Manuscript: WCD-2021-48

Response to the Reviewers' comments on

Summertime circumglobal Rossby waves in climate models: Small biases in upper-level circulation create substantial biases in surface imprint

By Fei Luo, Frank Selten, Kathrin Wehrli, Kai Kornhuber, Philippe Le Sager, Wilhelm May, Thomas Reerink, Sonia I. Seneviratne, Hideo Shiogama, Daisuke Tokuda, Hyungjun Kim, and Dim Coumou

We thank both reviewers for their insightful and positive comments. We find all the suggestions constructive, and we are delighted that both reviewers found this article interesting and of scientific importance. We address each comment point by point below. The reviewers' comments are given in **black** and our responses in **blue**.

Reviewer #1

General comments

In this manuscript, the authors assess the representation of wavenumber 5 and 7 patterns in three different climate models (EC-Earth, CESM and MIROC), and search for the reason of model biases with respect to the ERA5 reanalysis data. They find that the models represent these wave patterns reasonably well, however, small biases in the upper level circulation lead to large biases in surface variables, like temperature, precipitation and mean sea level pressure. They show a significant improvement of model performance in case the upper level circulation is nudged based on the observed one. In contrast to previous studies, the soil moisture plays only a minor role in the representation of high amplitude wave events, which can be also a consequence of the chosen weekly time scale.

The paper has a very clear structure. The scientific message is clear, useful and very relevant, it can help to improve climate models and produce more realistic future climate scenarios. The language of the paper is understandable, except some (actually, a bit too many for a submitted manuscript...) grammatical and typographical errors. Based on the importance of the message of the paper, I suggest the manuscript to be accepted for publication in WCD, however, only after the authors have taken care of the below mentioned deficits.

My main criticism concerning this work is that it concentrates only on wavenumber 5 and 7 patterns, and there is no critical discussion related to the applied methodology. I understand that previous studies have found these patterns to be relevant for simultaneous extremes, and it is important to test whether the models can represent them or not, but it would increase the scientific value of the paper if the introduction, at least, presents summertime circulation anomalies and related surface extreme events from a broader perspective. At the end, extreme events can be observed during other wavenumbers as well, thus it is important, that models are able to reproduce a wide range of wave numbers not just the mentioned 5 and 7. Although the authors show the spectrum for a range of waves in their results, this is only very shortly discussed in the paper.

We agree with the reviewer that including analyses of more wavenumbers will improve the scientific value of the manuscript. In the revised manuscript we will provide additional figures of waves 4-8 to provide a broader range of wavenumbers relevant for extremes.

In this work, the model experiments are compared to ERA5 reanalysis data, which is supposed to represent reality. Reanalysis data have, however, their biases and deficits too, which should be mentioned in the paper as well. The relatively short period of ERA5 can lead to additional biases of the statistical estimation – it can lead to an under- or over-sampling of certain wavenumbers, as shortly mentioned in the manuscript too. A short discussion of these issues would increase the degree of objectiveness of the paper. A critical evaluation of the used technical tools – Fourier decomposition of the atmospheric field, nudging - in the discussion part would be beneficial as well and would increase the scientific quality.

Overall, this is a good paper with an important scientific message, however, the scientific quality needs to be improved before publication mainly by

1) increasing the accuracy of the wording and presentation of the results,

2) improving the objectiveness related to the used methods and data set, and

3) the work should be put in a broader scientific context.

Thank you for the list and suggestions. Some key aspects of our responses and revisions regard the above suggestions are:

1) We will revise the manuscript carefully to improve the overall accuracy and scientific presentations of the results.

2) We will add one subsection in “Data and Methods” to include the details on the atmospheric nudging including its vertical profile. We will provide a critical discussion on limitations of ERA5.

3.a) For the main concern from Reviewer #1 about presenting that models can reproduce a broader range of different waves numbers and also their related surface extreme events. We decided to add one paragraph on this topic in the introduction part during the revision process.

3.b) We will add one small subsection in “Discussion” to discuss the limitations of the methods we used such as Fourier decomposition of the atmospheric field and nudging to improve this paper’s scientific quality. We will also discuss the potential biases in ERA5 data to improve the objectiveness of this paper.

Specific comments

Figure 2: The probability densities are smoothed. This is not mentioned in the manuscript nor is the smoothing procedure and bandwidth. This information should be included, and the non-smoothed histograms should be shown as well to give a realistic picture of the involved uncertainties.

Thank you for this comment, we have included the bandwidth information in the caption in Figure 2 as shown below. We also constructed the histograms for Figure 2.

