

“The ExtremeX global climate model experiment: Investigating thermodynamic and dynamic processes contributing to weather and climate extremes”

By Kathrin Wehrli, Fei Luo, Mathias Hauser, Hideo Shioyama, Daisuke Tokuda, Hyungjun Kim, Dim Coumou, Wilhelm May, Philippe Le Sager, Frank Seltner, Olivia Martius, Robert Vautard, and Sonia I. Seneviratne

Dear author, co-authors,

First of all, sorry for the delay in this editor's decision, due a misunderstanding, on your ms submitted for publication in ESD. Having read in detail your response to the reviews as well as the suggested revisions of the manuscript, I thought that I could basically accept the revision for publication in ESD but then reading once more carefully the revised ms (including the one with track changes, note the differences in line numbers), I came still across quite some (revised) statements that were not optimally formulated/require further editing/triggered some additional questions. In addition, there were also some other issues being raised in reading the ms in terms of the overall context. Below you can find all these and would like you to handle these comments for my final decision on your paper.

Dear editor,

Thank you for carefully reading our manuscript and we appreciate your thoughts on the content. We have addressed the questions and comments raised and believe that this has augmented the quality of the paper further. The only additional change made to the manuscript is that paragraphs from the introduction have been rearranged such that we first state the research questions and scope of the paper, then mention possible further applications and finish by an overview of the content of the paper which now also indicates the section numbers. These changes in our opinion read better with the other changes made in the introduction. In the following, answers are given below the statements. Please also find with the re-submission the document with track changes (changes with respect to the previous revised manuscript from 28.02.22) and the newly revised manuscript itself.

Best regards

Kathrin Wehrli (on behalf of all authors)

Line 37: “and new, **yet unseen**, **not yet seen** extreme intensities are appearing”
Something that has not been seen yet has also not appeared..... I would suggest here to use wording; “are anticipated”

We thank the editor for this suggestion and the line was changed accordingly.

Line 51: “predecessor paper”, suggest to refer here to “**previous study**”

done

Line 58/59: “atmospheric nudging and/or **prescribed** soil moisture”

done

Line 60-61: “ This nudging approach has previously been verified for CESM (Wehrli et al., 2018) also analyzing biases....”

The sentence was changed as suggested.

Line 61: “Here, the same experiments are carried out for three ESMs”. Now that it becomes much more clear how the presented study compare to the previous studies only having used the CESM system, it would be good to stress this: “Here, building further upon the studies with CESM, we here present the results of the same experiments but carried out for”

Thank you for this suggestion, we changed the line as suggested (except for cutting the second “here” and specifying the model version CESM1.2).

Line 64/65: “Other applications have not been tested so far and will be left to explore in future studies”. Now that you have included this to address the reviewers comments, e.g., on more detailed analysis of the water balance but now this statement, especially mentioning “other applications” is way to vague. Can you indeed indicate how the ExtremeX experiment results can be diagnosed in a different manner to further unravel the role of land surface processes versus circulation (and the oceans)?

This statement refers to more detailed analysis on drivers and model biases but also to the application of the experiments to different types of events. One possible application is shown in Luo et al. (2021) using the ExtremeX experiments to study biases of anomalies in near-surface temperature, precipitation, mean sea level pressure and meridional wind speed at 250hPa during certain Rossby wave events in the Northern Hemisphere.

As we explained to Reviewer #2 and state in the conclusions, the experiment setup is not ideal to study the role of the ocean but instead simulations with interactive ocean would be more suitable if the focus was on the ocean. Also, some of the drawbacks of the setup are already mentioned in the conclusions (e.g. that the analysis is co-temporal and thus a separation of initial source vs. response is not possible). If by detailed analysis on the water balance you are referring to the comment by Reviewer #1 (the first general comment), then this is something that is not possible with the present setup. We would prefer to not mix up things that might be possible with ExtremeX with analysis that would require other experiments. Reviewer #2 had the suggestion to look into the moisture budget, this is indeed one of the “other applications” we think could be mentioned here. We rephrased the lines to better explain what we mean:

“The ExtremeX experiments could also be used to examine other types of events than heatwaves. They are suitable for more in-depth analysis of model biases by examining for example the atmospheric moisture and heat budgets or the surface energy balance. In Luo et al. (2021) the ExtremeX experiments are used to study the origin of model biases in the anomaly of upper-level winds and near-surface climatology during certain summertime Rossby wave events in the Northern Hemisphere by constructing composites. Other applications have not been tested so far and will be left to explore in future studies.”

