Dear Editor,

Thanks for your satisfaction with our response, and your time to review it. We would also like to thank both reviewers for their helpful and useful comments on the manuscript.

Per your request, we have acknowledged both reviewers in our acknowledgement section in the revised manuscript. A point-by-point reply to the comments, and a marked-up manuscript version are also provided at the end of this letter.

Thank you in advance for reviewing the manuscript again, we look forward to your further advice.

Best regards
Minchao Wu (on behalf of all authors)

Below is our response to the reviewers' comments, following the structure:
(1) comments from Referees
(2) author's response,
(3) author's changes in manuscript are in “”
Comments from anonymous Referee #1

Recommendation: Accepted with major revisions

General Comments:

This paper investigates the impacts of the model formulation and resolution on the ability of two Swedish RCMs to simulate precipitation in Africa. The two RCMs were used at 200, 100, 50 and 25 km resolutions and one of them has two different formulations. This experimental setup allows disentangling the improvements related to either the resolution or the model formulation. The topic is of interest and relevant for the RCM community and deserve to be considered for publication. However, I am not sure that the journal Earth System Dynamics is the best journal to convey this study since I very rarely read RCM papers from that journal. I let the editor to decide whether the topic of this paper is suitable for this journal or not. The paper is very well written and the literature review is very good although the introduction could include more papers related to the topic. Few papers suggested below could be added in the literature review of the introduction. The abstract is generally fine, but few sentences are not clear and should be improved. The introduction is generally clear and interesting, but it should be improved to emphasize the full motivation of the analysis. The methodology is appropriate to address the objectives of the study, but I am concern about the relevancy to run an RCM at 200 and 100 km and the utility of those simulations in the paper. The results are interesting and address the objectives raised at the beginning of the paper. The figures are clear and support the analysis. The conclusions are in line with the analysis and are of interest for the community. Thus, I recommend this paper to be accepted with major revisions.

Major Comments:

1. Introduction: The introduction is interesting and fully explain the motivation of the study. However, it is a bit short and it lacks a more complete literature review of the challenges to simulate precipitation over Africa. Thus, I recommend to extend the paragraph from the line 98 to 117 in 2 or 3 paragraphs to include more RCM studies that paid attention to the challenges to simulate precipitation in Africa with RCMs.

Here is a short list giving examples of papers that could be added to the literature review in the introduction: ...

Response: We completely agree with this comment and extended Introduction by including more RCM studies. There are really many RCM-based evaluation studies for Africa and we focus mostly on studies with large RCM ensembles.

2. Methodology: Even with the warnings at lines 183-184 and 488-490, I really wonder if it is relevant to use an RCM at 200 and 100 km resolutions and I also wonder if the use of those simulations adds substantial information to the paper. I think that 200 and 100 km are
excessively far from the RCM range of resolution or comfort zone for which it is configured and calibrated and I think that little is gained from those simulations in this paper. Thus, either the authors should be really convincing that those resolutions are relevant and add substantial content to the paper or either they should remove at least the 200 km resolution simulations from the paper. In some sense, the 100-km resolution simulations may be relevant since they are at a resolution close to ERAInterim and can be used as a “no added-value experiment”. Additionally, by removing the 200 km resolution simulation, only an aggregation to 100 km would be necessary for the analysis, leading to more details of the simulated precipitation in the results section.

Response: From the beginning, our experiment was developed to include simulations at coarse resolution outside of a RCM comfortable zone following experiment setup in Moufouma-Okia et al. (2015) with the coarsest resolution - 150km for their RCM (HadGEM3-RA). Our point of view is that such coarse-resolution simulations are a useful supplement to simulations at RCM comfortable resolution and help us to understand RCM behaviour without additional, resolution-dependent tuning. Our results show that performance of the RCA4 and ALADIN RCMs at 200km is, in general, consistent and fits well to what can be expected for moving from the highest (25km) to coarsest (200km) resolution. This, for example, includes among others i) a common tendency to precipitate less at coarser resolution for both RCA4 and ALADIN in JAS (Fig. 2) and ii) the deterioration of simulated daily rainfall intensities with decreasing resolution (Fig. 6). We also found different behaviour of RCA4 and ALADIN with decreasing resolution in DJF (Fig. 3). This shows that the impact of coarser resolution on the simulated precipitation climatology is not the same in different seasons and regions and depends on RCM formulation. We would prefer to keep the coarse resolution simulations as the study becomes less complete if the 200km simulations are excluded.

