The paper introduces and validates a high-resolution configuration of the EC-Earth model, utilizing the IFS T1279 atmosphere and NEMO ORCA12 ocean components. The validation confirms that the model is technically ready for use and, in many aspects, shows reduced biases compared to its lower-resolution predecessors, although some tuning is still required, particularly for the Southern and Arctic Ocean regions. The authors definitely did a great job. I therefore recommend this paper for acceptance after minor revisions, which I have summarized below.
We thank the Reviewer for the comments on our work. In the following, we answer each specific point (in blue).

My main concern is that the plots should be updated to meet the standards of the 202Xs. This is especially relevant for Figures 7 and 14. Based on visualizations I have seen on YouTube and in several presentations, the presented plots appear outdated.
In our opinion, figures follow many of the current standards in plot making, such as clean and concise design, accessible color schemes, clear labels, and high data-ink ratio. Nonetheless, we have modified certain figures as requested by both Reviewers, either to increase readability, or to add new information, or both. High-quality rendered figures will also be provided during the publications process.

L43: Although the message is clear I would reformulate this sentence which currently states that biases lead to something improved.
Clarified.

L155: Regarding the effective resolution for the ORCA grid, could you cite the appropriate reference or tell how you compute it?
A reference has been added.

L166: independent on which grid, atmospheric or ocean? In general which type of remapping between ocean and atmosphere do you use? Is it conservative? Do you conserve the global freshwater balance? If not, how large is the imbalance?
Coupling variables are communicated between the different model components via the OASIS3-MCT coupler, using the parallel SCRIP library as interpolation method. Some variables are conservative, to ensure global freshwater and energy conservation. The conversion process requires MPI collective communications, to fix potential imbalances and ensure that any possible imbalance is negligible over time. This approach is described in detail for the CMIP6 configuration of EC-Earth (Döscher et al., 2022). The reference has been added to the reviewed manuscript.

References
Döscher, R., et al.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geosci. Model Dev., 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

I am also curious if you couple the ocean velocities for computation of the wind stress as this could explain some of the biases. Also I would list the number fluxes and surface fields which are being exchanged.
The fields that are passed from the atmosphere to the ocean are:

- Momentum fluxes for the ocean and sea ice on the U and V grids of the NEMO model.

- Non-solar and solar radiation over the ocean, total evaporation, and precipitation (which is conserved).
- Solar and non-solar radiation over sea ice, the sensitivity to temperature of non-solar energy flux over sea ice, and evaporation over sea ice (which is not conservative).

These are the ocean fields passed to the atmosphere:

- SST; temperature, albedo, fraction, and thickness of the sea ice; and snow thickness over ice.

Such coupling characteristics are the same as those described for the CMIP6 version of EC-Earth, which are described in Döscher et al. (2022) and thus do not include ocean velocities.

References

Döscher, R., et al.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geosci. Model Dev., 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Fig. 2 Does this include the I/O cost? How many cores were assigned to OIFS, and how many to NEMO? What would the scaling for the individual uncoupled components look like, or at what number of grid points per core would scaling saturate according to the technical reports of these two models? If I understand correctly, for a 50x50 distribution in NEMO, this would be (4322*3059)/(4800/2) > 5,500 grid points per core. Given that NEMO scales linearly until around ~500 grid points per core, the saturation is reached much earlier than expected, possibly due to coupling issues or I/O. Is this correct?

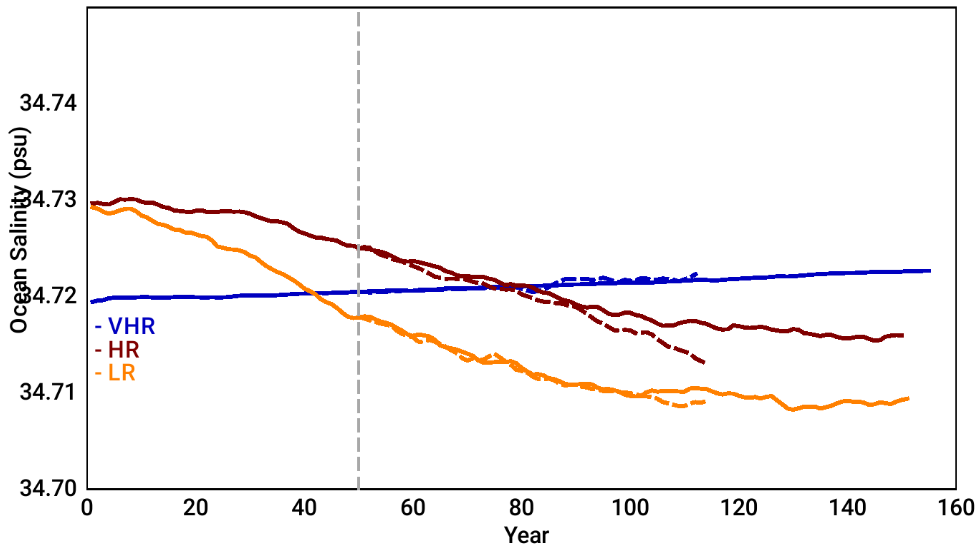
We thank the Reviewer for pointing this out. The results do include the I/O cost, and that is one of the reasons why it does not fully scale. The saturation is reached much earlier than expected, because of the coupling (exchange between atmosphere and ocean due to waiting times and costly conservative exchanges), and also because of the IFS I/O approach. In EC-Earth3, while NEMO uses a parallel and asynchronous I/O server (XIOS), IFS cycle36 does not use it, which leads to the model treating the flush of the output serially. This is computationally more expensive and the reason why IFS saturates sooner. We therefore consider we do not need an independent scalability test for each model component.