Reference:

Luo, F., Selten, F., Wehrli, K., Kornhuber, K., Le Sager, P., May, W., Reerink, T. Seneviratne, S. I., Shigamata, H., Tokuda, D., Kim, H., and Coumou, D. (2021), Summertime

circumglobal Rossby waves in climate models: Small biases in upper-level circulation create substantial biases in surface imprint, *Weather and Climate Dynamics Discussions*, 1–30, doi: [10.5194/wcd-2021-48](https://doi.org/10.5194/wcd-2021-48), *in review*

Line 82: for consistency use the term “observation-based data”

done

Line 85: “All experiments prescribe SSTs, sea ice and vegetation..” would be good indicate here what sea-ice and vegetation characteristics are prescribed, cover? Or other properties?

Sea ice cover fraction is prescribed and the land use (vegetation type or plant functional type in CESM) is prescribed. The sentence now reads:

“All experiments prescribe SSTs, sea ice cover fraction, and land use (i.e. vegetation) but differ in the simulation of the atmospheric circulation and soil moisture that are either interactive or constrained.”

Line 92-93; had to read this sentence multiple times: suggest to change to “...experiments both components are constrained prescribing soil moisture varying over time or prescribing soil moisture using a climatological soil moisture, respectively”

We agree that the original sentence is hard to read and changed the lines based on the suggestion to say:

“... both components are constrained prescribing soil moisture time-varying or prescribing soil moisture using climatological soil moisture (but including the seasonal cycle), respectively.”

Line 95: “As ExtremeX was initiated in 2017, no longer time horizon was proposed”. Should have commented on this earlier but what is meant with this? I don’t get what you exactly want to express with this statement.

This was solely to explain why the experiments end in 2015/2016. But as it seems to cause confusion the sentence was removed.

Line 103: “observation-based products” or “observation-based data”? be consistent

We decided to replace “observation-based products” and “observation-based values” by “observation-based data” throughout the manuscript to be consistent.

Line 113: “..to generally refer to the applied method of nudging the atmospheric large-scale circulation and prescribing soil moisture”

We agree with the suggestion made and adapted the sentence accordingly.

Line 130: “2.2.2 Prescribing Soil moisture”

We adapted the section title according to the suggestion.

Line 146: having referred before to the specific method of prescribing soil moisture and then introducing here in a very general way “2.3 Reference data sets”, I would put soil moisture here under the same heading. But here you are introducing more the “Atmospheric circulation and hydrology data”

Indeed, the subchapter on soil moisture prescription is a lot about how to generate the soil moisture data that is prescribed because it is an essential part of the method description. Whereas on the other hand for the chapter on atmospheric nudging, it is more the specifications (like the relaxation time scale and the nudging profile) that have to be provided to ensure reproducibility. In the “2.3 Reference data sets” chapter we introduce the data sets that were used to validate the results of the experiments. It is true that these data sets provide information on the atmospheric circulation and hydrology. We hope that this helps to understand our choice of the structure of the manuscript and we are open for suggestions to rename and/or rearrange chapters.

Then reading the introduction again of the three models that have been included in these ExtremeX experiments, I realized that I miss(ed) a motivation on the actual of these three models. What has been the main reason that these experiments include these three different modelling systems, is there a specific reason why these models have been selected and what would be the expected differences based on the main (different) features of these models? It would be good to shortly include this in the introduction of the paper.

In the beginning we (ETH; CESM) reached out to different modeling groups that might be interested to contribute to ExtremeX. The motivation to carry out a multi-model experiment was to find out whether the results would agree. If certain phenomena/behaviour is found across all models, despite their differences, this adds to the confidence we have in the results. In the end two other modeling groups were found that were willing to run all the required experiments. Therefore, the selection of the models was in the first place due to availability of experiments and less an intentional choice of the specific models (or search for exactly these models). However, we were keen to have models that do not too closely resemble (like for example two setups of CESM). Comparing the studies by Knutti et al. (2013; for CMIP5 models) and Brunner et al. (2020; for CMIP6 models) shows that this is true for the models chosen.

