2nd Review of the paper "Non–linear intensification of Sahel rainfall as a dynamic response to future warming" by Jacob Schewe and Anders Levermann

The authors carefully took into account and precisely responded to all my comments, and the manuscript results improved in comparison with the previous version. However, some major issues are still open, specifically concerning the robustness of the results and the final conclusions of the study. I find the results presented in this paper very interesting, with potential impact in the understanding and modelling of climate change in the Sahel. Therefore, I recommend that the authors fix these issues before the paper is published. See details below.

Major comments

a) Robustness of the results.

1) Model selection. The authors select the Wet7 models on the basis of the projected precipitation at the end of the 21st century. Three models are selected because of the magnitude of the response (>100%), four models are selected on the basis of the spatial pattern, even though some are "drier" than other models not selected. E.g. it can be noticed that CanESM2 and NorESM1-M are "drier" than HadGEM and ACCESS models (see Figure 1). I find the selection method not robust because no quantitative approach is presented to measure similarity in patterns. Moreover, the reader cannot verify whether not selected models show more or less similar patterns. I suggest to present a robust quantitative method to select the "wet" model subset.

2) Domain selection. In my previous review, I requested to describe the methodology to select the domain, but the authors' response is incomplete. Although the choice of the boxes appears quite obvious (especially for the oceanic domains), and small differences in size/position should not produce large differences, I think that the domain selection should not just be based on "substantial" differences. Does "substantial" mean that differences are statistically significant? Or above a threshold? This should be clarified, not to give the impression of a subjective choice. Moreover, presenting relative differences does not help much. Indeed, for some models the highest differences are outside the boxes, in deserted areas where even very small absolute differences mean large relative differences (see Figure 5).

3) Rainfall-SST relationship. When presenting the precipitation-GMT relationship (Figure 11), authors state that non-linearity is not reproduced as in the precipitation-SST plots (Figures 7 ad 8). Actually, I cannot see such a large difference between the two cases. Also in this case I recommend the statement to be supported by a quantitative demonstration. A simple functional shape for the precipitation-SST relationship could be hypothesised and tested. This method could also help to identify the specificity of the Atlantic and Mediterranean basins, in driving the increase in precipitation, compared to the global warming.

4) Comparison with "neutral" and "dry" models. In order to understand the specificity of the Wet7 subset, a comparison with the "neutral" and "dry" model subsets would be ideal. Indeed, the reader has no mean to understand whether the Wet7 models project such large anomalies because of their ability in simulating the precipitation-SST non-linearity, or because they are just “warmer” than the others. In other word, does the precipitation-SST non-linearity hold also in “neutral/dry” models? Or are these model just not “warm” enough?

b) Conclusions. While the objective of the study is clearly stated in the Introduction, the conclusions (and the title as well) are quite misleading concerning the achievements of the study. The reader could have the impression that this study explain how precipitation in the Sahel will increase in the 21st century, but this is not the case. Indeed, on the one hand, the authors clearly state that they aim to identify the key mechanism leading some models to be wetter than others. On the other hand, they conclude that “this explanation of an abrupt intensification of inland monsoon rainfall in the Sahel region is consistent with studies suggesting a substantially wetter Sahel, and Sahara, region in past climates compared to today (DeMenocal et al., 2000; Gasse, 2000). It is also consistent with theories linking rainfall changes in the Sahel to a combination of a local (through radiative forcing changes) and a remote (through tropical SST impacts on atmospheric stability) forcing mechanism (Giannini, 2010; Giannini et al., 2013; Seth et al., 2010)”; and that “the mechanism we suggest here would act on top of these two mechanisms, and help explain the abruptness of the Sahel rainfall response to global warming. It would particularly affect the more continental parts of the region”. They finally reconcile with the initial objective stating that “consideration of this mechanism may help to make sense of the diversity of model projections, and eventually establish a more consistent understanding of the Sahel’s future climate in a warming world”. I recommend the authors to make clear in the Conclusions that their results focus on the mechanisms underlying the projections of a “wet” model subset, not implying that these models are the good ones.

Minor comments

Page 2, line 18: “pronounced rainfall increase north and west”, I see the rainfall increase north and east actually.

Figure 1: I don’t see the discussion on the ability in modelling of drought crucial for the paper, which basically focuses on trends rather than multidecadal oscillations. Moreover, the discussion on the ability to reproduce multidecadal oscillations is rather concise and does not add much to the discussion. Even the ability of the Wet7 subset in simulating the big drought is questionable. Indeed, though Wet7 are above the average, FGOALS-g2 simulates almost no drought, and CanESM2, BNU-ESM and NorESM1-M are just slightly dry. I suggest to remove this figure, because not crucial for the model selection.

Figure 2: The MIROC5 ability in simulating the multidecadal variability is quite good in the second half of the century, and less good in the first half. Moreover, the assessment is only qualitative, not supported by any metric. Similarly to the comment above, I suggest to remove this argument to support the model selection.

Figures 4-6: Statistical significance of the differences should be assessed.

Figures 7, 8, 9 and 11: XY axes should share the same range, for better comparison.

Conclusions: The inclusion of the Mediterranean role in the discussion is helpful for understanding the main idea of the study in a broader context. I think that this could be emphasised in the conclusions.

Figure S2: is it the same as Figure 3?

Figure S3: in caption, reference should be to Figure 4 (not Figure 2).

Figure S4: it is redundant with Figure 9. I see more useful to show the evolution of rainfall along with the Mediterranean SST and the GMT (similarly to Figure 9).