles, differing from the atmospheric model's resolution of 110 km. It is configured with 60 vertical layers, with finer resolution (~10 m) near the surface and coarser resolution (~200 m) at depth. The vertical mixing scheme employed is the K-profile parameterization, as described by Van Roekel et al. (2018).

Comment#2:

On section 2.2: Just by reading this, for an external reader, it is not very clear what is this product (EN4.2.1) and thus what is assimilated. Those are profiles with spatiotemporal variability (right?), what are the typical depths where observations are available? Typically, what are the regions where we have a lot or few observations, and thus can/should we expect improvement in these areas or not? In summary, a better description of this product here would help us understand better the results presented after.

Response:

Thank you for pointing out the need for more clarity regarding the EN4.2.1 dataset. The EN4.2.1 product is developed based on quality-controlled ocean temperature and salinity profiles from four input sources: Argo, ASBO (Arctic Synoptic Basin Wide Oceanography), GTSPP (Global Temperature and Salinity Profile Program), and WOD09 (World Ocean Database). The EN4.2.1 dataset includes observations from a wide range of profiling instruments, such as Argo floats, expendable bathythermographs (XBTs), and mechanical bathythermographs (MBTs) (Chen et al., 2020).

According to Good et al. (2013), observations in EN4.2.1 are most abundant in the upper 100 meters, with vertical resolution refined to ~1 m in the top 100 m. Spatially, data density is high in regions such as the North Atlantic and western Pacific but decreases significantly in high-latitude and deep ocean regions. Areas with higher observational density, such as the upper North Atlantic, are expected to show greater improvement in the assimilation results, while regions with sparse observations may exhibit limited improvements.

Following your advice, we have revised the manuscript to include a more detailed description of the EN4.2.1 dataset, emphasizing the spatiotemporal distribution of observations and their implications for data assimilation (L135-146).

L135-146: Produced by the Met Office Hadley Centre, the EN4.2.1 product is developed based on quality-controlled ocean temperature and salinity profiles from four input sources: Argo, ASBO (Arctic Synoptic Basin Wide Oceanography), GTSPP (Global Temperature and Salinity Profile Program), and WOD09 (World Ocean Database) (Good et al., 2013). The EN4.2.1 dataset includes observations from a wide range of profiling instruments, such as Argo floats, expendable bathythermographs (XBTs), and mechanical bathythermographs (MBTs) (Chen et al., 2020). According to Good et al. (2013), observations in EN4.2.1 are most abundant in the upper 100 meters, with vertical resolution refined to ~1 m in the top 100 m. Spatially, data density is high in regions such as the North Atlantic and western Pacific but decreases significantly in high-latitude and deep ocean regions. This distribution in data availability influences the assimilation results. Areas with denser observational coverage, such as the upper North Atlantic, are expected to show greater improvements through assimilation, while regions with sparse observations may exhibit limited improvements.

Comment#3:
Line 150: “significantly reducing”: by how much the computational resources are reduced?

Response:
To clarify, the computational efficiency of the DRP-4DVar system used in this study has been demonstrated in previous studies. Based on Zhu et al. (2022), the DRP-4DVar system reduces computational time by approximately 50% compared to traditional 4DVar systems. For example, in a 6-hour assimilation window, the 4DVar system required 25 minutes, while the DRP-4DVar system only took 13 minutes due to its efficient computation.

In response to this comment, we have revised this sentence (L162-163) for clarification in the revised manuscript.

L162-163: Zhu et al. (2022) demonstrated that the DRP-4DVar method significantly reduces computational time by approximately 50% compared to traditional 4DVar systems.

Comment#4:
Line 207: what do you mean by “observed external forcing”? what is this?

Response:
The “observed external forcing” refers to external factors influencing the climate system, such as solar radiation and greenhouse gas and aerosol concentrations. These transient historical external forcings are prescribed following the CMIP6 protocol (Eyring et al., 2016). These observed external forcings directly influence the atmospheric component of the model and subsequently influence other components (e.g., land and ocean) through their coupling with the atmosphere.

We have revised this sentence (L218-222) to include a more precise explanation of “observed external forcing” for better clarity.

L218-222: driven exclusively by observed external forcings (e.g., solar radiation and greenhouse gas and aerosol concentrations). The observed external forcings, prescribed according to the CMIP6 protocol (Eyring et al., 2016), directly influence the atmospheric component and subsequently affect other components (e.g., land and ocean) through their coupling with the atmosphere.

Comment#5:
Line 211: “across sixty ocean layers”. Is this all the model ocean layers? if not what’s the depth of the 60 layer and why only 60. Please clarify this.

Response:
Yes, the referenced "sixty ocean layers" corresponds to all the vertical layers in the ocean component of the E3SMv2 model. We have revised this sentence (L226-227) for better clarity.

L226-227: across all sixty ocean layers spanning the entire ocean depth

Comment#6:

Line 237-238: what is the reason for the initial jump in the cost function reduction from -12% to -4% ?

Response:

The observed initial jump from -12% to -4% can be attributed to the rapid adjustments of the model during the first two years of assimilation. At the start of the assimilation cycle, the model state undergoes rapid adjustments to align with the reanalysis data, leading to a sharp reduction in the cost function. As the assimilation progresses, subsequent iterations refine these adjustments, resulting in a slower rate of reduction

To clarify this point, we have provided a brief explanation of this phenomenon (L263-266).

L263-266: The initial sharp reduction rate of the cost function reflects the rapid adjustments made by the model to align with the reanalysis data. As the assimilation progresses, subsequent iterations refine these adjustments, resulting in a slower rate of reduction.

Comment#7:

Line 246: “nine ocean layers”: why did you choose to show specifically these 9 layers? Is 85m the last layer where observations are assimilated ? if not why not showing anything below? This need to be clarified.

Moreover, is it necessary to show these nine as they relatively show the same results, maybe you could only show 5m, 45 and 85m ..? Reading further, you’re showing profiles up to 1000m so why not showing maps of the deeper ocean here as well? 85m is not ‘deep ocean’, depending on region and/or season this is still in the ocean mixed layer.

On that note, it could be interesting to show seasonal maps as well. Is there any seasonal variability on these results, i.e. if maybe there are fewer observations during winter months does the ASSIM still perform better?

Response:

Thank you for your insightful comments. The original nine layers were simply chosen as the first nine layers of the ocean model, ranging from 5 m to 85 m. However, we recognize that this approach limited the depth representation and did not fully capture the assimilation impacts across deeper layers. Based on your suggestion, we have revised the depth of the nine layers to span a broader depth range. The updated layers are at depths of 5 m, 45 m, 85 m, 135 m, 327 m, 528 m, 708 m, 879 m, and 1106 m. We have revised Figure 4 to include the nine newly selected ocean layers from 5 m to 1106 m and updated the description of assimilation performance (L285-287) in the deeper ocean layers (879 m and 1106 m).

L285-287: In the deeper layers, the assimilation still shows notable improvements in regions such as the North Pacific and parts of the Southern Ocean, though with more pronounced degradation observed in the equatorial Atlantic and parts of the Indian Ocean.

We have also added supplementary figures (Figs. A1 & A2) and the corresponding sentences (L283-285) to show the seasonal performance of RMSE differences for summer (Fig. A1) and winter (Fig. A2) across the same nine depths. In the upper ocean layers, we observe better RMSE performance during winter compared to summer in some regions, such as the tropical Pacific.

L283-285: In the upper ocean layers, RMSE performance is better during winter compared to summer in some regions, such as the tropical Pacific (Figs. A1 & A2).

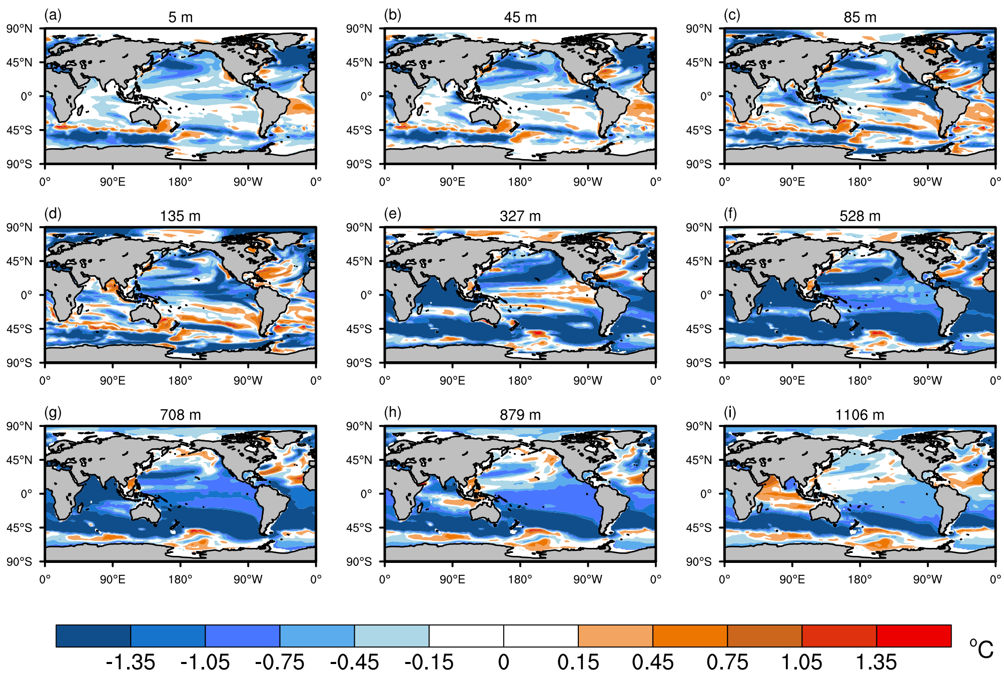

Figure 4. Spatial patterns of root mean square error (RMSE) differences in ocean temperature (unit: °C) between ASSIM and CTRL across nine ocean layers from 1950 to 2021. The RMSE differences are shown for nine different ocean depths: (a) 5 m, (b) 45 m, (c) 85 m, (d) 135 m, (e) 327 m, (f) 528 m, (g) 708 m, (h) 879 m, and (i) 1106 m.

