Reply to Reviewer 2, August 2023

Thank you for the thorough review and the suggestions to improve the manuscript! In the following I will respond to the comments in blue. Not all responses are full-fledged, but rather indicate how to implement the required changes in the revised manuscript. I will start revising the manuscript accordingly after the editor's decision to proceed further.

Summary

This manuscript evaluates boreal summer season mean surface fluxes, soil moisture, land surface temperature (LST) and precipitation in two EC-Earth3 earth system model (ESM) simulations – one using the land-surface component only and one coupled to the atmospheric component. The objective of the work is to assess the quality of the land-surface component and to determine the impact of coupling on the simulation of surface climate.

Review

In order to make the best use of ESMs for climate research, it is critical to understand how well they simulate the real Earth climate. The current work evaluates EC-Earth3 and the results would be informative for scientists looking to use that model to answer research questions. In particular the current work focuses on the impact of using the coupled configuration versus the land-only configuration on biases in the simulation of surface climate. This is an important consideration, as it is necessary to understand whether biases in the land-surface model propagate through the coupled model and how they may be enlarged or supressed by the model representation of land-atmosphere interactions.

However for a paper with the title "the role of land-surface interactions for surface climate in the EC-Earth3 earth system model" and the aim of evaluating biases globally I find the usefulness of the current results limited due to the sole focus on boreal summer and the slim selection of observational datasets used for evaluating LST and precipitation.

Land-atmosphere coupling in the southern hemisphere is maximised in boreal winter, and I suspect the biases in surface climate will be also. An analysis for multiple seasons (at least June-August for the Northern hemisphere and December-February for the Southern hemisphere) is required. The analysis would also benefit from the inclusion of additional observational datasets including monthly mean satellite LST (infrared and microwave for comparison) and at least one additional precipitation dataset.

Yes, I think the observation on the choice of the season is correct. I had chosen to present only season to limit the number of figures (which already is quite large) and motivated the focus on JJA in the text with the domination of the Northern and Southern Hemisphere. I will extend the analysis to also include DJF, focussing on the Northern Hemisphere (10 °S – 80 °N) in JJA and the Southern Hemisphere (60 °S – 10 °N) in DJF. Please find more details below.

I think it is correct that observations cannot be a "perfect" representation of the true climate state and that, therefore, several independent observational data sets should be considered. For the land-surface temperature, we face the problem that may of the satellite-derived estimates consider only retrievals under clear-sky conditions, limiting the choices. For precipitation, there are several options based on gauge data, satellite retrievals or combinations thereof, but in this case the spatial scale of the data must be taken into account to avoid differences that are due to the different representations of orography or coast lines. I will include an additional data set for land surface temperature and precipitation. Please find more details below.

I would therefore recommend that the current manuscript be revised before being considered for publication. Please find my specific comments below.

Specific Comments

Focus on boreal summer

Line 300 is the first line in the main text that mentions the study focuses only on boreal summer (June – August). The study focus should be reflected in the abstract, introduction and possibly the title, with the justification provided at the start of the manuscript, not in the Discussion.

Yes, I agree. The choice of the JJA season should be more obvious throughout the manuscript.

On that note, I am not convinced by the reasoning provided on Lines 587-593. If the objective is to understand the effect of atmospheric coupling on surface climate biases globally the work should consider the different seasons when the coupling is maximised in the different regions.

As mentioned above, I will extend the analysis to also include DJF, focussing on the Northern Hemisphere (10 °S – 80 °N) in JJA and the Southern Hemisphere (60 °S – 10 °N) in DJF. Not presenting the whole globe (60 °S – 80 °N) in the maps is necessary for not increasing the number of figures. At the same time, the maps need to be combined with the histograms, unless the histograms can be omitted as such. Moreover, I consider omitting the figures with the averages for the IPCC-regions (Figs. 15, 16 & S12) and Table S1, because the regions are located in different hemispheres and, thus, represent different seasons.

Datasets – LST

The author does not provide a justification for the exclusion of satellite LST products in the data section. Instead the author includes an inaccurate statement in the results section:

"A serious shortcoming of existing satellite-based data sets for land-surface temperature is that they depend on clear-sky conditions and, thus, do not incorporate periods with partly or fully cloudy conditions".

This is actually only true for infrared satellite observations. Microwave satellite observations (for example AMSR2 skin temperature) do not depend on clear sky conditions. Furthermore the limitation of infrared satellite observations does not prevent a comparison when considering seasonal mean values. You could apply a mask to the MODIS Terra/Aqua monthly mean daytime LST products to exclude land points with less than a minimum number of valid observations (clear sky days) contributing to the value (for example 10 days in the season).

Therefore it is possible to include satellite LST products in the evaluation rather than rely solely on reanalysis.

One final query on this point, why did the authors use LST from ERA5 rather than ERA5-Land?