L303: You say here that the spin up for 50 years is insufficient but from the plot we clearly see that even 150 years are insufficient. In Fig. 3 (lower panel) I note that the warming drift in VHR is the smallest for the global ocean but it is the largest for the upper ocean (upper panel). Does this imply that the overly large warming drift in the upper ocean is compensated by a negative drift in the deep ocean?

We agree that the extent of the spin-up is insufficient to equilibrate the full ocean, as discussed in the Discussion section. Concerning the drifts in VHR, in both the global ocean (Fig. 3b) and the upper-100-m (Fig. 3a), VHR has the smallest drifts (solid blue lines) out of the three configurations. This is also described with trend values in the manuscript, at the beginning of Section 3.1: "Across all three model resolutions, the length of the spin-up (50 years) appears to be insufficient to equilibrate the full ocean (Fig. 3b); in fact; the ocean temperature is still drifting about 0.001–0.002 °C/yr towards warmer conditions at the end of the control simulation (computed over the last 50 years) in the three configurations. In the upper ocean, however, VHR shows a substantially smaller warming drift than the other two configurations: about 0.00005 °C/yr compared to 0.0025 °C/yr and 0.0062 °C/yr in HR and LR, respectively (computed over the last 50 years; Fig. 3a). "

Regarding my previous comment, is there drift in global salinity or the system is conserved?

The Figure below shows the evolution of the mean salinity in the upper 100 m and in the whole ocean in the spin-up, control, and historical simulations. The drift in the full ocean in the three configurations is no larger than 0.00005 psu/yr in the last 50 years. VHR again shows the smallest drift out of the three. The result points to very small adjustment in the ocean at the end of the control period, similar or smaller than the changes over the historical period (dashed lines). Although we have not added this extra figure to the revised manuscript, the results are now mentioned in Section 3.1.



Response Figure 1. Mean global oceanic salinity (in psu) in the LR (yellow), HR (red), and VHR (blue) models in the spin-up runs (0–50-year period), control runs (50–150-year period; solid lines), and historical runs (50–114-year period; dashed lines). The vertical dashed line marks the end of the spin-up period.

L342: While the precipitation bias over the ocean, especially the “dipole structure” bias, will be mixed by ocean eddies, likely resulting in a low ocean salinity bias, the precipitation bias over land, especially in northeastern South America, will lead to heavily reduced runoff from land, as seen in the Amazon River. I observe this bias in all three simulations. What is the origin of this bias? Is it also present in other CMIP-type models?

Monteverde et al. (2022) study the precipitation biases in CMIP6 models over the Amazonia, including the CMIP6 version of EC-Earth. They find that models fail to reproduce the seasonal cycle in precipitation, as well as extreme events. The three configurations of EC-Earth show a similar bias in their precipitation, which suggest similar mechanisms might be at play. This has been added to the reviewed manuscript.

References

Monteverde, C., De Sales, F. and Jones, C.: Evaluation of the CMIP6 performance in simulating precipitation in the Amazon River basin. Climate, 10(8), 122, https://doi.org/10.3390/cli10080122, 2022.

L408: I also observe large deep mixing which takes place in the Arctic Ocean north of the Bering Strait. Seems like the Arctic Ocean is well mixed there which shall not be the case according to my knowledge. Nevertheless the ice patterns look nice but you didn't provide the sea thickness. Could you augment Fig. 7

with sea ice thickness such that the story becomes more consistent. As you say thick ice and the excessive brine rejection are the likely reasons for what is happening.

L654: In my opinion, the negative aspects are also strongly reflected in the biases observed in the Arctic Ocean, which should be addressed in the discussion section.

We answer the last two comments together, because they are related. The Reviewer is right in pointing to the excess of ocean mixing in the Arctic in VHR. As proposed, we have updated Figure 7, so now the shading represents the bias in sea ice thickness in VHR with respect to PIOMAS. As we discussed in the initial manuscript, the sea ice excess could be connected with the excess of ocean mixing in the central Arctic. Too thick ice would lead to overly brine rejection, explaining the positive surface salinity bias. Since more salinity makes the waters denser, this will lead to a less stratified water column, which is easier to mix. The results are now also highlighted in the Discussion Section and Abstract.