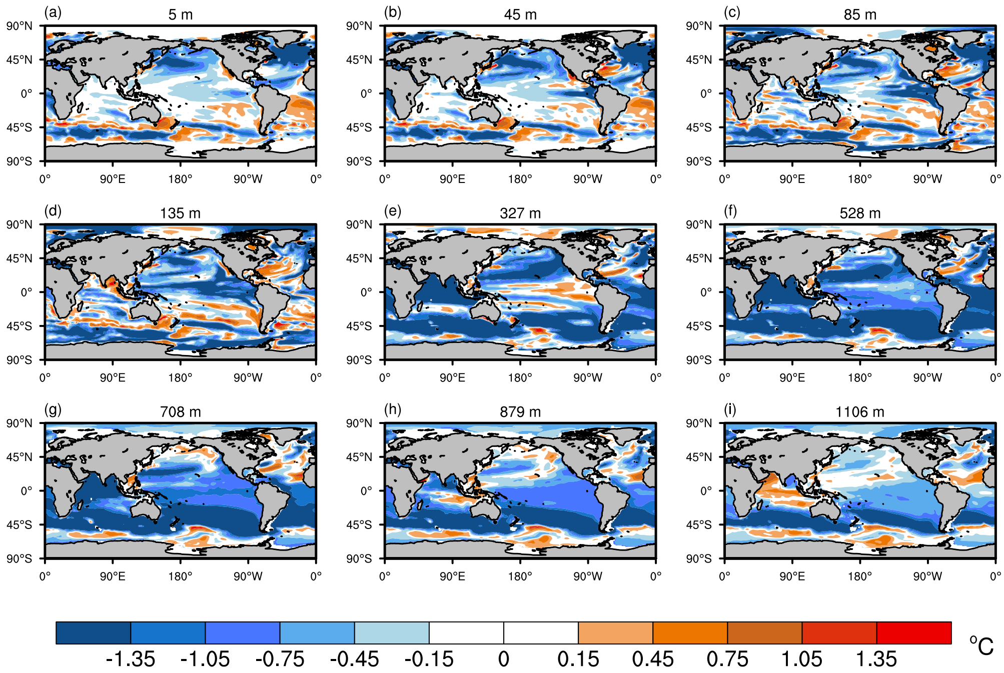

Figure A1. Similar to Figure 4 but during summer.

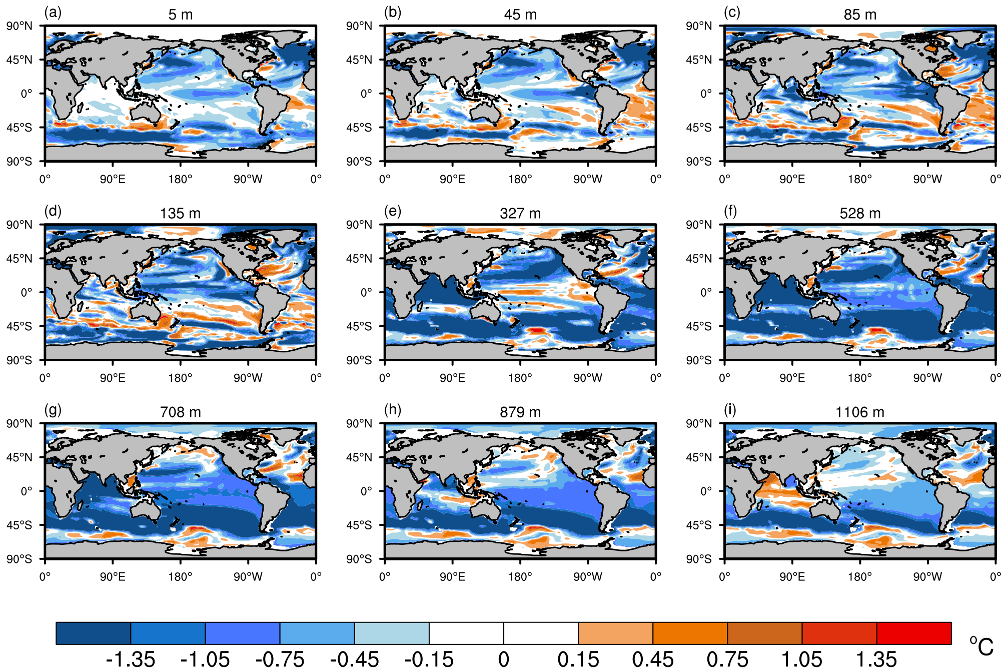

Figure A2. Similar to Figure 4 but during winter.

Comment#8:

On Figure 4 : Is there any ocean current-related impact? As it seems that the RMSE is somewhat increased near strong ocean currents (i.e. Gulf Stream separation, ACC from Agulhas to East Australian current with some sort of dipole increased/decreased RMSE)? and/or near upwelling regions California coast, Northwest Africa? any reasons on why the assimilation would not do better there? As in Figure 6 and 7 statistical significance should be added on both Figure 4 and

5, this will make the result more robust and these increased RMSE might not turn out to be significant (?). Also please indicate how the significance is calculated.

Response:

Thank you for highlighting the potential impact of ocean currents and upwelling regions on RMSE differences in Figure 4. We agree that the assimilation results show degradation near strong ocean currents and upwelling regions. These regions are characterized by strong horizontal gradients and mesoscale variability, which are not well captured by MPAS-O at relatively low resolution and hence pose challenges for the assimilation system and likely contribute to increased RMSE.

We have revised the manuscript to include additional discussion (L279-283) to highlight the impact of strong ocean currents and upwelling regions on assimilation performance.

L279-283: However, increased RMSE values are observed near strong ocean currents and upwelling regions, such as the Gulf Stream, Agulhas Current, and the California coast. These regions are characterized by strong horizontal gradients and mesoscale variability, which are not well captured by MPAS-O at relatively coarse resolution and hence present challenges for the assimilation system and likely contribute to diminished performance.

In response to your suggestions, we have applied the paired t-test to assess the statistical significance of the RMSE differences in Figures 4 and 5. Specifically, the mean and standard deviation of the square error differences were calculated at each grid point. The t-statistic was computed as the ratio of the mean difference to the standard error to determine statistical significance. A confidence level of 95% ($p < 0.05$) was used to identify regions where the RMSE differences are statistically significant. Our results show that some regions of increased RMSE also pass the significance test.

We have updated Figures 4 and 5 to use dotted regions to illustrate statistical significance. Regions where the RMSE differences are statistically significant at the 95% confidence level are now marked with dots in the revised figures.

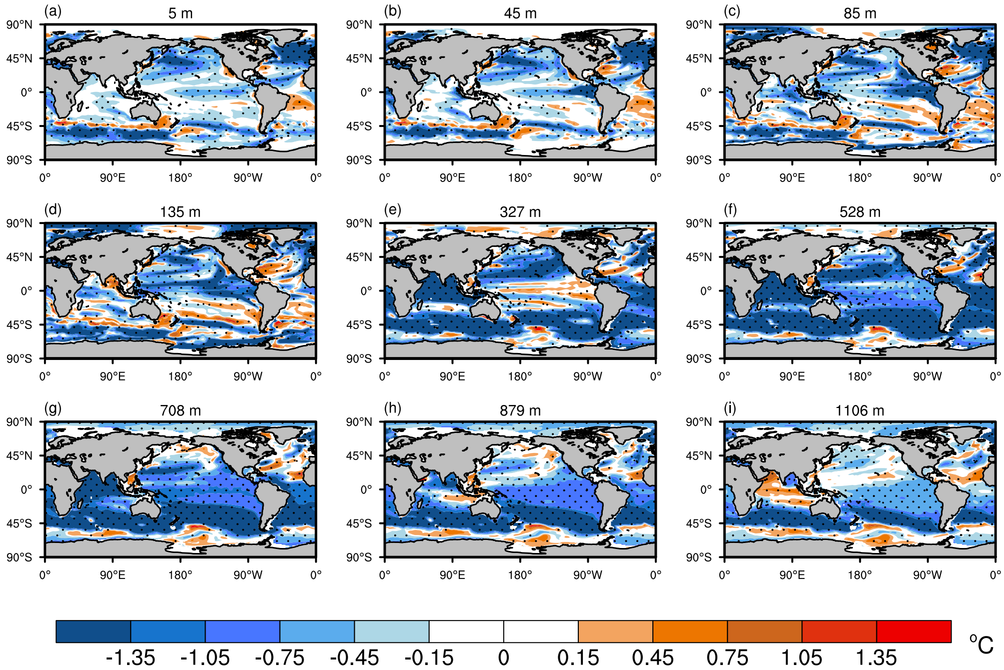

Figure 4. Spatial patterns of root mean square error (RMSE) differences in ocean temperature (unit: °C) between ASSIM and CTRL across nine ocean layers from 1950 to 2021. The RMSE differences are shown for nine different ocean depths: (a) 5 m, (b) 45 m, (c) 85 m, (d) 135 m, (e) 327 m, (f) 528 m, (g) 708 m, (h) 879 m, and (i) 1106 m. Dotted areas represent statistical significance at the 95% confidence level.