We also need to note that running a RCM at resolution outside of its comfortable resolution range can sometimes bring unexpected results, different from what was previously thought. Vergara-Temprado et al. (2020) found that an explicit representation of convection in a RCM (CCLM) may be beneficial in representing some aspects of climate over Europe at 12-25km resolution that is far away from a few km resolution typical for convection permitting RCMs. Running a hydrostatic RCM (RCA3) in the grey zone (6.5 km), Güttler et al. (2015) showed that many aspects of precipitation climatology over Europe are improved at 6.5km resolution compared to coarser resolutions (50, 25 and 12.5km).


Güttler et al. (2015) [https://journals.ametsoc.org/doi/full/10.1175/MWR-D-14-00302.1](https://journals.ametsoc.org/doi/full/10.1175/MWR-D-14-00302.1)

We added a short explanation to “2.2 Experiment design”:

“We note that in general, both regional models - RCA and HCLIM-ALADIN were developed to operate at a range of tens of km resolution and their performance at 100 and especially at
200km may not be optimal. A potential caveat here is that very few RCM physical
parameterisations are automatically scaled at very coarse resolution. Thus, results at the
coarser resolutions may be partly related to the lack of model retuning. We think that such
coarse-resolution simulations are a useful supplement to simulations at a RCM comfortable
resolution zone and help us to understand RCM behaviour without additional,
resolution-dependent tuning. "

3. Due to the large African domain and that no spectral nudging was used, I am wondering if the
internal variability as mentioned in the line 175 would be large enough to produce large
differences between simulations from the same model? Thus, I would suggest the authors to
rerun the 50-km resolution simulation of one of the two RCMs with different initial conditions or
different starting time and repeat the analysis to see if the IV could affect the simulated
precipitation.

Response: To respond to this and to other experiment-related comments from both reviewers
we extended section “2.2 Experiment design” by providing more details on our experiment
setup, potential caveats and additional sensitivity experiments. We also performed two
additional simulations (RCA4-v1 and ALADIN) at 50km resolution but starting them on 1st
January 1980 instead of 1st January 1979 as for all other simulations in the study (see updated
“2.2 Experiment design”).

4. Conclusion: The discussion of the results in the conclusion is a bit thin and the opening
towards additional studies that could follow that one is missing. I would suggest the authors to
add some discussions about the results and provide few ideas towards additional studies that
could follow that one.

Response: We agree with this comment. In context of more discussions, we should note that
there are almost no studies focusing on multi-resolution RCM experiments over Africa, including
an analog of the no added value experiment (NAVE). We've already proposed that the NAVE
can be used as an additional experiment within the CORDEX framework. In the revised
manuscript we also added that the next step is to focus on i) other variables and especially on
processes and ii) on applications of the NAVE for RCM-based future climate projections (many
thanks to John Scinocca for providing an detailed description of the NAVE in the climate
projection context).

"In our study, as the first step, we focus only on precipitation that has large relevance for climate
change impact studies. As the next step, we foresee similar studies looking also at other variables
and especially at processes and drivers relevant for regional climate.
Moreover, the same NAVE framework can be used for quantifying the added value in
RCM-based future climate projections. For this, one needs to downscale GCMs at their native
resolution in addition to the standard CORDEX resolutions (25 or 50km). The RCM projections
at the native GCM resolution serve as the NAVE in the climate change context. A potential
caveat, already mentioned in our study, is that RCMs are generally developed and tuned to operate at resolution of tens of km. “Downscaling” a GCM at its native resolution, for example 150 or 200km, may lead to artefacts related to a lack of RCM retuning for coarser resolution. Nerveless, more and more GCMs, for example in CMIP6, have resolution finer than 100km that allows application of the NAVE.

Minor Comments:

1. Title: I think that abbreviations should not be used in titles in general. Thus, I suggest to replace “RCM” by “regional climate model” in the title.

   Response: We changed the title accordingly.

2. Lines 25 and 183: Please add a “–” after ALADIN or use parentheses to name the two models.

   Response: changed to (SMHI-RCA4 and HCLIM38-ALADIN)

3. Line 27-29 and 42-43: Something is wrong with these sentences. Please correct them.

   Response: We reformulated these sentences:

   l. 27-29 is now “By contrasting different downscaling experiments, it is found that model formulation has the primary control over many aspects of the precipitation climatology in Africa.”

   l. 42-43 is now “Such model formulation related improvements are strongly model dependent and can, in general, not be considered as an added value of downscaling.”

