Responses to Anonymous Reviewer 2

General Comment:

This is an interesting paper assessing the impact of the GCM as compared to the RCM on a domain in southern Africa. However, I think there are some concerns and questions that need to be addressed before the paper is ready for publication.

Variance between the models (RCMs or GCMs) has been employed to compare the model performance for the simulation of the monthly precipitation. In my reading, this study aims to show how much each RCM output is affected by its driving GCM. However, some concerns that may be critical should be addressed if the information provided in this study will eventually be used by both climate and non-climate scientists, as the authors mentioned.

RESPONSE: We would like to thank the Anonymous Reviewer #2 for the positive interpretation of the manuscript. Based on the suggestions and comments, we provide the following replies.

1st Comment:

RCM output can be affected not only by the input GCM datasets but also by its parameterization, configuration options, and setup. A simple comparison of the variance of monthly data between the complex models does not guarantee whether the impact of the driving GCM is dominant. All possible outcomes or limitations should be discussed to show the relationships between the RCMs and GCMs.

RESPONSE: Thank you for this comment. Indeed, in a multi-RCM/multi-GCM set of simulations there are multiple sources of uncertainty that may eventually strongly influence the climatological output of a single variable, let alone of a specific atmospheric process studied in a climate-scale context (such as the climatological expression of the Angola Low pressure system). Therefore, monthly variance of a specific variable examined in a multi-RCM ensemble, cannot be attributed solely to the different GCMs that provide the lateral boundary conditions. By no means do we imply that the comparison of variances for monthly precipitation between the driving GCMs and the respective RCMs guarantees that the impact of the GCM is dominant (or not) in all possible aspects. In the revised version of the manuscript, we make clearer statements concerning the limitations and underlying assumptions of our work, however, it would be impossible to technically discuss all possible outcomes or limitations. For example, assessing the impact of the cumulus parameterization scheme on all RCM simulations, would require that all RCMs participating in the CORDEX-Africa ensemble, perform a series of sensitivity runs with -at-least-two cumulus parameterization schemes and make the model output available to the climate modeling community on a database such as ESGF or the Climate Data Sore. Unfortunately, this is not the case. Each research group participating with model runs to the CORDEX-Africa ensemble makes a series of modeling choices that may or may not introduce considerable amounts of uncertainty in the final model simulations. These kinds of studies are performed within specific Flagship Pilot Studies (FPS[1]), which are beyond the scope of the current work. We agree however, that the way through which we frame the argument about the impact of the driving GCMs on RCM simulations may cause some misunderstanding. So, we make clearer statements about some main limitations of our work. These statements are made in the last paragraph of the *Discussion* section and are the following:

"*Lastly, it is imperative to highlight that the impact of the lateral boundary conditions on RCM simulations comprise only a portion of the potential sources of uncertainty in the CORDEX-Africa ensemble examined, therefore attributing entirely the variance of RCM simulations to the driving GCMs would be erroneous. Therefore, we mention that uncertainty in RCM simulations can have a plethora of sources that are mainly categorized as parameter or structural uncertainty (Günther et al., 2020; Howland et al., 2022). These types of uncertainty sources may relate to the parameterization schemes employed by each RCM or assumptions and numerical choices involved in the dynamics of each specific RCM. However, since within CORDEX-Africa only a limited number of variables is being made available to the community, it would be impossible to meticulously*

---

[1] https://cordex.org/experiment-guidelines/flagship-pilot-studies/endorsed-cordex-flagship-pilote-studies/

comment on all possible sources of uncertainty and access the impact of their variance on monthly precipitation."

In addition, the following sentence is now introduced in the first paragraph of the Discussion section:

"*Our work examines monthly precipitation variance caused by the lateral boundary conditions and does not examine parameter and structural uncertainty separately in the multi-RCM and the multi-GCM ensembles analyzed.*"

2nd Comment:

How the relaxation zone has been defined to simulate each RCM should be addressed. Bias can be amplified through the lateral boundary conditions.

RESPONSE: Thank you very much for this comment. We have now included this information for all RCM used.

CCLM4-8-17.v1 details are provided in Dosio and Panitz, (2016), according to which 10 grid points are used on every side of the domain. Relaxation is performed using the Davies scheme (Davies, 1976, 1983).

RCA4 solution is relaxed towards the forcing boundary data across an eight-point wide relaxation zone following the Davies' boundary formulation method (Davies, 1976), with a cosine-based relaxation function.