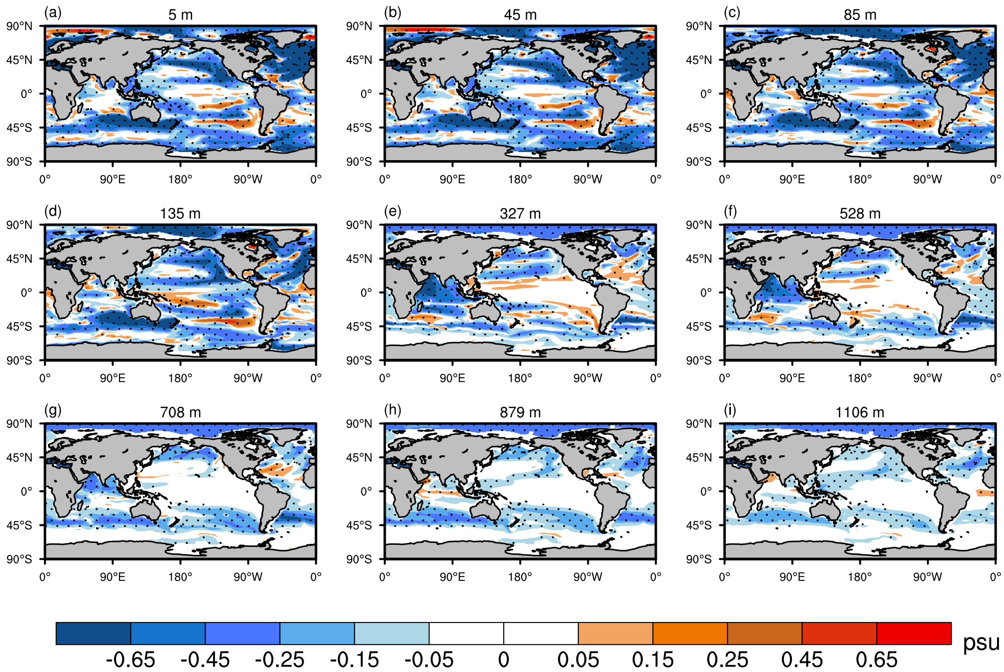

Figure 5. Similar to Figure 4 but for ocean salinity (unit: psu).

Comment#9:

Line 256: Again, 85m is not "deep-ocean dynamics"

Response:

Yes, we have included deeper ocean layers at 879 m and 1106 m in the updated Figure 4, and replaced the term "deep-ocean dynamics" with "deeper ocean processes" (L289).

L289: deeper ocean processes

Comment#10:

Line 274-276: It seems a good hypothesis with maybe hints of this in Figure4 as well showing reduced RMSE on temperature in this region. Since you have both the observations and simulations it would be interesting to confirm and show this, if assimilation actually helps representing El Nino/La Nina better. Otherwise one can wonder why the correlation is largely increase in the Pacific compare to, i.e, the Atlantic.

Response:

Thank you for your insightful comment. We have conducted an additional analysis of the Niño 3.4 index and included a new supplementary figure (Figure A3). Figure A3 compares the time series of the winter Niño 3.4 index from the observation, ASSIM, and CTRL. Our analysis shows that the correlation coefficient between the time series of the winter Niño 3.4 index and observation is improved from 0.06 in CTRL to 0.79 in ASSIM, confirming that the assimilation enhances the representation of El Niño/La Niña variability.

We have included this new Figure A3 and the associated results (L315-317) in the revised manuscript to support the better representation of El Niño/La Niña variability from ASSIM.

L315-317: Further analysis of the winter Niño 3.4 index (Fig. A3) confirms that the assimilation improves the representation of ENSO variability, with the correlation coefficient increasing from 0.06 in CTRL to 0.79 in ASSIM.

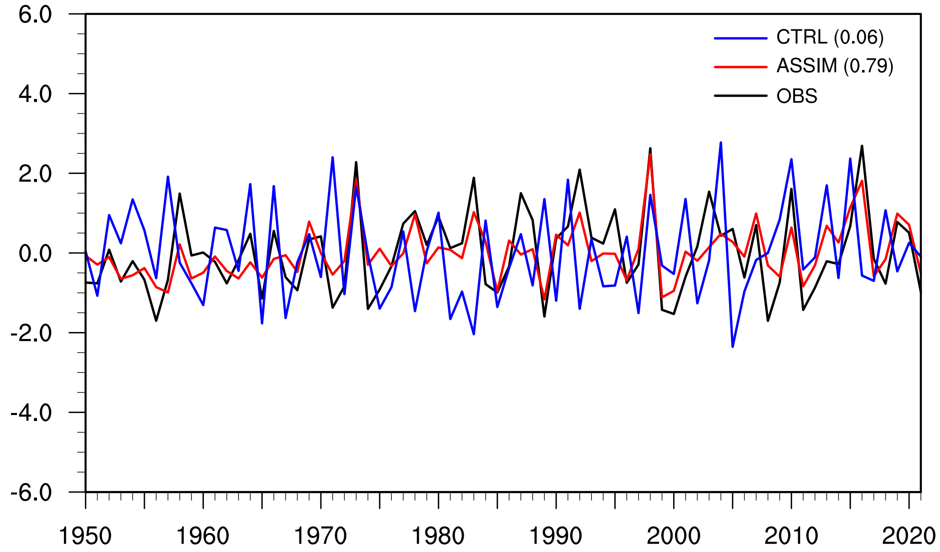

Figure A3. Time series of the winter Niño 3.4 index from 1950 to 2021 for the observation (black line), ASSIM (red line), and CTRL (blue line). The correlations of the Niño 3.4 index with the observation in ASSIM and CTRL are also shown.

Comment#11:
Line 276: “considerable”: considering the values in the Indian Ocean this vocabulary might be too strong.. otherwise the Southern Ocean have similar magnitude, is that also “considerable improvements”..?

Response:
We agree that the term "considerable" might overstate the magnitude of the improvements in the Indian Ocean. We have revised this sentence to remove the mention of the Indian Ocean and replaced "considerable" with "noticeable" (L317).

L317: Moreover, parts of the North Pacific also exhibit noticeable improvements.

Comment#12:
Line 277-278: “complex ocean dynamics” This term was used before in the text, but this is a bit too generic to explain differences and especially to explain the diminished performance, you have to be more specific or don't use this. what is "complex ocean dynamics"? In that regard, then it means that the simulation without assimilation (CTRL) is doing better at these "complex ocean dynamics", then why? I believe here, from the stippling on the figure, it doesn't seem that these reduced correlations are statistically significant (whereas all the increased ones are) and thus, this might just be due to internal variability of the ocean model?

Response:
Thank you for your insightful suggestion. We agree that the term "complex ocean dynamics" is too generic and does not adequately explain the observed diminished performance. In response to your feedback, we have removed the term "complex ocean dynamics" and added more specific factors. Additionally, when deeper layers (e.g., 708 m, 879 m, and 1106 m) are included, we observe that some regions of reduced correlations also pass the significance test. For these areas, the diminished performance may result from sparse observational coverage introducing higher uncertainty into the assimilation process or imbalances between ocean state variables during the assimilation (Edwards et al., 2015; He et al., 2020).

We have revised this sentence (L319-321) to include more specific factors contributing to the diminished performance.

L319-321: However, certain areas exhibit diminished performance, possibly due to sparse observational coverage introducing higher uncertainty into the assimilation process or imbalances between ocean state variables during the assimilation (Edwards et al., 2015; He et al., 2020b).

Comment#13:
Line 286: Again, this doesn't seem to be statistically significant

Response:

Yes, we have revised this sentence (L327-328) to remove the mention of the Indian Ocean and focus only on regions with significant results.

L327-328: Noteworthy improvements are evident in the tropical Pacific, North Pacific, and parts of the North Atlantic.

Comment#14:

Figure 8: Related to previous comment on ocean model depth. Here, why stop the profiles at 1106m, nothing below? 1106m which is rather a very specific depth..? and thus why showing only up to 85 m on the previous panels, Figure 4 to 7? As previously mentioned, it would be interesting to see maps at deeper ocean levels as well as it seems here that the assimilation is improving significantly deeper in the ocean? Instead of showing 9 layers in the first 90m which are showing very similar results.