That is because ERA5 has the lower resolution, which is closer to the resolution of the model. Using the high-resolution data from ERA5-Land would introduce some artificial biases in areas with high and steep different orography. See, for example, Figure 4 in Muñoz-Sabatier et al. (2021).

AMSR2/GCOM-W1 surface soil moisture (LPRM) L3 1 day 10 km x 10 km ascending V001 (contains microwave skin temperature) DOI:10.5067/B0GHODHJLDA8

MODIS/Aqua Monthly mean Day-Time Land Surface Temperature at 1x1 degree V005 DOI:10.5067/2YCD3NSNDMRM

MODIS/Terra Monthly mean Day-Time Land Surface Temperature at 1x1 degree V005 DOI:10.5067/4SI45J6G6BW5

The two doi's point to the monthly mean day-time LST at a resolution of 1° from the Aqua and the Terra satellites, respectively. Data from Terra are available for the period March 2000 to June 2015 and data from Aqua for the period August 2002 to June 2015, resulting in only 12 years of data. But there are also the corresponding data sets for the night-time LST. Although combing the four data sets seems to overcome much of the issue with the clear-sky conditions (Chen et al. 2018), Muñoz-Sabatier et al. (2021) noticed remaining problems in regions with permanent cloud cover sauch as the tropical rain forests. I will include the combined MODIS LST data in the study, having the limitations in mind when interpreting the differences.

Datasets – precipitation

Were other datasets aside from GPCC considered? For example CHIRPS?

It would be more informative if more than one dataset were included for comparison.

The author needs to justify their choice of datasets in the data section.

https://www.chc.ucsb.edu/data/chirps

I am aware that different data sets of precipitation can give somewhat different estimates, even for long-term climatologies, given the differences in how these data sets are assembled. For the purpose of this study, I hadn't considered using another precipitation dataset, mainly because I did not consider the precipitation bias not very essential for the study. The main intention was to indicate a potential bias in the precipitation in ERA5, which had been used as forcing in the offline simulation.

Unfortunately, the CHIRPS "global" data (starting in January 1981) do not cover the entire globe but only the land areas between 50 °S and 50 °N. Also, they are at a very high resolution of 0.05°. The GPCP monthly data (starting in January 1979), the other hand, are at a rather low resolution of 2.5°, whereas the high-resolution (0.5 °) GPCP daily data do not start before June 2000. I will try different options and include the best suited in the manuscript.

Abstract

The abstract could be simplified to make it more clear and concise for the reader. For example I find this line from the Discussion makes the aims of the work immediately apparent:

“the intention is to assess a) the quality of the land-surface component and b) the effects of the coupling with the atmosphere”

Compared to the equivalent line from the abstract which is much longer and includes model details that could be reserved for later:

“The aim of this study is twofold, first to evaluate the quality of the simulation of surface climate by the land-surface component of the EC-Earth3 ESM, combining the HTESSEL land-surface model and the LPJ-GUESS dynamic vegetation model, and second to assess the role of the coupling of the land surface with the atmosphere for the simulation of the surface climate in EC-Earth3.”

Yes. I agree. I will rephrase this part accordingly.

Minor point on Line 21-22, I assume one of the instances of “underestimate” should in fact read “overestimate”?

Yes, this is correct.

Introduction

In paragraph 2 the author states that EC-Earth3 represents a distinct step forward, then proceeds to list the flaws with EC-Earth3. How do the biases in EC-Earth3 represent an improvement compared to the earlier version of EC-Earth?

Yes, I realize this is somewhat contradictory. I will consult Döscher et al. (2022) for details on the improvement compared to the previous model version.

There is a distinct lack of references in paragraphs 3-5 although many different processes are discussed. Here are a few examples that could be included in paragraph 3:

K. L. Findell, E. A. B. Eltahir, Atmospheric controls on soil moisture-boundary layer interactions. Part I: Framework development. J. Hydrometeorol. 4, 552–569 (2003)

R. A. Pielke, Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. Rev. Geophys. 39, 151–177 (2001)

Bhowmick, M. and Parker, D.J. (2018) Analytical solution to a thermodynamic model for the sensitivity of afternoon deep convective initiation to the surface Bowen ratio. Quarterly Journal of the Royal Meteorological Society, 144, 2216–2229.

Thank You for pointing this out and providing some references. I will include additional references in relation to the different processes mentioned.

Concerning paragraph 3, the concepts are a little disorganised. The paragraph is predominately discussing the surface impact on the atmosphere, yet ends with a sentence on the atmospheric impact on the surface. Also the discussion on surface impacts on the atmosphere jumps between temperature and humidity/precipitation couplings rather than discussing each in turn.

I tried to keep this part concise, apparently this means that the line of thought got disturbed. I will rephrase and reorganise the paragraph.

Equations

The flux equations, currently in the results section, should be included in the data section.

Yes, I agree that the equations should be part of Section 3.1. However, in response to the other reviewer's comment I am thinking to drop these equations.