REMO2009.v1 also uses (Jacob et al., 2012) an eight-point wide relaxation zone, following the Davies' boundary formulation method also (Davies, 1976).

Since the relaxation zone used in all three RCMs employs the same method and since eight (RCA4 and REMO2009.v1) and 10 grid points (CCLM4-8-17.v1) are used, it can be assumed that the relaxation zone is not introducing a substantial bias capable of affecting monthly precipitation climatology, trends, and climate change signal, that varies among the three RCMs.

The following portion has now been added to the *Data* section:

"All RCMs employed a relaxation zone which was either 10 grid-points wide (CCLM4-8-17.v1) or eight points wide (RCA4.v1 and REMO2009.v1). Relaxation in all RCM simulations was performed using Davie's method (Davies, 1976, 1983)."

3rd Comment:

The reasons why the three RCMs have been chosen should be addressed in the data section. If they have shown good performance in the domain, I suggest adding related references.

RESPONSE: The reason for including these three RCMs was not related to their performance over southern Africa, but to the calculation of variances, as stated within the manuscript in the *Data* section: "*More specifically, the CORDEX-Africa simulations selected are those that were driven by more than two GCMs…*".

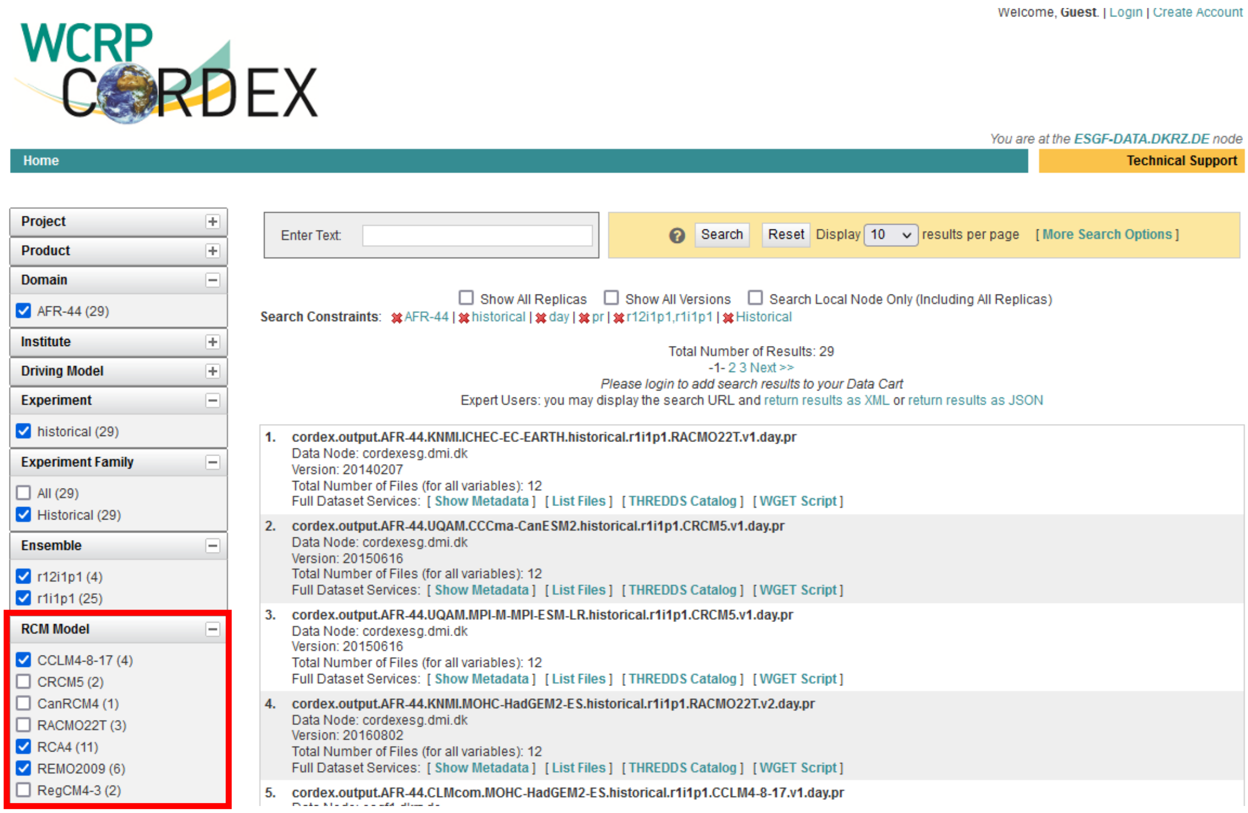
More specifically, in order to examine the intra-GCM and intra-RCM variance we select all RCM simulations that are driven by at least two (2) GCMs, using the sample variance (n-1), as indicating in the following equation:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The sample size required for an acceptable calculation of variance is a non-trivial statistical issue (e. g. https://stats.stackexchange.com/questions/7004/calculating-required-sample-size-precision-of-variance-estimate). Even using three (3) RCM simulations (three simulations of the same RCM receiving lateral boundary conditions from three different GCMs) can be problematic, however, using the "at least three driving GCMs" rule provides an acceptable compromise for the purposes of this analysis.