Response:

Thank you for pointing this out. We appreciate your suggestion to extend the depth range and improve the layer selection in the figures. The ocean model consists of sixty ocean layers from 5 m to 5375 m. The original profiles stopped at 1106 m because this depth corresponds to the 41st layer of the 60-layer ocean model, which is the closest to the upper 1000 meters of the ocean. Based on your feedback, we have added Figure A4 to extend the vertical profiles to cover all layers from 1106 m to the deepest layer at 5375 m. The results in Figure A4 show that the RMSE differences between ASSIM and CTRL become notably smaller for both ocean temperature and salinity, with temperature differences generally within 0.10 °C and salinity differences within 0.02 psu below 2000 m. This suggests the limited impact of assimilation in the deeper ocean.

We have included this new Figure A4 to extend the vertical profiles from 1106 m to 5375 m, and added the relevant results (L347-349) to the revised manuscript.

L347-349: The extended profiles in Figure A4 indicate that below 1106 meters, the RMSE differences between ASSIM and CTRL gradually decrease for both ocean temperature and salinity, suggesting the limited impact of assimilation in the deeper layers.

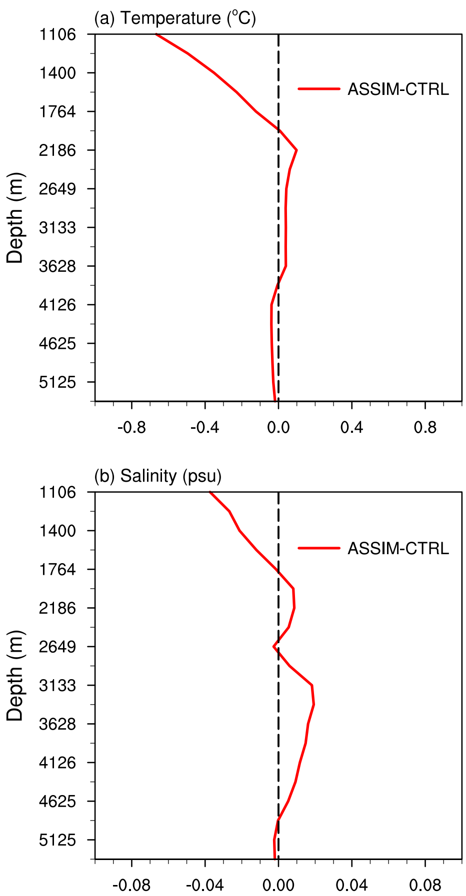

Figure A4. Vertical profiles of the globally averaged RMSE differences between ASSIM and CTRL for (a) ocean temperature (unit: °C) and (b) ocean salinity (unit: psu) with depths from 1106 m to 5375 m.

In response to this comment, we have revised Figures 4 to 7 to include nine different ocean layers: 5 m, 45 m, 85 m, 135 m, 327 m, 528 m, 708 m, 879 m, and 1106 m. Additionally, we have updated the corresponding descriptions of the assimilation results in the deeper layers (879 m and 1106 m) for Figure 4 (L285-287), Figure 5 (L295-298), Figure 6 (L317-318) and Figure 7 (L328-331).

L285-287: In the deeper layers, the assimilation still shows notable improvements in regions such as the North Pacific and parts of the Southern Ocean, though with more pronounced degradation observed in the equatorial Atlantic and parts of the Indian Ocean.

L295-298: In the deeper layers, the improvements are less extensive but remain evident in regions such as parts of the North Atlantic and North Pacific. However, RMSE degradation becomes notable in the equatorial Atlantic and parts of the Indian Ocean, highlighting the need for further improvements in these regions.

L317-318: In the deeper layers, improvements are observed in the western Pacific and parts of the Southern Ocean.

L328-331: In the deeper layers, the improvements in correlation become more localized, primarily concentrated in the western Pacific and parts of the Southern Ocean. Meanwhile, reductions in correlations are observed in parts of the equatorial Pacific and the South Atlantic, indicating the need for further improvements.

Comment#15:

Line 299: "gradually decrease as depth increases" but it does increase again below 300m? Is it not significant ?

Response:

Thank you for your comment. We agree that the phrase "gradually decrease as depth increases" does not fully capture the observed trend. The RMSE differences for salinity gradually decrease from 155 meters to 305 meters, but a slight increase is observed between 305 meters and 1106 meters. We have revised this sentence (L343-345) to clarify this result.

L343-345: The RMSE differences gradually decrease as depth increases from 155 meters to 305 meters, but a slight increase is observed between 305 meters and 1106 meters.

Comment#16:

Line 306-307 and Figure 9: it could be useful to show the average over different depths not only 0-1000m. i.e. 0-300m, 0-700m, 0-1000m. From fig 8, doesn't this "systematic overestimation of temperature" come from 300m and below? if looked at the surface or over different depths it may not systematically overestimate the temperature? This would give more insight into what is improved or not depending on the depth, as in Fig8. Similar thing could be done for salinity.

Also, is there a spin-up period to take into account to analyze the result or when doing this kind of assimilation? it seems to take 10~15 years for the bias to approach the 0 line (Fig 9a,c). Maybe a comment on that would be useful.

Response:

Thank you for these insightful suggestions. We have included additional analyses showing the temporal variations of bias and RMSE averaged over different depth ranges: 0-300 m (Figure A5) and 0-700 m (Figure A6). The results in Figures A5 and A6 show that the systematic overestimation of temperature in CTRL primarily originates from depths below 300 meters, consistent with the vertical profiles shown in Figure 8. In contrast, the salinity bias in CTRL is already prominent in the upper 300 meters.

We have added Figures A5 and A6 to include these depth-specific averages and revised the corresponding text (L354-355 and L360) to reflect these findings.

L354-355: This overestimation in ocean temperature primarily originates from depths below 300 meters (Figs. A5 & A6).

L360: This salinity bias in CTRL is already prominent in the upper 300 meters (Figs. A5 & A6).

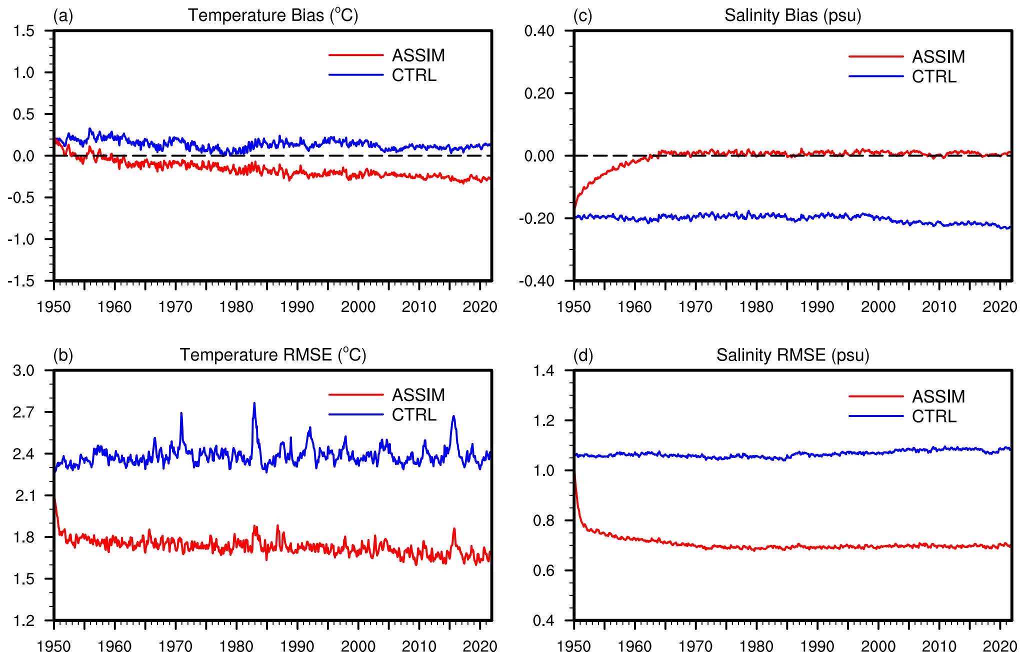

Figure A5. Temporal variations of the global mean bias (a, c) and RMSE (b, d) for ocean temperature (unit: °C) and salinity (unit: psu) averaged over the upper 300 meters. The red lines represent ASSIM, while the blue lines represent CTRL.

