Reply to the comments of referee 1 for Second Revision of HESS-113

Thank you for revising the manuscript. As in the part 1 of this study, there are still some remaining comments that need to be addressed. Please see the comments of reviewer #2, especially the comment related to significance test. Please make sure to provide response to each of their comment and specify what changes were made or not, in the manuscript.

Major/Specific Comments:

This is the third revision of HESS 113. Overall I find Part-II to be interesting, but I don't feel that the authors fully responded to my comments in my second revision. Also, as noted in my review of HESS 112, while I appreciate the authors sending the manuscript for editorial review, I think the authors still need to read over the paper for some missed edits, some of which are noted below. Overall my recommendation is minor revisions as again I think these are mostly cosmetic suggestions, but please note that the authors should carefully respond to each comment.

1. ***Comments:***

In my second review, comment 2, I stated: I'm still having some issues/concerns with your significance tests. Almost the entire region is significantly different in most of your Figures, to the point where it would be easier to show areas where there is NO significant difference. You did use a 90% confidence level, perhaps this is too low given your datasets? Is your sample size too low to form a proper significance test?

The authors response did not fully answer the questions posed. One could simply remove the significance test at this point, but if not, then do the authors feel that 90% CL is appropriate here, and why? In addition, I feel like your sample size is far too low to perform a proper significance test, but I'm unsure of the sample size used as I do not see it stated.

Thank you very much. As agreed in Part 1 of this study, the main shortcoming was that we performed the significance test with monthly values leading to samples of small size. In this revised version, instead of doing the Student t-test with monthly means, we did it with daily values (from June to September) for each year (2003 and 2004) and thus, with samples of 115 days (without the 7 days spin-up period).

The two paragraphs below were added (lines 127-141) at the end of the section 2.1 (Model description and numerical experiments) to introduce the experiments design and the Student t-test as follows:

In the part 1 of this study, we designed three experiments (reference, wet, and dry), each with a set of five (5) simulations starting from June 1st to September 30th. The difference between these three experiments is the change in the initial soil moisture condition (reference initial soil moisture condition, wet initial soil moisture condition, and dry initial soil moisture condition) during the first day of the simulation (June 1st, 2001, 2002, 2003, 2004, and 2005) over the West African domain. Then, we selected the two years most affected by the wet and dry initial soil moisture conditions (2003 and 2004) to estimate the limits of the impact of the internal soil moisture forcing on the new non-hydrostatic dynamic core of RegCM4.

For these two years most sensitive to soil moisture initial conditions, the Student t-test is used to compare the significance of changes in climate extreme indices between a wet or dry sensitivity test (sample 1) and the control (sample 2) in assuming that this method performs well for climate simulations (Damien et al., 2014) and knowing that it is extensively used for climatological analysis (Menedez et al., 2019; Talahashi and Polcher, 2019). In this study, the t-test at the 95% confidence level was used to consider statistically significant.

2. ***Comments:***

There is an exclamation point on line 156 after the word "neighboring"

Thank you. In this revised version, this sentence has been deleted. A new paragraph on the Student t-test has been added at the end of the section 2.1, as mentioned above (in the answer to the first comment).

3. ***Comments:***

Line 191: "The TRMM datasets underestimate..." Compared to CHIRPS?

Thank you for the comment. In agreement with your comment 5 (below), this sentence has been deleted in this revised version. We agree with your suggestion and remove the comparison between CHIRPS and TRMM to ease the reading and to be more focus on the aim of the study.

4. ***Comments:***

Line 193 should read "The strongest underestimation was found over the central Sahel..."

This sentence referred to the comparison between the two precipitation products. As said above, it has been deleted in this revised version.

5. ***Comments:***

Given the authors response to my note in my second review, comment 1, on the use of the two observational datasets TRMM and CHIRPS, I'm unsure why this comparison is indeed necessary for this particular manuscript. If in fact CHIRPS is considered more appropriate for extremes in West Africa with the given references, why re-assess it here? It may clarify things to remove this comparison, unless the authors think it is strictly necessary to include. If the authors do think that this comparison is necessary to include, then very clear reasoning should be provided.

Thank you for the comment. We agree with your suggestion and we removed the comparison between CHIRPS and TRMM to ease the reading and to be more focus on the aim of the study. In this revised version, we used CHIRPS products and we introduced them as follows (Lines 147-149): ***We have chosen CHIRPS as reference in this study, mainly because this product has been widely assessed and used for the study of extreme events in West Africa by Bichet et al. (2018a, b) and Didi et al. (2020).***

6. ***Comments:***

Are there similar references that discuss the ability of the temperature datasets over West Africa? I could see this as a reason for including the comparison if there are no other references, but again this does not seem to be the focus of this particular manuscript.

Thank you for the comment. You're right. In this revised version, we removed the comparison between the temperature datasets to be more focus on the aim of the study. While the temperature product used remains the same, we changed its name in this version (CPC-T2m instead of common name GTS temperature dataset) to be more specific. The following lines have been added to introduce the T2m product (Lines 150-157): ***We validated the 2-m temperature using the NOAA/NCEP/CPC daily maximum and minimum global surface air temperature. The NOAA/NCEP/CPC global daily surface 2-m air temperature (CPC-T2m) is a land-only gridded global daily maximum (Tmax) and minimum (Tmin) temperature analysis from 1979 to the present, available at two spatial of 10 min × 10 min and 0.5° × 0.5° (latitude × longitude). This product provides an observational T2m estimate for climate monitoring, model evaluation, and forecast verification (Fan Y. and Huug van den Dool, 2008; Pan et al., 2019). In this study, the daily Tmax and Tmin are used at spatial resolution 0.5° × 0.5°..***

7. ***Comments:***
   As in Part-I, I suggest again noting that these experiments are done in a highly-idealized framework and are intended to show the potential impact of very strong soil moisture conditions on extremes, and should thus be used as a guide or first look at the influence of soil moisture on extremes.

Thank you very much. The following paragraph has been added at the end of the conclusions section (lines 566-572) as follows:

This study is the first to investigate the impact of soil moisture initial conditions on climate extreme indices over West Africa. These experiments were done in a highly-idealized framework and were intended to show the potential impact of very strong soil moisture initial conditions on climate extremes. Consequently, it should be considered as a first overview of the influence of initial soil moisture on climate extremes with a RCM (RegCM4). In perspectives, this study will benefit from being performed in a multi-model framework with several RCMs within CORDEX-Africa initiative (Coordinated Regional Downscaling Experiment).