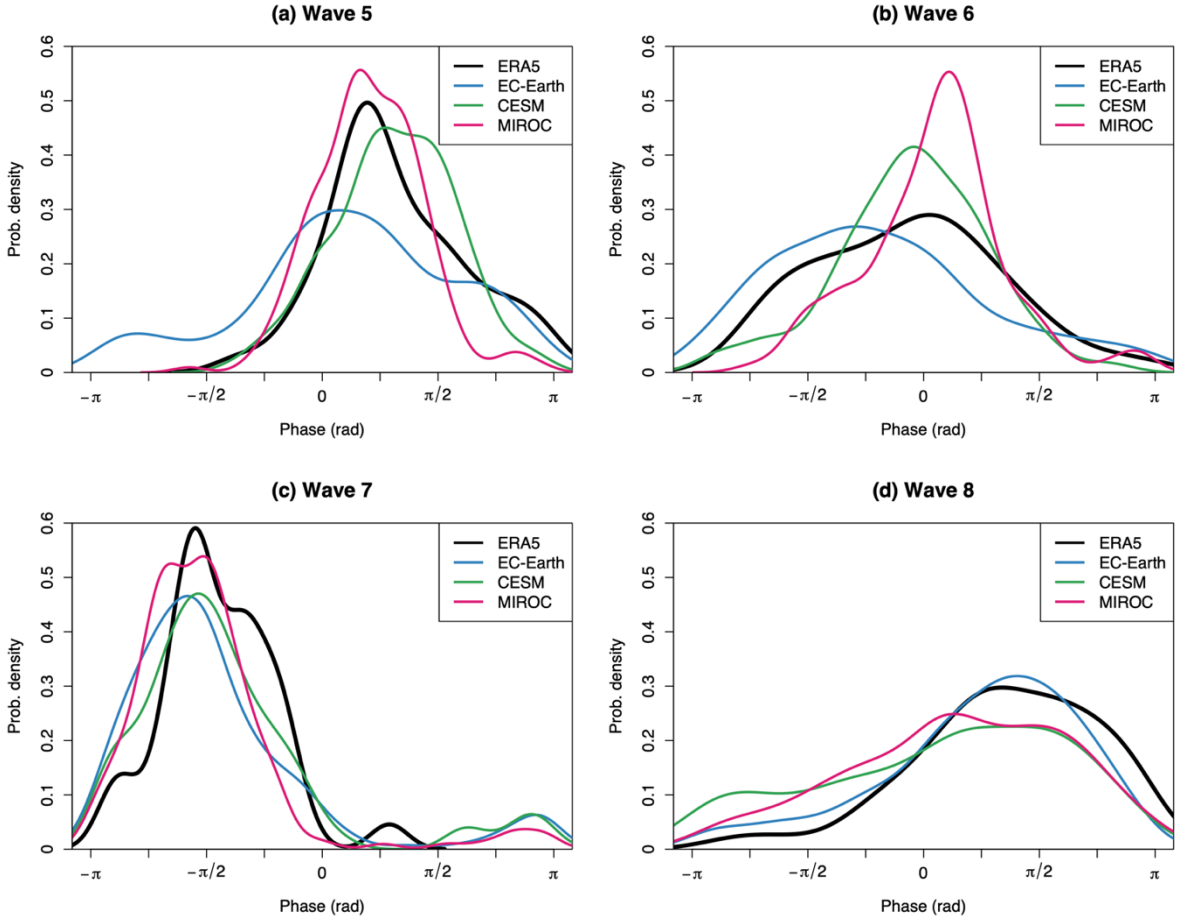
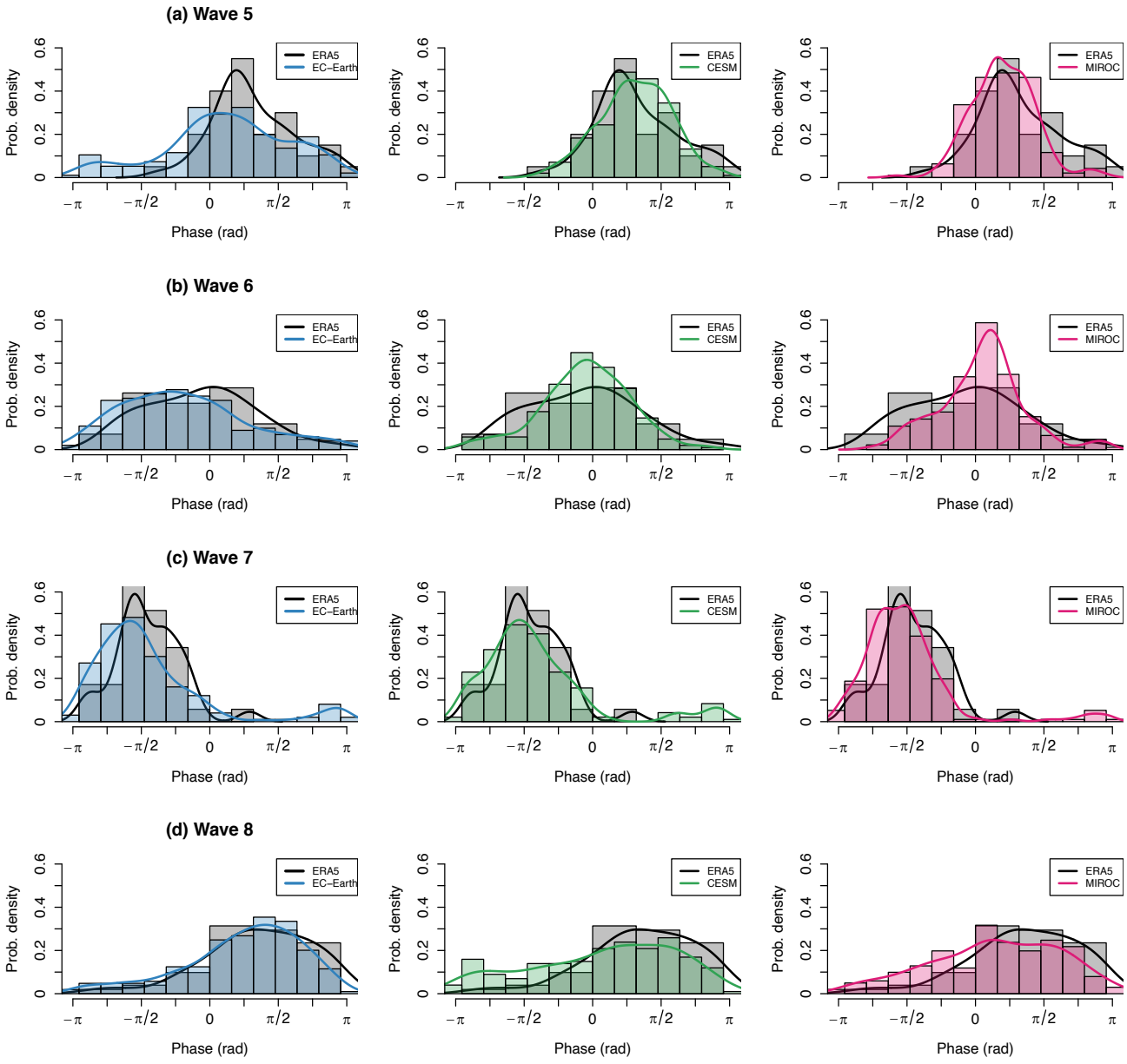