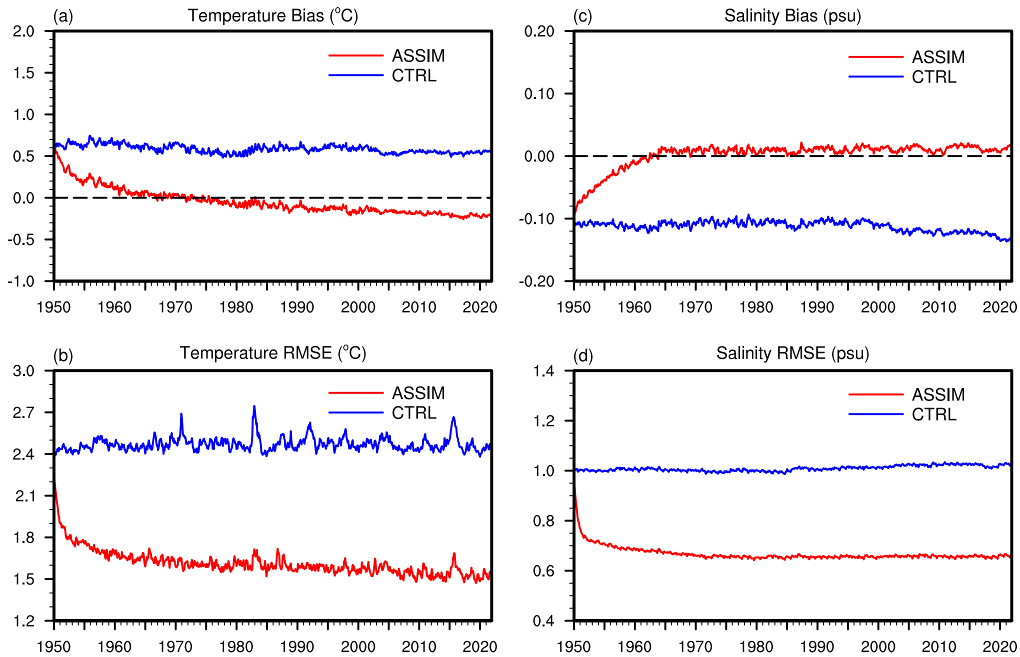

Figure A6. Similar to Figure A5 but averaged over the upper 700 meters.

Regarding the "spin-up" period for the assimilation system, we also acknowledge that it takes approximately 10-15 years for the biases in temperature and salinity to stabilize near the zero line. This behavior reflects the system's equilibration during the assimilation process, where the model takes time to adjust to the assimilated reanalysis.

We have included the following sentence (L363-365) to note that the first 10-15 years were considered as part of the adjustment period.

L363-365: Notably, it takes approximately 10-15 years for the biases in both temperature and salinity to stabilize near the zero line, reflecting an adjustment period where the assimilation system equilibrates.

Comment#17:

Figure 10: What are the units here? The colorbar is the same for both temperature and salinity? is it up to 3 Deg C difference and 3 psu difference or is it in %? This need to be clarified. if in % please quantify in the text how much in degC/psu.

Response:

The color bar in Figure 10 represents differences, not percentages. The units are °C for temperature, and psu for salinity, and the same color bar is used for both panels. To address this, we have added the units (°C for temperature and psu for salinity) to both the figure and its caption (L665-666) to ensure clarity.

L665-666: **Figure 10.** Climatological mean differences in sea surface temperature (left, unit: °C) and salinity (right, unit: psu) from 1950 to 2021.

Comment#18:

Figure 10b,d Should be "ASSIM minus Obs", it would then be much easier to appreciate the improvements in ASSIM in comparison to panels (a,c). Right now it is not clear or difficult to see actually how much the biases are reduced (or not) in ASSIM.

Response:

Thank you for your valuable suggestion. We have modified panels (b, d) in Figure 10 to show "ASSIM minus OBS" instead of "ASSIM minus CTRL". The figure caption has also been updated to reflect this change.

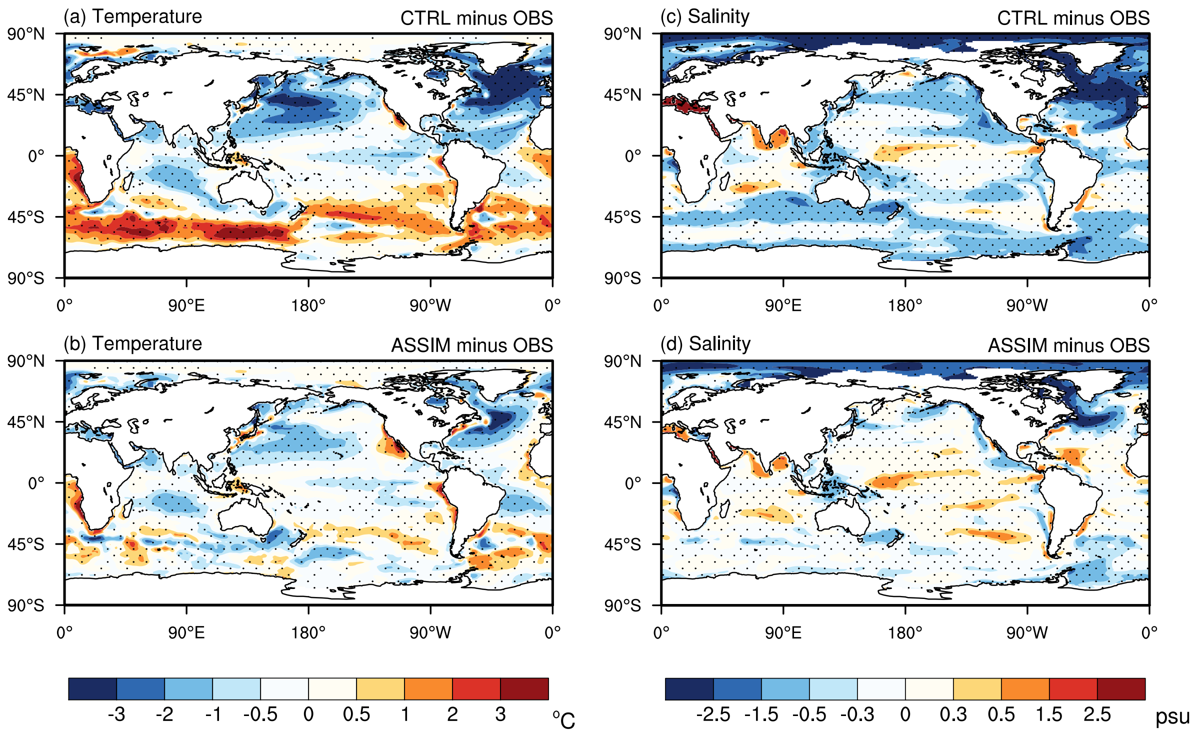

Figure 10. Climatological mean differences in sea surface temperature (left, unit: °C) and salinity (right, unit: psu) from 1950 to 2021. The top panels show the differences between CTRL and observation, while the bottom panels show the differences between ASSIM and observation. Dotted areas indicate regions where the differences are statistically significant at the 95% confidence level.

Comment#19:

Line 323-324: Again, quantify better how much.

Response:

We have revised the text (L373-376) to include specific numerical values to quantify the improvements. Specifically, in the North Pacific and North Atlantic, the cold biases in CTRL are reduced by approximately 1-2 °C. Similarly, in the Southern Ocean, the warm biases are corrected by about 1.5-2.5 °C.

L373-376: In contrast, these SST biases found in CTRL are substantially reduced by ASSIM (Fig. 10b), with cold biases in the North Pacific and North Atlantic diminished by approximately 1-2 °C, and warm biases in the Southern Ocean corrected by about 1.5-2.5 °C.

Comment#20:

Line 325-326: could be worthwhile to note the large high bias in salinity in the Mediterranean Sea as the bias trend seems to be opposite to the global ocean, and it seems to be improved in the ASSIM as well.

Response:

Thank you for pointing this out. In response, we have revised the text (L379-380) to highlight the large positive salinity bias in the Mediterranean Sea from CTRL.

L379-380: Notably, in the Mediterranean Sea, CTRL exhibits a large positive salinity bias exceeding 2.5 psu.

Comment#21:

Line 378-380: In light of my previous comment, I believe this statement should be also removed.

Response:

We agree with your suggestion and have removed this statement from the manuscript.

References:

Chen, J., Liu, H., Bai, C., Yan, H., Lu, K., Bao, S., and Liu, K.: Identifying climate modes contributing to sea surface salinity decadal variation in the North Pacific Ocean, Journal of Geophysical Research: Oceans, 125(10), e2019JC016011, https://doi.org/10.1029/2019JC016011, 2020.

Edwards, C. A., Moore, A. M., Hoteit, I., and Cornuelle, B. D.: Regional ocean data assimilation, Annual Review of Marine Science, 7(1), 21–42, https://doi.org/10.1146/annurev-marine-010814-015821, 2015.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, Journal of Geophysical Research: Oceans, 118(12), 6704–6716, https://doi.org/10.1002/2013JC009067, 2013.

He, Y., Wang, B., Liu, L., Huang, W., Xu, S., Liu, J., Wang, Y., Li, L., Huang, X., Peng, Y., Lin, Y., and Yu, Y.: A DRP-4DVar-based coupled data assimilation system with a simplified off-line localization technique for decadal predictions, Journal of Advances in Modeling Earth Systems, 12(4), e2019MS001768, https://doi.org/10.1029/2019MS001768, 2020.

Van Roekel, L., Adcroft, A., Danabasoglu, G., Griffies, S. M., Kauffman, B., Large, W., Levy, M., Reichl, B., Ringler, T., and Schmidt, M.: The KPP boundary layer scheme for the ocean: Revisiting its formulation and benchmarking one‐dimensional simulations relative to LES, Journal of Advances in Modeling Earth Systems, 10, 2647–2685, https://doi.org/10.1029/2018MS001336, 2018.

Zhu, S., Wang, B., Zhang, L., Liu, J., Liu, Y., Gong, J., Xu, S., Wang, Y., Huang, W., Liu, L., He, Y., and Wu, X.: A Four‐Dimensional Ensemble‐Variational (4DEnVar) Data Assimilation System Based on GRAPES‐GFS: System Description and Primary Tests, Journal of Advances in Modeling Earth Systems, 14(7), e2021MS002737, https://doi.org/10.1029/2021MS002737, 2022.