Based on our search in the ESGF-DKRZ database (https://esgf-data.dkrz.de/search/cordex-dkrz/), the RCM simulations that are driven with more than two (2) GCMs are: CCLM4-8-17, RCA4, REMO2009, as displayed in the figure below (last accessed at 25/7/2022).

The RACMO22T simulations were performed using lateral boundary conditions from HadGEM2-ES and EC-EARTH (r1i1p1 and r12i1p1) for the historical simulations and from HadGEM2-ES and EC-EARTH (r1i1p1) for the RCP8.5 simulations, hence it did not meet the >2 GCMs criterion and was therefore excluded from the analysis.



Domain
☑ AFR-44 (29)

Experiment
☑ historical (29)

Experiment Family
☐ All (29)
☑ Historical (29)

Ensemble
☑ r12i1p1 (4)
☑ r1i1p1 (25)

RCM Model
☑ CCLM4-8-17 (4)
☐ CRCM5 (2)
☐ CanRCM4 (1)
☐ RACMO22T (3)
☑ RCA4 (11)
☑ REMO2009 (6)
☐ RegCM4-3 (2)

☐ Show All Replicas ☐ Show All Versions ☐ Search Local Node Only (Including All Replicas)

Search Constraints: ✖AFR-44 | ✖historical | ✖day | ✖pr | ✖r12i1p1,r1i1p1 | ✖Historical

Total Number of Results: 29
-1- 2 3 Next >>
Please login to add search results to your Data Cart
Expert Users: you may display the search URL and return results as XML or return results as JSON

1. **cordex.output.AFR-44.KNMI.ICHEC-EC-EARTH.historical.r1i1p1.RACMO22T.v1.day.pr**
Data Node: cordexesg.dmi.dk
Version: 20140207
Total Number of Files (for all variables): 12
Full Dataset Services: [Show Metadata] [List Files] [THREDDS Catalog] [WGET Script]

2. **cordex.output.AFR-44.UQAM.CCCma-CanESM2.historical.r1i1p1.CRCM5.v1.day.pr**
Data Node: cordexesg.dmi.dk
Version: 20150616
Total Number of Files (for all variables): 12
Full Dataset Services: [Show Metadata] [List Files] [THREDDS Catalog] [WGET Script]

3. **cordex.output.AFR-44.UQAM.MPI-M-MPI-ESM-LR.historical.r1i1p1.CRCM5.v1.day.pr**
Data Node: cordexesg.dmi.dk
Version: 20150616
Total Number of Files (for all variables): 12
Full Dataset Services: [Show Metadata] [List Files] [THREDDS Catalog] [WGET Script]

4. **cordex.output.AFR-44.KNMI.MOHC-HadGEM2-ES.historical.r1i1p1.RACMO22T.v2.day.pr**
Data Node: cordexesg.dmi.dk
Version: 20160802
Total Number of Files (for all variables): 12
Full Dataset Services: [Show Metadata] [List Files] [THREDDS Catalog] [WGET Script]

5. **cordex.output.AFR-44.CLMcom.MOHC-HadGEM2-ES.historical.r1i1p1.CCLM4-8-17.v1.day.pr**

Figure 1: ESGF-DKRZ database search at 25/7/2022.

The statement in the *Data* section is now changed from:

“More specifically, the CORDEX-Africa simulations selected are those that were driven by more than two GCMs and for which there are runs available for both the historical and the future period under RCP8.5.”

to:

“More specifically, the CORDEX-Africa simulations selected are those that were driven by more than two GCMs **(at least three simulations available using the same RCM driven by at least three different GCMs)** and for which there are runs available for both the historical and the future period under RCP8.5.”