Figure 2: Phase-locking of Rossby waves for JJA ERA5 and model waves 5-8 in control run AISI for high wave amplitude events (> 1.5 s.d.): (a)-(d), Probability density functions of the phase positions of waves 5-8 in ERA5, EC-Earth, CESM, and MIROC during JJA for the period of 1979-2015/2016 (wave 5 (a), wave 6 (b), wave 7 (c), wave 8 (d)). The bandwidth for ERA5 and models are as follows: (a) Wave 5: 0.35(ERA5), 0.40(EC-Earth), 0.25(CESM), 0.22(MIROC), (b) Wave 6: 0.53(ERA5), 0.45(EC-Earth), 0.30(CESM), 0.25(MIROC), (c) Wave 7: 0.25(ERA5), 0.29(EC-Earth), 0.27(CESM), 0.22(MIROC), (d) Wave 8: 0.49(ERA5), 0.39(EC-Earth), 0.52(CESM), 0.46(MIROC).

Below is the figure where the histograms of ERA5 and model data are shown for high amplitude wave-5 to wave-8 phase-locking distributions: (a) wave 5, (b) wave 6, (c) wave 7, and (d) wave 8. We think it's not ideal to plot everything together (i.e. all the model and ERA5 histograms in one figure), as that way one cannot differentiate between datasets anymore, hampering the interpretation of the figure. Thus, we prefer to show the histograms between ERA5 and the models one by one, and we will add this figure in Appendix.