Since EC-Earth3 is the most recent and was used for CMIP6, it has the highest horizontal and vertical resolution of the tree models used here and the most recent forcing data. A recent study by Brands (2022) shows that EC-Earth3 is among the best-performing models with regard to the representation of atmospheric dynamics/circulation in the Northern Hemisphere mid- to high-latitudes. MIROC5 is among the models with higher circulation errors, while CESM (comparable to the TaiESM1.0 in the paper by Brands (2022)) is in-between (for EC-Earth3 and MIROC5 this is also confirmed in the study by Fernandez-Granja et al. (2021)). Apart from that there are no fundamental structural differences between the models. Different behaviour of the models is discussed throughout the paper, we currently do not see a way to bring this to the introduction without explaining the results. The models behave differently in the CMIP experiments but as the experiments only cover the historical time period, we do not see how this might be indicative of the models' behaviour in the ExtremeX experiments.

In the introduction we added the following sentences:

“Here, building further upon the studies with CESM1.2, we present the results of the same experiments but carried out for three ESMs that were contributed each by one of the collaborating modeling groups. The models do not show high interdependence and thus are an optimal selection for a small ensemble (Brunner et al., 2020; Knutti et al., 2013). EC-

Earth3 is the most recent of the three models and, hence, has the highest horizontal and vertical resolution of the three models used here. Since EC-Earth3 was used for CMIP6, it also has the most recent forcing data. Among CMIP5 and CMIP6 models, EC-Earth3 is one of the best-performing models with regard to the representation of atmospheric circulation in the Northern Hemisphere mid-to-high latitudes (Brands, 2022; Fernandez-Granja et al., 2021). MIROC5 has contributed to CMIP5 as well as the 1.5°C versus 2.0°C global warming experiments (e.g. Hirsch et al., 2018; Mitchell et al., 2017; Shiogama et al., 2019)."

References:

Knutti, R., Masson, D., and Gettelman, A. (2013), Climate model genealogy: Generation CMIP5 and how we got there, *Geophys. Res. Lett.*, 40, 1194–1199, doi:[10.1002/grl.50256](https://doi.org/10.1002/grl.50256).

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R. and Knutti, R. (2020), Reduced global warming from CMIP6 projections when weighting models by performance and independence, *Earth Syst. Dynam.*, 11 (4), 995–1012, doi: [10.5194/esd-11-995-2020](https://doi.org/10.5194/esd-11-995-2020)

Brands, S. (2020), A circulation-based performance atlas of the CMIP5 and 6 models for regional climate studies in the Northern Hemisphere mid-to-high latitudes, *Geoscientific Model Development*, 15 (4), 1375–1411, doi: [10.5194/gmd-15-1375-2022](https://doi.org/10.5194/gmd-15-1375-2022)

Fernandez-Granja, J.A., Casanueva, A., Bedia, J. et al. Improved atmospheric circulation over Europe by the new generation of CMIP6 earth system models. *Clim Dyn* **56**, 3527–3540 (2021). doi: [10.1007/s00382-021-05652-9](https://doi.org/10.1007/s00382-021-05652-9)

Hirsch, A. L., Guillod, B. P., Seneviratne, S. I., Beyerle, U., Boysen, L. R., Brovkin, V., Davin, E. L., Doelman, J. C., Kim, H., Mitchell, D. M., Nitta, T., Shiogama, H., Sparrow, S., Stehfest, E., van Vuuren, D.P., & Wilson, S. (2018). Biogeophysical Impacts of Land-Use Change on Climate Extremes in Low-Emission Scenarios: Results From HAPPI-Land, *Earth's Future*, 6. doi: [10.1002/2017EF000744](https://doi.org/10.1002/2017EF000744)

Shiogama, H, T Hasegawa, S Fujimori, D Murakami, K Takahashi, K Tanaka, S Emori, I Kubota, M Abe, Y Imada, M Watanabe, D Mitchell, N Schaller, J Sillmann, E Fischer, J. F. Scinocca, I. Bethke, L Lierhammer Jun'ya Takakura, Tim Trautmann, Petra Döll, Sebastian Ostberg, Hannes Müller Schmied, Fahad Saeed, Carl-Friedrich Schleussner (2019) Limiting global warming to 1.5°C will lower increases in inequalities of four hazard indicators of climate change. *Environ. Res. Lett.* 14, 124022. doi: [10.1088/1748-9326/ab5256](https://doi.org/10.1088/1748-9326/ab5256)

Line 233: "...some variability in wind speed between the five members.."