We thank Reviewer #2 for the constructive comments and suggestions, which greatly help to improve the quality of our manuscript. We have made revisions and replied to all the comments. Please find the point-by-point responses to the comments below. Our responses are shown in "Blue" and the changes in the manuscript are shown in "Red". The line numbers correspond to those in the clean version of our revised manuscript.

Response to the comments from Reviewer #2

General Comment:

Shi and co-authors evaluate an assimilation run produced with the 4DEnVar-based weakly coupled ocean assimilation method applied to the Earth system model E3SMv2. The 4DEnVar method follows the dimension-reduced projection 4DVar method of Wang et al (2010). The present study, together with a recently published work also by Shi et al (2024) [https://doi.org/10.5194/gmd-17-3025-2024] on a weakly coupled land data assimilation for E3SMv2 using the same methodology, appear to be part of the authors effort to produce realistic initial conditions for decadal climate predictions with E3SMv2. This is welcome news for the decadal climate prediction community and climate prediction users at large.

The present work briefly describes the authors implementation of the 4DEnVar method, discussed at length by Shi et al (2024), and presents evaluations of an E3SMv2 assimilation run. These include evaluations of the model 3D ocean temperature and salinity, which are constrained directly with the EN4.2.1 observational dataset, and air temperature and precipitation over the contiguous United States, which are constrained indirectly from the effects of the constrained ocean on the atmosphere. Since the methodology has already been discussed and tested elsewhere, and the results discussed here appear to be robust, I do not have major critical concerns. I do however have several suggestions/comments that I exhort the authors to address to hopefully improve the presentation and enhance the relevance of their work.

Response:

We would like to express our sincere gratitude for your time and effort in reviewing our manuscript. We truly appreciate your constructive comments and suggestions, which have significantly contributed to enhancing the quality of our work. We have carefully addressed each comment, as outlined below, and have made the necessary revisions to our manuscript.

Comment#1:

The evaluation of the E3SMv2 assimilation run only uses a control simulation that does not ingest observational data (other than the external forcing, common to both runs). That is, there is no assimilation method to benchmark against (e.g., simple nudging, or just imposing the observed temperature and salinity fields) so as to assess the effectiveness of the 4DEnVar-based methodology. For example, while Fig. 3 shows that the assimilation run outperforms the pre-assimilation background, it is unclear whether the 4.20 % error reduction (L239-240, Section 3.1.) can be considered “good” enough to justify the complexity of 4DEnVar. I do not suggest to produce a benchmark simulation using a different assimilation method, but I wonder whether the authors (or someone else) have tried a simpler approach on the E3SM model, or whether the authors can comment on previous work showing comparisons between different initialization methods that can shed light on this.

Response:

Thank you for your insightful comments. We agree that comparisons with simpler approaches could strengthen the effectiveness of the 4DEnVar method. Unfortunately, to the best of our knowledge, no previous studies have applied a simpler method (e.g., nudging) for ocean data assimilation in the E3SM model, and implementing such methods for comparison with 4DEnVar would require substantial effort that goes beyond the current scope of this study. However, He et al. (2020) implemented a similar 4DEnVar-based ocean data assimilation system for the FGOALS-g2 model and demonstrated improvements over the previous assimilation system in that model. Specifically, the 4DEnVar method in their study showed an average monthly reduction rate of the cost function of 4.4%, better than 3.0% from the previous system. In our study, we observed a comparable 4.20% reduction, further supporting the effectiveness of the 4DEnVar approach. Furthermore, previous studies have shown that 4DVar-based methods outperform simpler methods, such as nudging and 3DVar, by maintaining dynamical consistency with the model and minimizing initial shocks in the forecasts (Sugiura et al., 2008; Zhang et al., 2020).

Based on your suggestions, we have revised the manuscript to include the advantages of using 4DVar-based methods over simpler assimilation techniques (L87-89) and note that the 4.20% reduction rate in our study is comparable to the 4.4% reported by He et al. (2020a) using a similar 4DEnVar-based assimilation system (L261-263), further supporting the effectiveness of this methodology.

L87-89: Previous studies have shown that 4DVar-based methods outperform simpler schemes (e.g., nudging or 3DVar) by maintaining dynamical consistency with the model and minimizing initial shocks in the forecasts (Sugiura et al., 2008; Zhang et al., 2020).

L261-263: This average reduction rate of 4.20% is comparable to the 4.4% reduction rate reported by He et al. (2020a), who used a similar 4DEnVar-based assimilation system in a different climate model, further supporting the effectiveness of the 4DEnVar approach.

Comment#2:

Section 2.4. In the experiment design, did the authors used a spinup run for equilibration before performing the assimilation, or is the assimilation applied directly from a piControl (L167) or historical (L205-207) run? Can the authors clarify and expand on this?

Response:

To clarify, the assimilation experiment (ASSIM) is initialized directly from the historical run in 1950, with no additional spin-up performed. We have added the following sentence (L228-230) to make this clear.

L228-230: The assimilation run is initialized directly from the historical run in 1950, using the fully coupled state at the start of the simulation.

Comment#3:

L222-224 Do "cost function", "cost function reduction" and "reduction rate of the cost function" refer to the same quantity? Please clearly name the quantity in Eq. 1 and use the same terminology subsequently.

Response:

Thank you for pointing out this inconsistency. To clarify, "cost function reduction" and "reduction rate of the cost function" refer to the same concept. The reduction rate of the cost function represents the percentage decrease in the cost function. The formula presented in Eq. 1 represents the reduction rate of the cost function.

To ensure consistency, we have revised the manuscript to consistently use the term "reduction rate of the cost function" throughout the manuscript, including the description of Eq. 1 (L240-241) and all subsequent mentions.

L240-241: The reduction rate of the cost function serves as a fundamental measure to assess the assimilation system's accuracy, calculated using the formula:

Comment#4:

L227 How is the observation error covariance matrix R computed? How is this matrix for EN4.2.1 ocean temperature and salinity? e.g., Is R diagonal or quasi-diagonal? If not, any insight on its spectral properties? How are the characteristics of R expected to impact the assimilation process?

Response:

The observation error covariance matrix ***R*** is determined statistically by estimating the variance of the EN4.2.1 ocean temperature and salinity data. In this study, ***R*** is assumed to be diagonal. The characteristics of ***R*** directly influence the weighting of observations in the assimilation process: larger values of ***R*** result in smaller weights for observations, whereas smaller values increase the weight of observations.

We have added this clarification (L245-248) to explain the computation of ***R***, its diagonal assumption, and its implications for the assimilation process.

L245-248: In this study, ***R*** is assumed to be diagonal and its diagonal elements are statistically computed based on the variance of the EN4.2.1 ocean temperature and salinity data. The characteristics of ***R*** directly influence the assimilation process, where larger values reduce the relative weight of the EN4.2.1 reanalysis and smaller values increase it.

Comment#5:

L234-242, Section 3.1. While the authors' message is clear, the use of negative percents is odd.

Consider showing positive percents specifying that they correspond to improvements due to the assimilation method.

Response:

Thank you for your suggestion. We have revised the manuscript to specify that positive percentages represent improvements due to the assimilation and to present the reduction rate of the cost function as positive percentages (L256-261).

L256-261: As noted earlier, negative values of the reduction rate of the cost function indicate the successful incorporation of reanalysis data into the coupled model. However, the reduction rate is presented here as positive percentages to represent improvements due to the assimilation. The reduction rate of the cost function reaches 12.03% in the first month. Over the entire 72-year period from 1950 to 2021, the average monthly reduction rate of the cost function is 4.20% for all months in ASSIM.

Comment#6:

L255 Can the authors expand on the two suggested reasons for the performance degradation in the deep ocean? If possible, can the authors provide some comments specific to the E3SMv2 model and the EN4.2.1 observational data?

Response:

We have expanded on the two suggested reasons for the performance degradation in the deep ocean, providing additional context specific to the EN4.2.1 observational data and the E3SMv2 model. For the EN4.2.1 reanalysis, the coverage and quality of observations tend to decrease with depth, which may lead to higher uncertainties in the deep ocean. This sparse observational coverage limits the constraints that data assimilation can impose on the model state in the deep ocean. Furthermore, in the E3SMv2 model, the complexity of simulating deep-ocean processes, such as vertical mixing and bottom water formation, may contribute to biases that are difficult to correct through data assimilation.

In response to this comment, we have incorporated this discussion into the revised manuscript (L298-305) to clarify the potential reasons for the performance degradation in the deep ocean.

L298-305: The degradation in the deeper ocean layers can be attributed to two main factors: observational data limitations and challenges in representing deep-ocean processes in the model. For the EN4.2.1 reanalysis, the coverage and quality of observations tend to decrease with depth, potentially resulting in greater uncertainties in the deep ocean. This sparse observational coverage limits the constraints that assimilation can impose on the model state. Furthermore, in the E3SMv2 model, the complexity of simulating deep-ocean processes, such as vertical mixing and bottom water formation, may contribute to biases that are difficult to correct through assimilation.

Comment#7:

L268-269 Is the seasonal cycle removed from the time series before computing the correlations?

Please specify. And is the linear trend removed?