4th Comment:

l57 - The authors should perhaps also cite the RCM lateral boundary papers focusing on southern Africa authored by Ditiro Moalafhi, given the focus there is the impact of changes to lateral boundary conditions, with bias assessed based on a reanalysis dataset. These papers are:

1 Moalafhi, D. B., Sharma, A., Evans, J. P., Mehrotra, R. & Rocheta, E. Impact of bias-corrected reanalysis-derived lateral boundary conditions on WRF simulations. Journal of Advances in Modeling Earth Systems (2017).

2 Moalafhi, D. B., Sharma, A. & Evans, J. P. Reconstructing hydro-climatological data using dynamical downscaling of reanalysis products in data-sparse regions–Application to the Limpopo catchment in southern Africa. Journal of Hydrology: Regional Studies 12, 378-395 (2017).

3 Moalafhi, D. B., Evans, J. P. & Sharma, A. Influence of reanalysis datasets on dynamically downscaling the recent past. Climate Dynamics 49, 1239-1255 (2017).

4 Moalafhi, D. B., Evans, J. P. & Sharma, A. Evaluating global reanalysis datasets for provision of boundary conditions in regional climate modelling. Climate Dynamics 47, 2727-2745, doi:10.1007/s00382-016-2994-x (2016).

RESPONSE: We thank Reviewer 2 for the suggested papers, which we have read. All four (4) papers provide interesting aspects to the discussion of the impact of the lateral boundary conditions on regional climate modeling, however, they are relevant to primarily model-specific issues and not to the broader discussion developed within the CORDEX community, with regards to the added-value of regional climate modelling (relative to the driving GCMs) and to how the RCM-GCM simulation matrix could be optimized, in order to advance the understanding of the research community on basic-science climatological issues and to also provide reliable tools for climate-services.

More specifically, in Moalafhi et al., 2017a (1), it is stated that there are "*… inconsistencies between the impact of the bias correction prior to downscaling and the resultant model simulations after downscaling. Mean and standard deviation bias-corrected WRF simulations are, however, found to be marginally better than mean only bias-corrected WRF simulations and raw ERA-I reanalysis-driven WRF simulations. Performances, however, differ when assessing different attributes in the downscaled field. This raises questions about the efficacy of the correction procedures adopted.*"

This statement summarizes the challenges posed by the selection of lateral boundary conditions on regional model simulations and also, the sensitivity of the assessment process to the specific atmospheric variable that is assessed. In such variable-specific assessments, a specific variable may display improved performance (prior and before to dynamical or statistical downscaling), however this may happen in isolation to the rest atmospheric variables. In brief, it is common that certain model-specific choices yield the right answer for the wrong reason. There are not absolute methodological choices on how to avoid such modeling "traps", however, a good practice is to perform a process-based assessment. In our work, we use the selected subregions and the specific months of the rainy season, as a proxy of the attributes of specific region-specific and month-specific climatic features. In such a way, we aim on performing an indirect process-based assessment of the impact of the lateral boundary conditions on RCM simulations. Therefore, we think that our work and the work of Moalafhi et al., 2017a (1) has fundamental differences on the methods applied and, on the conclusions drawn.

Moreover, Moalafhi et al., 2017b (2) is a model evaluation study in which the ERA-Interim reanalysis dataset is dynamically downscaled using WRF, with an emphasis on hydrological applications over the Limpopo catchment. Although the conclusions having applications for hydrological modeling are interesting, the dynamical downscaling of a reanalysis dataset such as ERA-Interim, is a standard procedure in regional climate modeling. This work addresses different research questions compared to those set in the current paper under review.

Furthermore, in Moalafhi et al., 2017c (3) WRF is driven by two different reanalysis products, ERA-Interim and MERRA. Although the context of Moalafhi et al., 2017c (3) is different to that described in the current paper under review, still its main conclusion on the impact of the lateral boundary conditions on the RCM simulations is in accordance with the context in which the current work is performed. For this reason, Moalafhi et al., 2017c (3) is used as a citation in Line 72.

The sentence from:

“Still, uncertainty arising from both the driving GCM and the downscaling RCM affect the final product… ”

Is now changed to:

“Still, uncertainty arising from both the driving GCM (Moalafhi et al., 2017) and the downscaling RCM affect the final product (Nikulin et al., 2012)… ”

Lastly, in Moalafhi et al., 2016 (4), five reanalysis datasets are evaluated with the purpose to identify the optimal lateral boundary conditions dataset. The work of Moalafhi et al., 2016 addresses very different research questions compared to those set in the current paper under review.

