Reply to reviewers' comments

We thank both reviewers for their careful second review of our manuscript. Below are our responses to their technical edits/minor comments. The reviewers' comments are in blue, our responses in black.

Reviewer 1

The reviewer thanks the authors for their comprehensive addressing of the initial comments on the manuscript, I am happy that these were all addressed. I only have a few further detailed edits to suggest.

Thanks.

Detailed comments:

L105: are represented explicitly
Corrected (l.105)

L274: rehology --> rheology
Corrected (l. 271)

L376: I think the start of this sentence could be improved: "Such a throughput would be possible already now on the new petascale Europe's..", for example "Such a throughput would now be possible on petascale machines such as Europe's.."
The sentence was changed according to the reviewer's suggestion (l. 377).

Figure 4: The caption refers to a), b), but these labels are not on the figures. Also I think HadCRUT stands for "Hadley Centre/Climatic Research Unit Temperature".
Labels a) and b) were added to Figure 4 and HadCRUT was changed to Hadley Centre/Climatic Research Unit Temperature in the caption of Figure 4.

Figure 5: Similarly the caption refers to a), b) etc, but there are no labels.
There were actually a), b) etc labels on Figure 5, they are to be found on the lower right corner of each panel. We put the labels on the lower right corners as the upper left corners are already busy with the curves (especially on Fig. 5c).

L563: "potentially pinpointing to.." – would this be better as "potentially pointing to.."
Changed accordingly (l. 564)

L677: Altantic --> Atlantic
Corrected (l. 678)

Reviewer 2

2nd review of Hohenegger's ICON-Sapphire coupled overview paper

As I said in the previous review, I think this is a well written paper on an important topic. I'm very pleased by how the authors responded to the suggestions of both reviewers, which I think makes the paper even better. I'm particularly happy with the injection of comments on my big-

picture questions (e.g. how does/doesn't km-scale resolution help, what can we learn from a 1 yr simulation, etc). I suggest the authors are given one more chance to make minor edits in response to the comments below, with the understanding that their next revision will be immediately accepted.
Thanks.

1. Reviewer 1 L304 comment (L324 in new manuscript): Saying you didn't spin up the land because it was unclear how to do so felt a bit naïve to me, but maybe that just reflects different capabilities at different modeling centers. In E3SM we frequently spin up the land model either by running with the land model interacting with an atmosphere that's continually nudged to reanalysis for the period leading up to our target start date or by running the land model in standalone mode driven directly by atmospheric observations. These approaches are explained in the most recent CAPT overview paper: https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015MS000490 . For SCREAM, we do land-model standalone runs at the full resolution of the model because they're cheap. With nudging, we use full-resolution for the land model but coarser resolution for the atmosphere. Another approach would be just to interpolate land conditions from coarser resolution coupled simulations (though this can lead to issues with land/sea masks). I mention these approaches not necessarily for you to mention in the paper, but in case it helps you come up with a plan for future runs.
Maybe we should have added in our response that it is unclear how to spin up the land *in global climate storm-resolving simulations*. We added this nuance in the manuscript on lines 318-319. As far as we are aware, no study investigated the effect of soil moisture initialization in this type of simulations. We agree with the reviewer that starting from soil moisture conditions close to observations can improve the forecast skill for short-range simulations, as also known from storm-resolving weather forecasts, but it would be more important to know if the soil moisture initialization can affect the climate, either in terms of mean or pdf.

More importantly, I didn't get the sense from Fig R1 that any soil layers in the tropics are equilibrated (except maybe the deepest layer, but that's the one we least expect to equilibrate fast!). Do you really think your "3 of 5 soil layers in the tropics are spun up" statement is accurate? To be clear, I don't think this detail affects the integrity of the paper in any way and I'm just pointing it out to make the paper as perfect as possible. In any case, I do think it would be worth articulating more clearly what "land not spun up" means in this case – it could mean that there's no moisture in the soil at all or that soil moisture is initialized to a single value everywhere, or that a spatially-varying climatology is used.
Our assessment on whether a soil layer was spun up or not was based on an assessment of whether a persistent trend in soil moisture was visible, e.g. soil moisture decreasing throughout the simulation. In Fig. R1, it can for instance be seen that in the lowest two layers, the soil moisture keeps on increasing with time, even though not always at the same speed. Hence the soil is not spun up. In contrast, in the first layer, soil moisture decreases first, in January and February, before beginning oscillating. Hence, for that layer, we only interpret January and February as months where soil moisture is not spun up yet. And so on for the other layers. To clarify how we assessed the spin-up period, we added on line 321 "in the sense that the continuous drying of the soil since simulation start has stopped".

2. Reviewer 1's comment on Fig 5 (Connection between TOA and large scale circulation): while the atmospheric general circulation responds rapidly to a given forcing change, SST changes slowly and the atmospheric general circulation will evolve in response to those slow changes. In that context, TOA energy imbalance will definitely cause sea level pressure and zonal winds to evolve if you ran your simulation longer. I don't think there's anything to change in your paper – adding the individual observed years in the graphic is nice and you've acknowledged now that the short length of your simulations precludes definitive analysis of ocean-related variables. I just thought that the response to reviewer 1 wasn't correct.

We agree with the reviewer that we cannot rule out at the outset that there might not be a connection, but as also noted by the reviewer, we acknowledged this in our manuscript by adding the individual observed years in the graphic and by explicitly mentioning the limitation due to the length of our simulation.

3. L511 in new draft: it felt odd to me that you say there are 2 parallel bands of precipitation in the W Pacific instead of saying you seem to have a double ITCZ. It reminded me of describing a camel as a horse-like animal with 2 bumps on its back – an accurate description but more verbose than needed and harder for people to connect to. It's fine if you have a reason to avoid calling it a double ITCZ, but if not you might as well use standard terminology.

We rephrased to "This leads to the formation of a double ITCZ over the western Pacific with two parallel precipitation bands, where the southern band is too zonal, reaching too far east and remaining south of the equator even during boreal summer" (l. 491-492).

4. Reviewer 1 comment on L623: is the effect of mesoscale eddies on phytoplankton production in your high-res and coarser rest simulations in the same direction as found in Harrison et al 2018? If so, that's stronger evidence for your finding. If not, it suggests the difference may be due to chance.

We totally agree with the interpretation of the reviewer and actually the effect of mesoscale eddies on ocean biogeochemistry and especially carbon is the topic of a current PhD study, which was also one of the main reasons for conducting the ocean-only high-resolution simulation with biogeochemistry. Hence we prefer not touching upon this topic in this paper.

5. Reviewer 2 comment 20: I agree that excessive precipitation implies too much radiative cooling within the atmosphere, but I disagree that excessive atmospheric radiative cooling implies a TOA radiative imbalance where more radiation is leaving the planet. My reasoning is that the heat capacity of the atmosphere is negligible compared to the surface, so TOA imbalances are more likely to reflect a surface that is too emissive or an atmosphere that reflects (rather than emits) too much energy.

We decided to remove the corresponding sentence in the manuscript, which was "A too strong radiative cooling would be consistent with the fact that G_AO_5km is loosing too much heat compared to observations" as it was not precise enough and could lead to false conclusions. The negative net TOA imbalance in our simulation compared to observations is primarily due to too much shortwave reflection, as noted in the manuscript when discussing Fig. 1, but too much shortwave reflection would affect similarly the net TOA and surface radiation balance, meaning that it would not be visible in the atmospheric radiative cooling. Instead, an excessive atmospheric radiative cooling could be consistent with too little shortwave absorption in the atmosphere, a too cold Earth's surface or too much incoming longwave radiation at the surface.