We changed the line to say "... some variability in wind speed and wind direction between the five members ..." in order to stay consistent with Fig. A1 which shows wind speed and direction (and not u and v vectors)

Line 234: "...representation of the reference wind speed", to stress that this evaluation/validation step is mainly relying on comparison of wind speeds and not other meteo. properties

Again, this was changed to say "... representation of the reference wind speed and wind direction ..."

Lines 270/271: “This contradicts the initial intuitive assumption and suggests that no sole component of the model is responsible for the biases, and hence the latter cannot be corrected in isolation”. I was somewhat puzzled by what you want to express by those last three words; I guess would like to express that the biased cannot be corrected by overcoming the causes of those simulated biases by improving the model representation of only one (in isolation) of the model components”. If this is indeed what you mean, possibly rephrase (more explicitly)

I believe the guess is correct. What we meant is that the biases (that we just discussed in the section; near-surface temperature and total precipitation) cannot be (substantially) corrected by improving the representation of single model components in isolation. We rephrased the sentence and for better readability we separated the statement into two sentences:

“This contradicts the initial intuitive assumption and suggests that no sole component of the model is responsible for the biases. Hence, the climatological biases that are discussed here cannot be corrected by improving the representation of the model components in isolation.”

Line 292: “Nonetheless, for EC-Earth AF_SF shows larger biases than AI_SI as the temperature biases introduced in AF_SF outweigh the biases corrected, and for precipitation, the biases remain of similar magnitude”. What do you mean here with biases corrected?? Corrected for what?

As AF_SF is the experiment with nudged atmosphere and prescribed soil moisture, the circulation and soil moisture should be “correct” or very close to observations. Therefore, the assumption would be that near-surface temperature and total precipitation are also well represented or at least substantially better in the AF_SF simulations compared to the AI_SI simulations (thus “corrected” with respect to AI_SI). We have changed the end of this paragraph to clarify:

“Nonetheless, for EC-Earth AF_SF shows larger climatological temperature biases than AI_SI (Fig. 4), and for precipitation, the biases remain of similar magnitude (Fig. 5). This indicates that even if some part of the temperature and precipitation biases is reduced by constraining the atmospheric circulation and soil moisture using observation-based data, other biases can be enhanced or change sign, resulting in a worse overall performance of the model.”

Line 359: “which both found a weak role of the ocean in??” explaining the Russian heatwave or heatwaves in general?

Both studies examined the Russian heatwave. We changed the line to say:

“This result is supported by the studies by Dole et al. (2011) using initialized forecasts and Hauser et al. (2016) using an ESM, which both found a weak role of the ocean in explaining the Russian heatwave of 2010.”

Line 374: “by Duchez et al. (2016) finding that the SST patterns set important preconditions for the 2015 summer”, same comment here; just for the 2015 summer in general or in explaining the 2015 European summer heatwave?

Yes, again more specifically the heatwave of the 2015 summer in Europe was meant. We clarified this in the sentence:

“This is in contrast to the modeling study by Dong et al. (2016) and the observation-based study by Duchez et al. (2016) finding that the SST patterns set important preconditions for the 2015 European heatwave.”

Line 399: “This is in line with the La Niña conditions that established in 2010 and 2011 and transitioned to cool to neutral during...”, this statement does not read well and suggest to change to “This is in line with the La Niña conditions that prevailed in 2010 and 2011..” but what do you want to express with “transitioned to cool to neutral”? a change to a more cool or neutral state/contribution by La Niña to conditions in Australia?

Based on the Oceanic Niño Index (ONI) by the NOAA Climate Prediction Center La Niña lasted from about May-June-July 2010 to March-April-May 2012 (three month running means). Then ENSO remained neutral for the rest of 2012 and during the year 2013. The sign of the ONI was slightly negative during 2013 (but not strong enough for La Niña, this is what we wanted to express with “cool to neutral”). The sentence was altered to say:

“This is in line with the La Niña conditions that prevailed from mid-2010 to early 2012 and then remained neutral for the rest of 2012 and during 2013 (NOAA Climate Prediction Center, 2022) as well as with the findings by Lewis and Karoly (2013).”