L47: "*These persistent weather extremes can have disastrous impacts on human health and societies such as wide spread crop failure, infrastructure damage and properties loss, especially when they are defined as compound events.*"

The word "defined" is inappropriate here: The impact is a consequence of the manifestation of a real extreme event, it does not depend on how the event is "defined".

Thank you, we agree on the above suggestion and we changed the original phrase to "especially when these extreme events co-occur".

It is not mentioned in the manuscript what kind of observational data set is used for the nudging. This is, however, relevant information, which should be included.

Thank you for this comment. Below is the vertical atmospheric nudging profile taken from the Figure 1 in a recent work by Wehrli et al. (2021, in review). More information on the atmospheric nudging setup and dataset used can be found in this paper. The winds dataset used for atmospheric nudging are from ERA-Interim reanalysis (Dee et al., 2011).

We will add this information in the revised manuscript.

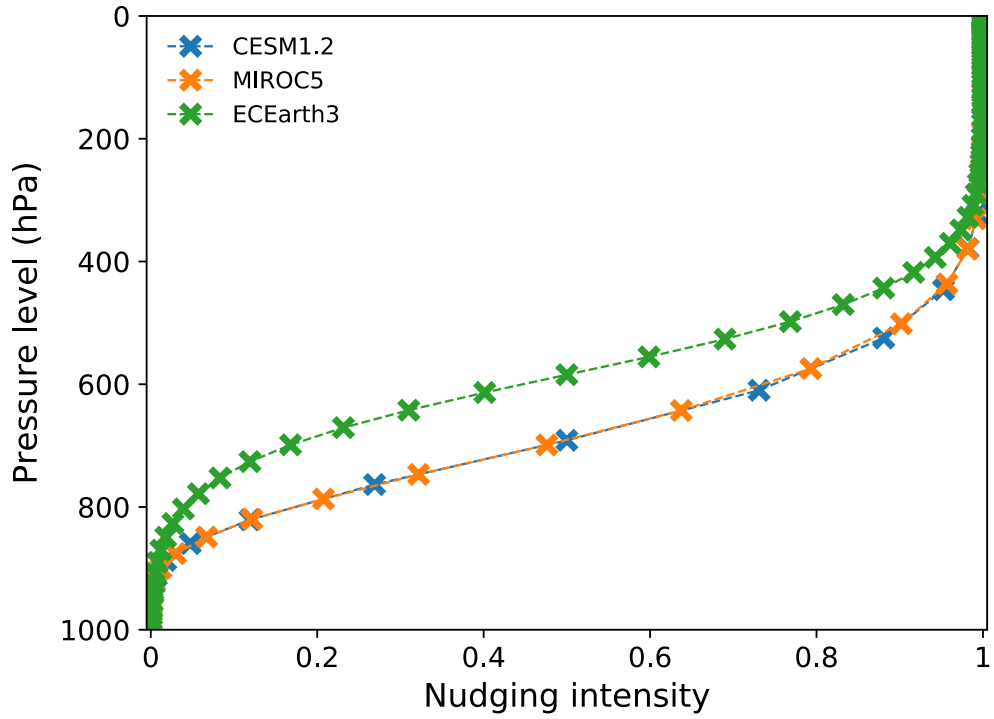


Figure 1. Nudging profile for the three ExtremeX ESMs. The actual pressure levels are marked with an x and joined with lines. The nudging intensity is given from zero (no nudging) to one (fully nudged) (Wehrli et al., 2021, in review). Taken with permission.

L172: "*ERA5 shows the peak for both the wave amplitude and variance at wavenumber 5*". According to Fig. 1, the peak in the spectrum of ERA5 is at wavenumber 6. Please clarify this.

Thank you for spotting this, indeed it should be "at wavenumber 6". The typo is corrected accordingly.

It is not clear why in the explanation of the Taylor diagram in L204, the root mean square error is mentioned twice.

Thank you, the second "the centered RMSE" is deleted.

L288-L290 "*The observations are complex and location specific as one component within a climate model might be tuned in such a way that it compensates for biases in other components. If so, nudging only one component might not necessarily reduce the overall biases, in this case, prescribing only soil moisture part.*"

The message here is not expressed clearly. It should be rephrased, explained more clearly and in a more elaborated way.

Thank you, the sentence is rephrased as follows:

"The aforementioned observations are location specific as one component within a climate model might, erroneously, be tuned in such a way that it compensates for biases in other components of the climate model. If so, nudging only that component would not reduce the overall bias. In this case, prescribing only soil moisture part does not guarantee the reduction of overall bias."