5th Comment:

l89 - I would like to point the authors to papers on correcting lateral and lower boundary variables focusing on Australia, where the focus was the representation of drought, and the RCM used was WRF. While here the authors are using an ensemble of GCMs, these papers focused more on what the representation of different attributes in the lateral and lower boundaries did to the overall monthly outcomes. These papers are:

1 Rocheta, E., Evans, J. P. & Sharma, A. Correcting lateral boundary biases in regional climate modelling: the effect of the relaxation zone. Climate Dynamics 55, 2511-2521, doi:10.1007/s00382-020-05393-1 (2020).

2 Rocheta, E., Evans, J. P. & Sharma, A. Can bias correction of regional climate model lateral boundary conditions improve low-frequency rainfall variability? Journal of Climate 0, null, doi:10.1175/jcli-d-16-0654.1 (2017).

Of special interest to this study is the first paper that assessed the progressing deterioration in the corrections as one focused deeper into the domain (away from the relaxation zone). I think the authors have missed this entire volume of work as I see no references to these papers. Please also note the different levels of impact lower versus lateral boundaries end up having on the simulations.

RESPONSE: Line 89 (and 88) to which this comment refers states the following:

> "*In this work we aim to assess whether it is the RCMs or their driving GCMs that dominate monthly precipitation climatology, monthly precipitation bias and climate change signal over SAF.*" We do not claim that GCMs control entirely monthly precipitation climatologies, biases and climate change signal, but rather, we aim to identify the dominant agent (between RCMs and driving GCMs).
>
> Both papers suggested above, involve technical methodological details about bias correcting the lateral boundary conditions provided by GCMs as input to the RCM simulations. The work currently under review exploits the dynamically downscaled RCM simulations performed within CORDEX-Africa. None of the RCMs analyzed here employs statistical downscaling methods or a prior bias correction of the lateral boundary conditions providing information to the RCMs.
>
> In addition, the papers proposed for citation above, refer to a different study region, with a different morphology and coastline, which is affected by entirely different large scale circulation patterns, than southern Africa. The impact of the domain set-up and size is very different between the two study regions and therefore we think that the results drawn for Australia are not transferable to the region of southern Africa.

6th Comment:

I also urge the authors to read the additional more recent papers:

1 Kim, Y., Evans, J. P., Sharma, A. & Rocheta, E. Spatial, temporal, and multivariate bias in regional climate model simulations. Geophysical Research Letters 48, e2020GL092058 (2021).

2 Kim, Y., Rocheta, E., Evans, J. P. & Sharma, A. Impact of bias correction of regional climate model boundary conditions on the simulation of precipitation extremes. Climate Dynamics 55, 3507-3526, doi:10.1007/s00382-020-05462-5 (2020).

Here the focus was on extremes, which were found to be impacted to a greater extent by the lateral boundary corrections, than the overall monthly attributes. Please also note the more recent of the two papers that attempted to quantify the impact of multivariate dependence bias in the lateral and lower boundaries, noting that the lack of this plays a significant role in the over quality of simulations.

I must confess that I am an author to the above papers and leave it to the authors (and editor's) judgement whether my suggestions above are essential to the present study. However, in my reading of the current paper, I did feel the above referenced works do add to the story the authors are trying to tell, as they focus on a similar domain (Moalafhi) and altered lateral boundaries (all).

RESPONSE: The 2nd paper suggested (Kim et al., 2020) has now been cited in the *Introduction* section (4th paragraph).

This sentence:

"For instance, when there is a strong large-scale circulation signal that is introduced to an RCM domain (e.g. advective mid-latitude storms), it is quite likely that the RCM will be able to reproduce the information that is received at its lateral boundaries."

Has now been changed to:

*"For instance, when there is a strong large-scale circulation signal that is introduced to an RCM domain (e.g. advective mid-latitude storms), it is quite likely that the RCM will be able to reproduce the information that is received at its lateral boundaries, **however, the GCM's impact on the RCM simulation may also vary depending on how far a region lies from the RCM domain boundaries (Kim et al., 2020).**"*

7th Comment:

l172 - Please clarify if this is for the mean or the monthly series.