Response:
Yes, we have removed the seasonal cycle and linear trend before computing the correlations. This clarification has been added to the revised manuscript in Lines 309-310.

L309-310: The seasonal cycle and linear trend have been removed before computing the correlations.

Comment#8:
L277-278 The authors suggest that the degradation in performance is "possibly due to sparse observational data or complex ocean dynamics". Can the authors expand on this? In particular, if the control run does not use observations (except for the external forcing, as it is the case for the assimilation run), how/why the sparse temperature and salinity observations would degrade the performance of the assimilation run relative to that of the control run?

Response:
Thank you for your insightful comment. Sparse temperature and salinity observations may introduce higher uncertainty into the assimilation process. In regions with sparse observations, the assimilation process may introduce biases or errors when attempting to fit the model to incomplete or uncertain data, leading to localized performance degradation relative to the control run. Additionally, possible imbalances between ocean state variables during the assimilation process may also degrade the assimilation performance in certain areas.

In response to this comment, we have revised this sentence (L319-321) to include more specific factors contributing to the diminished performance.

L319-321: However, certain areas exhibit diminished performance, possibly due to sparse observational coverage introducing higher uncertainty into the assimilation process or imbalances between ocean state variables during the assimilation (Edwards et al., 2015; He et al., 2020b).

Comment#9:
L305. According to the text, Fig. 9 shows global mean RMSE of vertically averaged temperature and salinity. From the caption to Fig. 9, it shows RMSE of the vertically averaged global mean ocean temperature and salinity. As these are two different quantities, please correct and clarify which one is shown.

Response:
Thank you for pointing out this inconsistency. We have corrected the caption of Figure 9 (L661-662) to accurately describe the global mean RMSE of vertically averaged ocean temperature and salinity.

L661-662: **Figure 9.** Temporal variations of the global mean bias (a, c) and RMSE (b, d) for ocean temperature (unit: °C) and salinity (unit: psu) averaged over the upper 1000 meters from 1950 to 2021.

Comment#10:

Figure 10. Panels (a) and (c) show CTRL minus OBS. However, from the caption and panels titles, (b) and (d) show ASSIM minus CTRL. Why? I would expect to see ASSIM minus OBS to assess the biases of the assimilating runs relative to those of the control. Please clarify, otherwise I would suggest to show and discuss the results for ASSIM minus OBS. This would imply changes to the discussion in L318-332.

Response:

Based on your suggestion, we have updated panels (b) and (d) in Figure 10 to display "ASSIM minus OBS" instead of "ASSIM minus CTRL". In addition, we have revised the discussion accordingly (L373-376, and L380-383) to reflect the updated figure and results.

L373-376: In contrast, these SST biases found in CTRL are substantially reduced by ASSIM (Fig. 10b), with cold biases in the North Pacific and North Atlantic diminished by approximately 1-2 °C, and warm biases in the Southern Ocean corrected by about 1.5-2.5 °C.

L380-383: Compared with CTRL, ASSIM significantly reduces the overall fresh biases in CTRL (Fig. 10d). Notable improvements are observed in the North Atlantic and North Pacific, where salinity biases are reduced by approximately 0.5-1 psu, and in parts of the Southern Ocean, where reductions reach up to 1.5 psu.

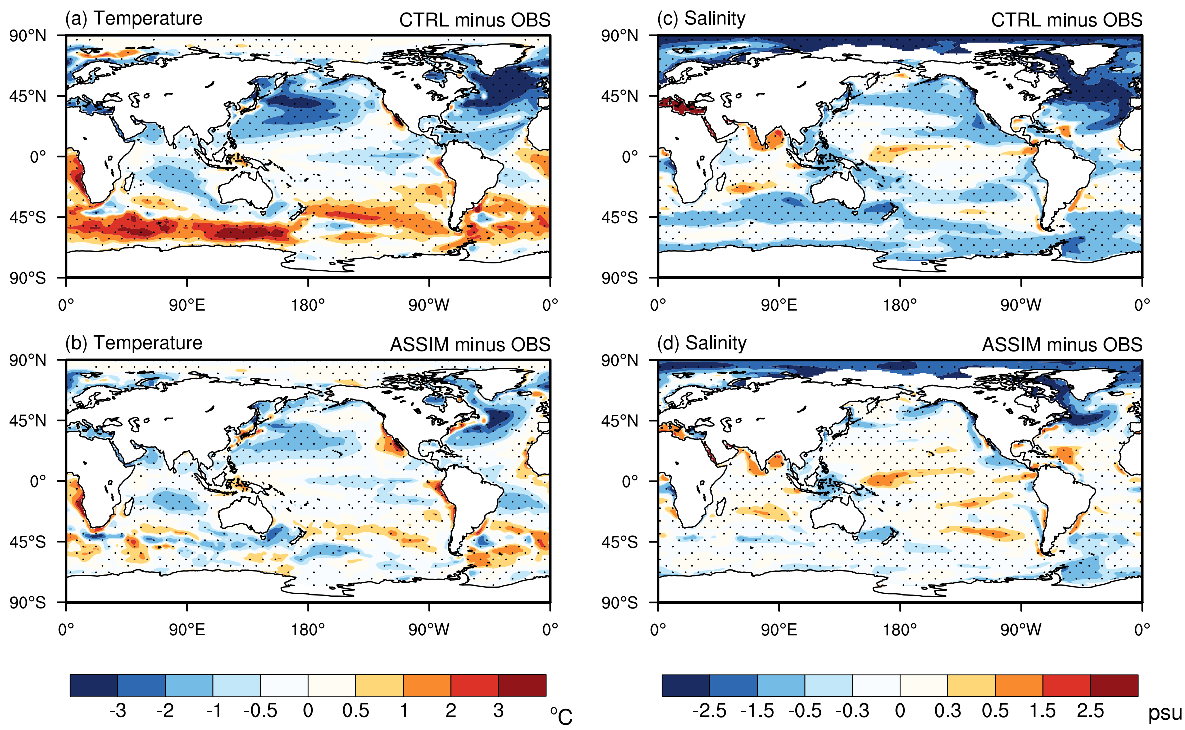

Figure 10. Climatological mean differences in sea surface temperature (left, unit: °C) and

salinity (right, unit: psu) from 1950 to 2021. The top panels show the differences between CTRL and observation, while the bottom panels show the differences between ASSIM and observation. Dotted areas indicate regions where the differences are statistically significant at the 95% confidence level.

Comment#11:
L318. What "mean differences"? Please specify in the text. See previous comment.

Response:
We have revised this sentence (L369-370) to show the climatological mean differences between CTRL and observation, as well as between ASSIM and observation.

L369-370: Figure 10 presents the climatological mean differences between CTRL and observation, as well as between ASSIM and observation, for both sea surface temperature (SST) and salinity (SSS).

Comment#12:
L322. From Fig. 10b it is unclear whether the "SST biases found in CTRL are substantially reduced by ASSIM". See comment above on Figure 10.

Response:
We have revised Figure 10 to replace "ASSIM minus CTRL" with "ASSIM minus OBS" in panels (b) and (d). Additionally, we have incorporated quantitative descriptions (L373-376) to better illustrate the magnitude of bias reduction.

L373-376: In contrast, these SST biases found in CTRL are substantially reduced by ASSIM (Fig. 10b), with cold biases in the North Pacific and North Atlantic diminished by approximately 1-2 °C, and warm biases in the Southern Ocean corrected by about 1.5-2.5 °C.

Comment#13:
In addition to Fig. 11, can the authors show the correlation and RMSE maps for both temperature and precipitation over the contiguous US (including statistical significance)? This will be useful to assess the regional impacts of the assimilated ocean. Perhaps the authors could show results for seasonal averages instead of annual means, choosing the seasons of strongest ENSO influence on US temperature and precipitation.

Response:
Thank you for your valuable suggestion. We have extended our analysis to include the correlation and RMSE maps for temperature (Fig. R1) and precipitation (Fig. R2) during boreal winter, when ENSO exerts its strongest influence on US climate variability. For winter temperature (Fig. R1), correlation improvements are observed across the northern and central US, while reductions in RMSE are primarily concentrated in the central and northeastern US. For winter precipitation (Fig. R2), correlation improvements are evident over the central and southern US, with RMSE reductions prominently observed in the southern US.

However, based on Reviewer #1's comment that the current results in Figure 11 are highly preliminary and require more rigorous analysis to draw robust conclusions, we have followed Reviewer#1's suggestion to remove Figure 11 from the revised manuscript. We sincerely appreciate your constructive feedback, which has been invaluable in refining our analysis. In future research, we plan to integrate Figure 11 and this analysis of correlation and RMSE maps into a separate study, with more comprehensive analyses to enhance the robustness and clarify our findings.