L295-L296: *"...are well captured in different climate models in terms of their climatology, variability, and phase-locking behaviors.*"

The phase locking behavior is captured by the analysed climate models, however I wouldn't call it "well captured" based on what Fig. 2 shows.

Thank you for pointing this out, we changed "phase-locking behaviors" to "The phase-locking bevaior is captured."

L327-L328: "*whereas in our soil moisture prescription experiments the soil moisture was prescribed with values from running the land component driven by atmospheric fields from reanalysis in the model offline, which thus represent much smaller forcings*."

The formulation is not clear enough, it should be rephrased.

Thank you, the sentence is rephrased as follows:

"whereas in our prescribed soil moisture experiments, the soil moisture is not set to zeros. Instead, soil moisture is set to more realistic values coming from the models' land component forced by atmospheric fields from reanalysis. Our experiment thus represents much smaller forcings than prescribing the soil moisture to zero as done in Teng et al. (2019)."

L356: "*increase from 0.71 and 0.99 to 0.63 and 1.06*". An increase from 0.71 to 0.63? Please clarify.

Thank you for spotting this, it should be "increase from 0.71 and 0.63 to 0.99 and 1.06".

Technical corrections

L37: "*When applying both soil moisture prescription and the nudging of upper-level atmosphere, both the correlation and n.s.d. values are quite similar to only atmosphere component is nudged experiments.*"

Grammatically incorrect, please rephrase.

Thank you, the sentence is rephrased as follows:

"When applying both soil moisture prescription and the nudging of upper-level atmosphere, both the correlation and n.s.d. values are quite similar to the experiments where only the atmosphere component is nudged."

L50: "Röthlisberger et al. 2019" instead of "Röthlisbergera et al. 2019".

Thank you, the typo is corrected accordingly.

L57: The citation "Hoskins and Ambrizzi 1993" appears twice.

Thank you for spotting this, the typo is corrected accordingly.

L72 and L74: "*some have analyzed*" and "*Some studies by Branstator at al...*" Formulation too general.

We implied the citations (Garfinkel et al., 2020; Wills et al., 2019) and Branstator et al. (2002 & 2017) as "some have analyzed" and "Some studies by Branstator" here. But we do think the suggestion above is helpful, thus we changed the original sentences as follows:

"Although studies such as Garfinkel et al. (2020) and Wills et al. (2019) have analyzed waves in models, their focus is not on summer and, also, they have not explored the phase-locking behavior of amplified, quasi-stationary Rossby waves. Furthermore, most studies have not analyzed waves above wave number 6. Studies by Branstator et al. (2002 & 2017) have also looked into models but focus on seasonal means and/or winter."

L192: "*Then we obtained the occurrences for JJA wave-5 and wave-7 events during 1979 to 2016 for ERA5 are 8.1% and 7.1%*". Grammatically incorrect.

Thank you, we deleted "Then we obtained" and changed "occurrences for" to "occurrences of". The sentence is now as follows: "The occurrences of JJA wave-5 and wave-7 events during 1979 to 2016 for ERA5 are 8.1% and 7.1%".

L212: "Taking ERA5 data as reference" instead of "taking ERA5 data as references".

Thank you, changed as suggested.

L283: “completely” instead of “completed”

Thank you, changed as suggested.

“And” is not a long word and can be written out, it does not need to be replaced by the symbol “&”.

We couldn’t find the “&” symbol in L283. Did you mean the symbol “&” in L283 or in general throughout the whole paper?

L293 “some extreme events” Formulation too general.

Thank you, we deleted the word “some”, and rewrote as “extreme events such as heatwaves and heavy precipitation”.

L340: *“... that persist more than 2 weeks in summer events*” Please rephrase.

Thank you for the suggestion, we rephrased the whole sentence as follows:

“In the summer, when the wave-5 or wave-7 events persist more than 2 weeks, the average reduction in crop production is 4% and even up to 11% on regional level (Kornhuber et al., 2020).”

L346: “flow” instead of “slow”

Thank you for spotting this, the typo is corrected accordingly.

L347: “large-scale circulation pattern” instead of “large circulation pattern”

Thank you, changed as suggested.

Reviewer #2

This study compares model errors with a set of nudging experiments using multiple AGCMs. They find nudging the upper-level circulation can significantly reduce surface biases, while nudging soil moisture can't lead to much improvement. This is a very interesting study that definitely deserves to be published. However, the interpretation and presentation may need some extra efforts. Below I list several concerns, mostly minor, for consideration during the revision process.