4. Line 32-34 and 35-39: The sentences are not clear and some points are repeated. Please improve all the sentences of those lines and simplify the message conveyed.

   Response: We reformulated these sentences:

   “However, the impact of higher resolution on the time mean climate is mixed. An improvement in one region/season (e.g. reduction of dry biases) often corresponds to a deterioration in another region/season (e.g. amplification of wet biases). At the same time, higher resolution leads to a more realistic distribution of daily precipitation. Consequently, even if the time-mean climate is not always greatly sensitive to resolution, the realism of the simulated precipitation increases as resolution increases.”
5. Lines 66, 72, 88, 94, 97, 122, 173 etc: The word “results” is used too many times, is too vague and sometimes inappropriate. Sometimes it means the outcome of downscaling. In another context, it refers to the outcome of the analysis. I would suggest to use other words to avoid confusion. As instance, the word “simulation” could be used at lines 66, 72, 173. Please pay attention to every time the word “results” is used and consider using another word or changing the sentence to be more specific.

Response: This comment has been taken into account. We made a number of changes and tried to use “results” for describing the outcome of the analysis.


Response: We think that both studies are relevant in the given context and added Giorgi et al. (2016) as well.

7. Lines 113 and 143: Remove “e.g.”

Response: removed

8. Line 146 and the rest of the paper: About the use of RCA4-v1 and RCA4-v4 to distinguish the two RCA model formulations. I think that v4 is not the best way to name the reduced turbulent mixing simulation since 4 brings in mind that a v2 and v3 are existing and that they are not used in this paper. I would suggest to use RCA4 and RCA4-RTB for Reduced Turbulent Mixing to name the two RCA simulations.

Response: At the moment there are three RCA4 configurations (small domain-related retuning) used in CORDEX and available through ESGF, namely: v1, v2 and v3. RCA4-v4 is a new configuration developed to deal with a large dry bias in Central Africa. New simulations generated by the RCA4-v4 for the Africa-CORDEX domain will be also available on ESGF and we would prefer to keep RCA-v1 and RCA-v4 for consistency. We also added necessary explanations.

“RCA4 has three configurations used for CORDEX simulations that are available through ESGF. They are named (so called RCM version) as v1 (Europe, Arctic, Africa, Southeast Asia, Central and North America), v2 (South Asia) and v3 (South America) and differ in some domain-specific re-tuning. In this study we also include a new configuration - v4. The RCA-v4 is based on RCA4-v1 but with a change in one parameter leading to reduced turbulent mixing in stable situations (especially momentum mixing). Such change in the parameter was applied to reduce a prominent dry bias found in the RCA4-v1 CORDEX Africa simulations over Central Africa.
(Tamoffo et al., 2019; Wu et al., 2016). Using two parameter settings of RCA4 allows us to examine how sensitive our results are to such small tuning of the same RCM.

9. Lines 171-178: I am confuse here about the size of the domains at different resolutions. Is the size of the free domain or full domain including the nudging zone the same between the simulations? Moreover, at line 175, it is mentioned that an additional experiment at 0.88° was performed, but this experiment is never mentioned later on in the analysis. Maybe the sentence of the line 176-178 refer to the two 0.88° simulations. Please pay attention to all the sentences of those lines and specify clearly, which simulation are referred.

Response: Table 1 shows the size of the full domain including the 8 grid point relaxation zone in all directions that is actually explained in l. 169-173. We also updated the title for Table 1.

“Table 1. The full domain configuration and time step for the RCA4 and HCLIM-ALADIN simulations. The full domain includes the 8 grid point relaxation zone.”

l. 175: We extended 2.2 Experimental design adding necessary explanations.

10. Line 180-181: For these simulations . . . Please specify which simulations?

Response: changed to “for these NAVE simulations”

11. Table 1. What the small “a” after 222x222 means?

Response: It’s a typo, deleted.

12. Line 203: Please specify the time period covered by TRMM and be more specific on the time period used for the analysis of Figures 5-6. I think that TRMM starts in 1997 or 1998. Moreover, considering the little amount of weather stations in Africa that are used to create CRU, UDEL and GPCC, I think that TRMM figures covering a subset of the full 1981-2010 could be used in Figures 2, 3 and 4 as it is done in Nikulin et al. (2012).