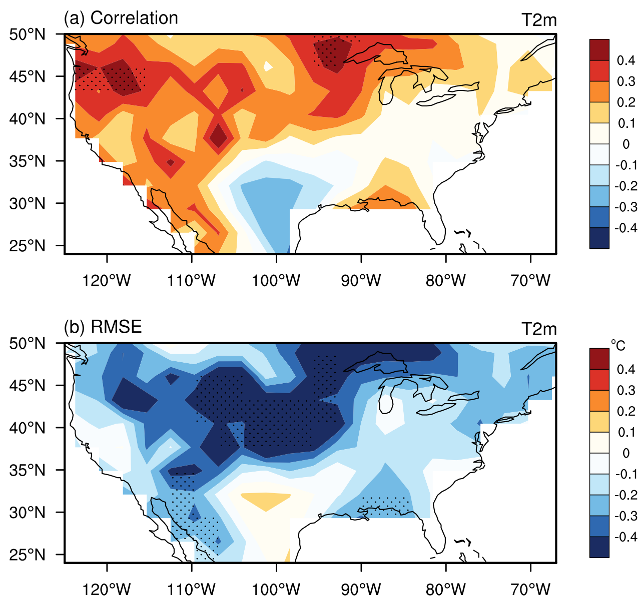

Figure R1. Spatial patterns of correlation and RMSE differences in surface air temperature during boreal winter between ASSIM and CTRL over the contiguous US from 1950 to 2021. Dotted areas represent statistical significance at the 90% confidence level.

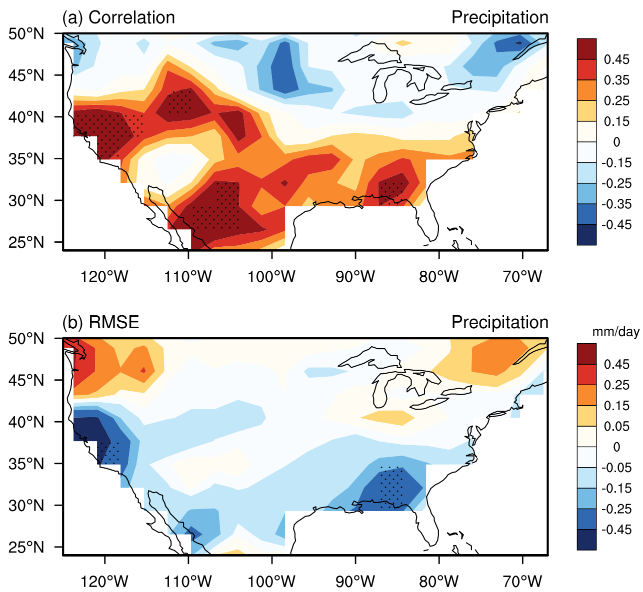

Figure R2. Similar to Figure R1 but for winter precipitation.

Comment#14:

While Fig. 6 shows improved ocean temperature for the assimilation run, this result uses difference of correlations relative to control. As part of the analysis in section 3.6, it would be useful to directly assess the SST variability of the assimilation run in the tropical Pacific, which is expected to influence the simulated climate over land. For example, consider showing time series of the seasonal averaged (e.g., DJF) Niño 3.4 index for the CTRL, ASSIM and OBS, and their correlation with OBS.

Response:

We have added a new Figure A3 to show the time series of the winter Niño 3.4 index for CTRL, ASSIM, and OBS. The correlation coefficient of the winter Niño 3.4 index with observation increases from 0.06 in CTRL to 0.79 in ASSIM, highlighting the enhanced representation of ENSO variability in the assimilation run.

In response to this comment, we have included this analysis as Figure A3 and added the corresponding results (L315-317) in the revised manuscript to underscore the enhanced SST variability in the tropical Pacific.

L315-317: Further analysis of the winter Niño 3.4 index (Fig. A3) confirms that the assimilation improves the representation of ENSO variability, with the correlation coefficient increasing from 0.06 in CTRL to 0.79 in ASSIM.

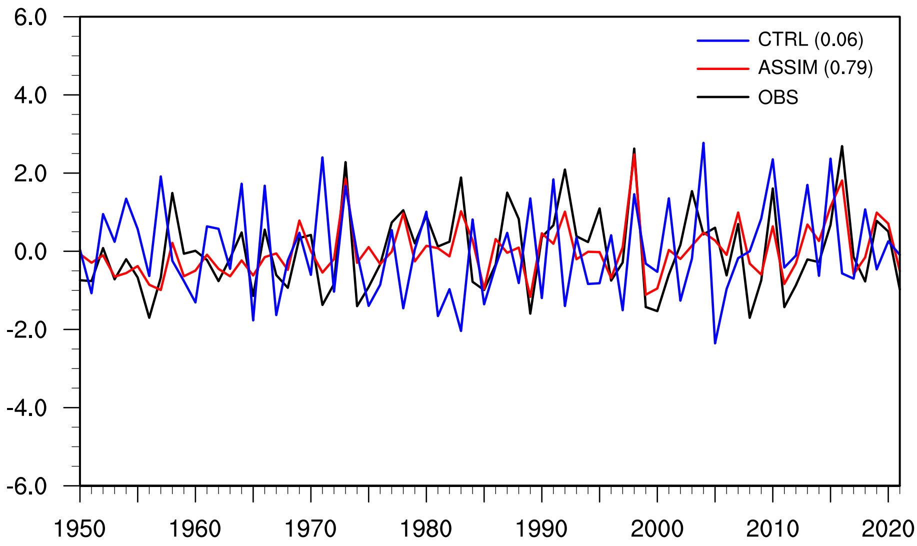

Figure A3. Time series of the winter Niño 3.4 index from 1950 to 2021 for the observation (black line), ASSIM (red line), and CTRL (blue line). The correlation of the Niño 3.4 index with the observation in ASSIM and CTRL are also shown.

Comment#15:

EN products have data-sparse periods and regions. Salinity in particular is sparsely observed and could lead to spurious static instability in the absence of dynamical constrains. Can the authors expand on the potential limitations/advantages of using EN.4.2.1 for the assimilation process instead of reanalysis products such as ORAS5 [https://cds.climate.copernicus.eu/datasets/reanalysis-oras5?tab=overview] or GLORYS [https://data.marine.copernicus.eu/product/GLOBAL_MULTIYEAR_PHY_001_030/description]? This would be useful information in particular for producing centers of decadal predictions. Such discussion could be added to the concluding remarks, aligned with the authors "aim to advance the predictive capabilities of E3SM for decadal predictions".

Response:

Thank you for your insightful comment. We agree that the sparse observational coverage in EN4.2.1, particularly for salinity, could pose limitations to the assimilation process, potentially introducing static instabilities in the absence of dynamical constraints. Reanalysis products such as ORAS5 and GLORYS offer promising alternatives for mitigating these limitations. Future efforts should explore incorporating these reanalysis products into the WCODA system to improve the assimilation performance in challenging areas.

In response to this comment, we have expanded the discussion (L410-415) in the concluding remarks to include the limitations of using the EN4.2.1 dataset and the potential benefits of employing alternative reanalysis products like ORAS5 and GLORYS.

L410-415: The reliance on the EN4.2.1 product could pose limitations to the assimilation process due to the sparse salinity observations and potential for static instabilities in data-sparse regions. Reanalysis products such as ORAS5 and GLORYS provide promising alternatives for mitigating these limitations. Future efforts should explore incorporating these reanalysis products into the WCODA system to improve the assimilation performance in challenging areas.

Comment#16:

L109 Change "employs sophisticated representations of" with "represents"

Response:

Done.

Comment#17:

L292 Change "variations in" to "of"

Response:

Done.

Comment#18:

L306 Add "over the top 1000 meters"

Response:

Done.

Comment#19:
L341 What “multiple US regions”? The contiguous US?

Response:
To clarify, "multiple US regions" refers to areas within the contiguous US. We have removed this term in the revised manuscript.

Comment#20:
L633 Change “in the” to “averaged over”.

Response:
Done.

Comment#21:
L382 Consider changing “challenges” with “limitations”

Response:
Done.

References:

Edwards, C. A., Moore, A. M., Hoteit, I., and Cornuelle, B. D.: Regional ocean data assimilation, Annual Review of Marine Science, 7(1), 21–42, https://doi.org/10.1146/annurev-marine-010814-015821, 2015.

He, Y., Wang, B., Huang, W., Xu, S., Wang, Y., Liu, L., Li, L., Liu, J., Yu, Y., Lin, Y., Huang, X., and Peng, Y.: A new DRP-4DVar-based coupled data assimilation system for decadal predictions using a fast online localization technique, Climate Dynamics, 54, 3541–3559, https://doi.org/10.1007/s00382-020-05190-w, 2020a.

He, Y., Wang, B., Liu, L., Huang, W., Xu, S., Liu, J., Wang, Y., Li, L., Huang, X., Peng, Y., Lin, Y., and Yu, Y.: A DRP-4DVar-based coupled data assimilation system with a simplified off-line localization technique for decadal predictions, Journal of Advances in Modeling Earth Systems, 12(4), e2019MS001768, https://doi.org/10.1029/2019MS001768, 2020b.

Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., Igarashi, H. and Ishikawa, Y.: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations, Journal of Geophysical Research: Oceans, 113, C10017, https://doi.org/10.1029/2008JC004741, 2008.

Zhang, S., Liu, Z., Zhang, X., Wu, X., Han, G., Zhao, Y., Yu, X., Liu, C., Liu, Y., Wu, S., Lu, F., Li, M., Deng, X.: Coupled data assimilation and parameter estimation in coupled ocean–atmosphere models: a review, Climate Dynamics, 54, 5127–5144, https://doi.org/10.1007/s00382-020-05275-6, 2020.