

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2020-446-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2020-446

Neil Swart (Referee)

Referee comment on "The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6" by Ralf Döscher et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-446-RC1>, 2021

General Comments:

This is a nice paper documenting the new EC-Earth model. It covers a lot of the material relevant for the community to understand this model. I particularly appreciated the detailed sections on model tuning, and replication. I think the paper could be improved by streamlining the introduction and conclusions, reducing repetition and addressing the specific questions raised below regarding specific details about the modelling system. Once that is done, I support publication.

By line:

45 "largely improved physical performance". In the following sections, various biases are documented, some of which have improved (e.g. winds) and others which have degraded (e.g. tas). No overall assessment of the model skill / improvement is made. This could be useful (also see comment for conclusions).

52: Possible citation is: Flato, G.M. (2011), Earth system models: an overview. WIREs Clim Change, 2: 783-800. <https://doi.org/10.1002/wcc.148>

54-61: Some references would be appropriate.

83-85: These web statistic citations probably need to be more robust given the journal style guide. Normally the date of the access and url is required. It is relevant

here as these statistics will change in time. I also note that this refers to a search, not a url, and Google customizes web searches to users, so different users will get different results. Personally, I cannot verify these numbers, and I get completely different results than reported here when searching cesm "climate model" (7900) and "ec earth climate model" (234). Or for example "community earth system model" (11200) or "community earth system model (cesm)" (5140). It's not clear therefore that this is really a robust metric.

87 "development has started in" -> "development started in"

102, 104: dynamical Greenland but not Antarctica? What is it physically justified including one major icesheet but not the other?

119-120: "and it is used in its version 3.3 for CMIP6" -> "and version 3.3 is used for CMIP6"

130: TM5 - > reference?

150-151: Is this initialization and forcing data publicly available? Is there a reference or a link to it?

165-166: Is this conservation assumed or is it verified? If it is verified, how? A comparison of 3D ocean heat content with TOA fluxes and surface fluxes in a long piControl can be used to verify to first order.

171-175: It would be interesting to know more (or see) what the drainage basins look like, and what the distribution over coastal points looks like. Is runoff inserted into a single ocean grid cell at the river mouth? Is it spread more widely? Inserted at the surface? Does runoff have properties (nutrient, temperature) or is it inserted at SST?

176-180: So moisture is not conserved within IFS? Why correct runoff to compensate and not E-P directly? The E-P correction could be distributed with the tendencies. Also, why is the correction diagnosed over the transient historical period as opposed to in a balanced piControl? Is there any evidence that this imbalance/correction is constant in time, or could it vary?

What about snowfall into the ocean? Is this being thermodynamically accounted for (i.e. latent heat to melt snow and bring it to SST)?

189-192: Is this duality lead to unphysical behaviour, e.g. with opposite moisture tendencies in the two components? How does this affect conservation of moisture?

195-205: tuning these parameters almost certainly has a significant influence on the ECS of the model. Probably worth noting.

208: "allowing to constuct" -> "allowing construction of"

265: The resolution of the standard "non-LR" ocean was not noted above as far as I see, but here ORCA1 is noted for the "LR" version.

298: How is 100 years of spinup selected? In my experience, this is not enough time, and the model is likely still drifting after 100 years. 500 years, and often 1000+ years is required for the long time scales of the ocean and deep soils to equilibrate in models of this nature.

311 : redundant title repeated twice

320 "has been" -> "had been"

399: "In this case, there is no coupling to the ice sheet model" - this is confusing, as this whole section is about coupling to the ice sheet model for PMIP. Perhaps this should say "For other resolutions"?

455: incomplete sentence.

483: Some of these version details have been mentioned above. Perhaps only specify them here to help with length.

518-523: This implies that snow albedo over land and sea-ice is computed differently. One might imagine a kind of non-physical line, where albedo changes as you move from "land snow" to "sea-ice snow" due to different parameterizations used in the sub-component models. A comment on the consistency of snow albedo and other properties across the land-sea-ice boundary is perhaps warranted.

595-600: a bit of repetition on the orbital parameters.

I do not see any discussion here about how the land mask for the atmosphere is derived. Is a fractional or binary land mask being used? How does this relate to the land mask in the ocean model? Is there tiling, and how are fluxes from ocean/land/ice combined?

500-700: I do not see any explicit mention of how lakes and other inland water bodies are handled. This is important for their regional impact, and also how they are treated with respect to conservation of the global water cycle.

773: Is this regular ORCA1 or the eORCA1 configuration?

Box: The protocol for testing replicability:

Why use a single ocean restart with a small perturbation? Basically, larger state variations in the ocean are being excluded here, but would contribute to internal variability. 20 year simulations would not be long enough for the oceans to diverge significantly within the ensemble simulations themselves.

Box: It's not clear whether the statistical testing (6) is on the standard metrics (5) only or also on the raw fields? The raw fields are shown in figure 2. If the testing were only on the standard metrics, it could lead to false negatives, because simulations on two platforms could have similar global-level biases with completely different underlying structure.

980-981: That a difference was detected and corrected proves the test can be useful, but it does not indicate instances when it might have failed to detect a difference. The test, especially if it is as above, definitely has a significant chance of failing to detect real differences. e.g. also see Baker, A. H., Hammerling, D. M., Levy, M. N., Xu, H., Dennis, J. M., Eaton, B. E., Edwards, J., Hannay, C., Mickelson, S. A., Neale, R. B., Nychka, D., Shollenberger, J., Tribbia, J., Vertenstein, M., and Williamson, D.: A new ensemble-based

consistency test for the Community Earth System Model (pyCECT v1.0),
Geosci. Model Dev., 8, 2829–2840, <https://doi.org/10.5194/gmd-8-2829-2015>, 2015.

Section 5:

In terms of the CMIP6 protocol, the DECK simulations should be submitted for each model configuration. Please discuss the status of this, as it seems piControl runs have only been done for a limited number of configurations.

1017-1018: Can you say the response is too high? What about internal variability?

1024-1030: The warm bias in the south is significant, and represents a large deterioration relative to the previous model version. This seems a bit inconsistent with the statements made in the abstract. Some discussion of the source of the new large bias would be appropriate.

Conclusions: I don't think it is necessary to summarize the result over every bias again here. Perhaps replace this with a shorter, more general overview of conclusions of the validation.

Table 1: What about ocean BGC?

Table 2+: timestep in the atmosphere or ocean or both? Presumably not the same. What about the coupling interval? That is noted in some of the subsequent tables (7,8) but not all of them (3-5, 10).

Figure 10: Showing both hemispheres on the same seasonal panel makes the changes very hard to see.

Figure 12: Would be clearer in two panels, one for Sep and one for Mar.