Response: The TRMM 3B42 (v7) precipitation dataset provides satellite-based precipitation estimates adjusted by large-scale monthly precipitation from gauge networks that is, in our case, the GPCC product. This means that monthly mean TRMM and GPCC7 precipitation in general do not differ too much and are basically almost the same if remapped to the same resolution or averaged over a region. The TRMM data set is used in the study only because of availability of daily and 3-hr precipitation for evaluation of the simulated daily and 3-hr precipitation but not to evaluate seasonal means and annual cycle. We added a few lines to 2.3 “Observations and reanalysis” to explain this issue.

“The TRMM product starts in 1998 and for evaluation of precipitation extremes and diurnal cycle we use a shorter period (1998-2010) in contrast to 1981-2010 used for evaluation of
seasonal means and annual cycle. We also need to note that the TRMM 3B42-v7 precipitation product provides satellite-based precipitation estimates adjusted by the GPCC gauge-based precipitation. This means that monthly mean TRMM 3B42 and GPCC precipitation are almost the same if remapped to the same resolution or averaged over a region."

13. Line 229: Replace “most northern” by “northernmost”.
Response: replaced

14. Lines 237-239: Please improve the sentence that is not clear.
Response: reformulated

"RCA4-v4 shows a similar pattern compared to RCA4-v1 but substantially reduces the dry bias over Central Africa at all four resolutions (Fig. 2i-l). For both configurations of RCA4, the smallest dry bias is found at the highest 25km resolution. At the same time, an overestimation of precipitation north of the region with the dry bias becomes more pronounced, especially for RCA4-v4."

15. Figure 2 and 3: Color scale on the left: The values above 15 mm/day could be removed as in Nikulin et al. (2012)
Response: We limited the color scale by 18 mm/day as there are a few grid boxes with values slightly above 17 mm/day.

16. Figure 2 and 3: Color scale at the bottom: I would suggest to use a white color between -0.5 and 0.5. This would prevent the color change at 0 that is misleading. As example, the Sahara desert is sometimes yellow or blue because there is almost no precipitation falling there.
Response: We agree and use white color between -0.5 and 0.5 mm/day.

17. Figure 2 caption: Please emphasize that the values are aggregated at 200 km.
Response: added

18. Line 284-286: Please give more details about the statement here.
Response: We added a sentence after.
“This indicates that HCLIM-ALADIN parameterisations may be better suited to work also at coarser resolution.”

19. Lines 333-335: Please clarify what is meant by “completely opposite behavior”.

Response: These lines were reformulated:

“HCLIM-ALADIN maintains similar behavior to that in Eastern Africa, although the difference in precipitation across the resolutions is small (Fig. 4l). On the other hand, for both configurations of RCA4 in Central Africa, increasing resolution leads to decreasing precipitation during the rainy seasons, especially in January.”

20. Lines 458-460: It is not clear to me that the 50 km HCLIM simulation shows higher frequency than the 25 km HCLIM simulation.

Response: We checked once more and saw the same result: the 50km PDF (yellow) is above the 25km one (blue) and even for a wider range of intensities (50 to about 200 mm/day) than we noted first. We changed these lines accordingly:

“An interesting detail is that the 50km HCLIM-ALADIN simulation shows higher frequency for intensities in the range of 50 to about 200 mm/day than the 25km simulation.”

21. Figure 6: Please emphasize in the caption that the season is different for the different regions.

Response: added

22. Lines 540-541: There are mistakes about the Figure numbers.

Response: fixed

23. Reference: Please remove the capital letters of the title of Sylla et al. (2013).

Response: removed

Comments from John Scinocca

Major Comments:
In this study the authors introduce a procedure to separate the impact of model formulation from the impact of resolution on the dynamical downscaling results of regional climate models (RCMs) driven by observations (reanalyses). The procedure involves performing the downscaling at several horizontal resolutions. The coarsest RCM resolution is set to match the resolution of the reanalysis model that provides the driving data. This is referred to as the "no added-value experiment", which I will refer to as the NAVE. The authors make the point that the NAVE biases vs the reanalysis biases (relative to an independent observational dataset) result from "model formulation" differences and so are independent of added value. Once NAVE biases are defined, higher resolution RCM simulations are employed to document the evolution of NAVE biases with resolution. It is argued that a reduction of NAVE biases with increasing resolution indicates added value in the RCM. The authors employ this procedure to precipitation biases in RCM downscaling results over the African CORDEX domain from two regional models. The results of the authors' analysis of model formulation vs resolution is often mixed with few clear results. But this is overshadowed by the introduction of the NAVE procedure itself, which is highly publishable as it provides a tool to the RCM community to make progress on the complex issue of added value in RCM studies. In fact, the NAVE approach would seem to have a logical extension to the much more important issue of value added by RCMs in climate-change experiments. In my detailed comments, I suggest a generalization of the NAVE approach to the issue of value added by RCMs in climate-change experiments. It is my recommendation that this manuscript be accepted for publication with only minor revision.