RESPONSE: In Lines 168-170 the following is mentioned:

> “*As an exploratory method of inspecting the differences between each RCM simulation from its respective driving (GCM) for* ***monthly precipitation*** *during both the historical and the future period, we subtract the downscaled precipitation field (*RCM_{DRI}*) from its driving (DRI), as in* ***Eq. 1****:”*
>
> In line 172 to which this comment refers we have now added the following:
> “If *DIFF>0* **(monthly precipitation)**…**”**

8th Comment:

l184 - There seems to be a mistake in the notation here, or also in equation 2. Please check. GCMi should refer to the variance of all RCMs driven by GCMi? Also, what is N?

RESPONSE: N refers to the number of available simulations contributing to either the inter-RCM or inter-GCM variance

The following changes have been made to the equations, to make notation clearer:

From:

$$RCM_{var} = \frac{1}{N_{RCM}} \Sigma_{RCM_j} \left(P_{ij} - \underline{P_j}\right)^2 \qquad \textbf{Eq. 2}$$

To:

$$RCM_{var} = \frac{1}{N_{RCM}} \Sigma_{RCM_j} \left(P_j - \overline{P}_j\right)^2 \qquad \textbf{Eq. 2}$$

From:

$$Inter\_RCMvar = \frac{\sum GCM_I}{N} \qquad \textbf{Eq. 3}$$

To:

$$Inter\_RCMvar = \frac{\sum GCM_j}{N} \qquad \textbf{Eq. 3}$$

From:

$$GCM_{var} = \frac{1}{N_{GCM}} \Sigma_{GCM_i} \left(P_{ij} - \underline{P_i}\right)^2 \qquad \textbf{Eq. 4}$$

To:

$$GCM_{var} = \frac{1}{N_{GCM}} \Sigma_{GCM_i} \left(P_i - \overline{P}_i\right)^2 \qquad \textbf{Eq. 4}$$

9th Comment:

l187 - similar confusion about the notation as earlier - please check and correct.

RESPONSE: This has been addressed in the previous comment. Thank you!

10th Comment:

Figure 2 and 3 - It would have been nice to also show the observed climatology in this figure.

RESPONSE: Thank you for this suggestion! We have now added a panel where precipitation climatology for the months of the rainy season (Oct-Mar) is shown for: ERA5, CHIRPS, CRU, UDEL, and MSWEP.

11th Comment:

Remaining figures and ANOVA analysis - I found this quite comprehensive and interesting to read. I realize the authors are already reporting a lot of information here, but I was interested on their comments on the following:

(a) What was the impact on temperature simulations and how the intra-RCM variances there compared against the precipitation?

(b) There is little focus on variability, although the main advantage an RCM brings is the added variability in both space and time. Could the authors comment on within grid variabilities across the RCMs and change in variability in time at each grid cell?

(c) Our results showed significant impact on precipitation extremes (from altered lateral and lower boundaries). It would be interesting if the authors could comment on this aspect of the RCM simulations compared to the GCM simulations.

RESPONSE: Thank you for this comment.

(a) In our work, the variable in concern is precipitation and the degree to which our conclusion can be generalized for other variables is a topic of further research. However, similar work has been performed for other study regions within CORDEX:

- Boberg F and Christensen J H 2012 Overestimation of Mediterranean summer temperature projections due to model deficiencies Nat. Clim. Change 2 433–6; -
- Christensen J H and Boberg F 2012 Temperature dependent climate projection deficiencies in CMIP5 models Geophys. Res. Lett. 39 L24705
- Sørland SL, Schär C, Lüthi D, Kjellström E (2018) Bias patterns and climate change signals in GCM-RCM model chains. Environ Res Lett 13:. https://doi.org/10.1088/1748-9326/aacc77).

b) Indeed, RCMs do increase the interannual variability of monthly means compared to their driving GCM. Also, RCMs -due to their higher horizontal resolution- increase the spatial variability of a field, especially precipitation which displays strong spatial heterogeneity. The increased spatial variability in RCMs can also be observed through visual inspection in the maps of monthly precipitation climatology and on maps displaying climate change signal. However, this work was concerned about analyzing climate mean values over a period indicative of the "current climate" and over a future period under RCP8.5. However, we do agree that investigating the modulation of interannual variability from the GCM(i) to the RCM(i) would be an interesting point to further study.

c) We do agree that investigating the degree to which precipitation extremes in RCMs are affected by the driving GCMs is also a very interesting topic. However, considering that precipitation extremes are expected during the core rainy season (DJF) during which precipitation is mainly caused by large-scale circulation that is significantly affected by the driving GCMs, this would introduce an additional level of complexity and it would perhaps comprise a complete analysis on each own.