Line 407: “whereas for the other two models the individual ratios are balanced to slightly circulation-dominated”, another puzzling statement, I guess you want to express that “whereas for the other two models the individual ratios reflect that there might be an even-to-a slightly circulation-dominated contribution by the two main drivers of the Australian heatwave according to the experiments”

The individual ratios show about equal contributions from circulation and soil moisture for approach A and a slightly larger contribution from atmospheric circulation for approach B. We rephrased the sentence based on the suggestion to say:

“... whereas for the other two models the individual ratios reflect that the contribution by the two main drivers of the Australian heatwave is equal to slightly circulation-dominated according to the experiments (Fig. A8d)”

Line 418: “– thus not favoring a heatwave –” a would rather say “– thus not strongly enhancing the heatwave–”

We changed the sentence to say “For the 2012/2013 heatwave in Australia, all models agree that the role of the ocean is negative – thus not enhancing the heatwave – ...” because “not *strongly* enhancing” reads a bit like it might in fact still be slightly enhancing the heatwave.

Line 442: “This shows the growing relative importance of the **land surface-atmosphere coupling** for long-duration events” (instead of only the state of the land surface, correct?)

We thank the editor for this suggestion as it indeed fits much better. The change was adopted as suggested.

Line 457: “Five experiments with varying levels of constraining were run with all models”; this is bad revised sentence, alternative; “Five experiments, with different sources of information to constrain these experiments, were run with all models”

What we meant to say was that some of the experiments are less constrained (AI_SI: only the ocean is prescribed) and some experiments are more constrained (AF_SF: ocean and soil moisture is prescribed, atmospheric circulation is nudged). We changed the existing sentence and added one sentence as follows:

“Five experiments were run with all models with one or more of the models’ components being constrained. SSTs and sea ice coverage fractions are prescribed in all experiments.”

Line 463: “even larger biases appear in the constrained experiments”, alternative also given that there were already biases in the default experiments, “the biases are even further enhanced in some of the constrained experiments”

In most cases it does not look like already existing biases are being enhanced in the constrained experiments. Instead, the magnitude of the bias is usually the same or slightly reduced (compare for example the temperature bias in Fig. 4a between AI_SI and AF_SI; for all three models the sign of the bias is the same in the two experiments and the RMSE is the same or slightly smaller in AF_SI). There are regions where the constrained experiments show biases of the opposite sign (see for example the positive precipitation bias in Europe and western Asia from 30°N to 60°N for AF_SI from MIROC in Fig. 5a or the negative precipitation bias east of the Black Sea for AI_SF from EC-Earth in Fig. 4a). Nevertheless, enhancement of already existing biases also happens, for example, negative temperature biases north of ~60°N seem to be enhanced in AI_SF in EC-Earth compared to AI_SI. To better address all of these points we changed the sentence to say:

“In some cases, biases are enhanced or even change sign in the constrained experiments (Fig. 3, Fig. 4 and Fig. 5)”

Line 468: “including atmospheric circulation and soil moisture (dynamics) interactions”

The suggested change works fine for us and was adopted in the manuscript.

Line 474: “For events that are mainly ocean-driven, we would recommend a setup with interactive ocean experiments to compute the ocean contribution more accurately.” Could you possibly shortly indicate/ hypothesize on extreme events that might strongly affected by ocean-atmosphere-driven process interactions. Would there be a role in heatwaves (in suppressing, or, enhancing) or other extreme events?

We think that this can affect heatwaves but also other types of extreme events such as floods or droughts. Essentially, all events strongly driven by ENSO, or other coupled ocean-atmosphere phenomena such as the Indian Ocean Dipole or the Pacific Decadal Oscillation would benefit from a better estimation of the ocean contribution by using interactive ocean experiments. We added the following sentence to the (new) Line 496 of the revised manuscript:

“This would apply for example to extreme events (i.e. droughts, heatwaves, floods) that are strongly driven by the El Niño Southern Oscillation and other coupled ocean-atmosphere phenomena such as the Indian Ocean Dipole or the Pacific Decadal Oscillation.”

Line 476: “..the potential ocean contribution..”

done

Line 493: “The largest contribution **by recent warming** is found for the U.S. heatwave..”

Yes, thank you, the sentence was changed accordingly.