Thank you for the positive feedback and we will consider the suggestions for our revision process.

Total field or anomalies. I get confused at different parts of the manuscript whether they were talking about the total field or the anomalies of v250. Wavenumber 5 and 7 may be more prominent and longitudinally phase locked in the total field, but it's unclear whether this is also true for the anomalies. I am afraid that for subseasonal variability and extreme events, what matters more is the anomaly, not the total field. I can't help wondering why they chose to only focus on events that project strongly onto wavenumber-5 and -7 of the total field.

Thank you for this comment, we have decided to do the follows in the revision process:

- We will expand to wave Nr. 4 to Nr.8 as we responded to Reviewer #1
- We prefer to work with absolute v250 values rather than anomalies as the former are easier to interpret. Still we will add the same figures using anomalies of v250 to the appendix

It would be far stretched to use the current experiments to address Question 3, whether model biases originate from the atmosphere circulation or land surface- feedbacks. First, prescribing soil moisture does not necessarily mean the model can accurately simulate land-atmosphere feedback. Secondly, it seems that the nudging they applied to the atmosphere circulation (above 700 hPa) is far more than just "the upper atmospheric levels". Therefore, they may need to rephrase the conclusion that "small bias in the upper atmospheric levels can result in big bias in surface weather conditions" (ln 305). See 4. for another concern on drawing this conclusion.

We agree with the reviewer that the current experiments cannot fully conclusively differentiate between the source of biases. Still, our experiments show relatively minor biases in upper level circulation propagate in models to create larger biases in surface variables. For completeness, we add the vertical atmospheric nudging profile which shows that nudging starts around 700hPa but only with a very weak nudging strength. The nudging strength increases gradually in the upward direction and full nudging is only applied above ca. 400hPa. Thus, we will reword 'upper atmosphere' into 'mid-to-upper atmosphere' as it is important to highlight that the planetary boundary layer is free to adjust in the nudged experiments.

Why are the pattern correlation coefficients in Figs. 5,6 all positive? What's the geographical domain used to construct the Taylor diagrams?

Because the Taylor diagram (Taylor 2001) calculates the field correlation coefficients between ERA5 and each model.

It can be seen in Fig.5 and Fig.6 that the spatial patterns are quite similar (highly correlated), thus the correlation coefficients are all positive. It would imply that the model fields are in the opposite phases as ERA5 to have negative coefficients (i.e. v250 has a positive sign in ERA5 whereas in EC-Earth to have a negative sign), which is not the case.

The geographical domain used to construct the Taylor diagram is the same as in Fig.3 and Fig.4.: the full longitudinal band between 35N to 60N.

Why unlike other three models, CESM (Fig.B1) B_land is not much smaller than B_atm? In fact, why are the magnitudes of B_tot, B_atm, B_land and B_res rather similar?

We don't agree on this point, as if we look at Fig. 5 and Fig. 6, B_land is different from B_atm. This is also shown quantitively in the statistical values in Table A1 to A4. Of course, the architect of each climate model is different, thus it's not surprising that models differ on smaller scale.

Significant test for the composite maps in Figs.3,4 is needed.

Thank you for the suggestion, we think this is a good idea and we will implement the significant test for the composite maps in Fig.3 and Fig.4 during the revision process.

The manuscript needs serious editing.

Thank you, we will revise the manuscript carefully during the revision process.

References

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, et al. "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System." *Quarterly Journal of the Royal Meteorological Society* 137 (656): 553–97. https://doi.org/10.1002/qj.828, 2001.

Röthlisbergera, M., Frossard, L., Bosart, L. F., Keyser, D., and Martius, O.: "Recurrent Synoptic-Scale Rossby Wave Patterns and Their Effect on the Persistence of Cold and Hot Spells." *Journal of Climate* 32 (11): 3207–26, doi: 10.1175/JCLI-D-18-0664.1, 2019.

Taylor, K. E.: "Summarizing Multiple Aspects of Model Performance in a Single Diagram." *Journal of Geophysical Research Atmospheres* 106 (D7), doi: 10.1029/2000JD900719, 2001.

Wehrli, K., Luo, F., Hauser, M., Shiogama, H., Tokuda, D., Kim, H., Coumou, D., et al.: "The ExtremeX Global Climate Model Experiment : Investigating Thermodynamic and Dynamic Processes Contributing to Weather and Climate Extremes," no. July: 1–31, doi:10.5194/esd-2021-58, 2021.