General Minor Comments :

1) NAVE procedure applied to Climate-Change experiments:
The NAVE procedure would seem to be equally applicable to climate change problems to help distinguish the impact of model formulation from the impact of resolution on RCM climate-change responses relative to those of its driving GCM. In the climate change context, two sets of RCM runs would need to be performed - NAVE runs at the resolution of the driving global climate model (GCM) and the usual high-resolution runs used for downscaling GCM climate-change results. Consider a typical time-slice experiment over a CORDEX domain performed at the end of the 20th and 21st centuries. For a given climate index (eg screen-level temperature, precipitation, extremes etc.), one could construct the three climate-change responses:

\[
\begin{align*}
R_{\text{GCM}}(X) &= \text{GCM}_{21\text{st}}(X) - \text{GCM}_{20\text{th}}(X) \\
R_{\text{NAVE}}(X) &= \text{NAVE}_{21\text{st}}(X) - \text{NAVE}_{20\text{th}}(X) \\
R_{\text{RCM}}(X) &= \text{RCM}_{21\text{st}}(X) - \text{RCM}_{20\text{th}}(X),
\end{align*}
\]

where each term on the right is a time (and/or ensemble) average at a given spatial location "X". In the above, \(R_{\text{NAVE}}(X)\) represents the climate-change signal associated with model formulation differences between the RCM and GCM. As for the authors' present-day analysis, the potential for value added due to the response associated with resolution changes may be expressed as:

\[
R_{\text{RES}}(X) = R_{\text{RCM}}(X) - R_{\text{NAVE}}(X).
\]
The NAVE analysis allows the decomposition:

\[ R_{RCM} = R_{NAVE}(X) + R_{RES}(X) \]

Given \( R_{RES}(X) \), and \( R_{NAVE} \) one can ask interesting questions like:
- Where is \( R_{RES}(X) \) significant in the RCM domain?
- Do these locations correlate well with where the authors found downscaling improvement in their NAVE analysis of reanalysis driven RCMs?
- Where is \( R_{RCM} \) appreciably different from \( R_{NAVE} \)?

The appreciable difference analysis presented in Section 5 of Scinocca et al. 2015 (JClim p. 17-35) would seem like an ideal approach to address this question. In locations where there exists an appreciable difference, there exists the potential for added value. However, where there is no appreciable difference, there can be no added value - irrespective of how one chooses to define added value. This is in line with the authors’ stated goals (ll.116-118). Clearly such climate-change questions are outside the authors’ present study but, they may want to discuss this potential application of the NAVE approach for future investigation.

**Response:** We really appreciate such detailed description on how the NAVE approach can be used for climate change projections and completely agree. Now, we also use the abbreviation “NAVE” in the study. Actually, we’ve already completed downscaling of two global models (RCP8.5) over Africa at their native resolution, in addition to the standard 0.44deg CORDEX resolution. The first results were presented at EGU2019 (https://meetingorganizer.copernicus.org/EGU2019/EGU2019-7631.pdf) and at ICRC-CORDEX2019 (http://icrc-cordex2019.cordex.org/wp-content/uploads/sites/2/2019/11/AbstractBook_20191114.pdf, A1-P-38). A paper is in preparation. We added a paragraph in “Summary and Conclusion”

Moreover, the same NAVE framework can be used for quantifying the added value in RCM-based future climate projections. For this, one needs to downscale GCMs at their native resolution in addition to the standard CORDEX resolutions (25 or 50km). The RCM projections at the native GCM resolution serve as the NAVE in the climate change context. A potential caveat, already mentioned in our study, is that RCMs are generally developed and tuned to operate at resolution of tens of km. “Downscaling” a GCM at its native resolution, for example 150 or 200km, may lead to artefacts related to a lack of RCM retuning for coarser resolution. Nerveless, more and more GCMs, for example in CMIP6, have resolution finer than 100km that allows application of the NAVE.”

2) Interpretation of the NAVE:
It is assumed here that differences in the NAVE and driving model results arise from differences the RCM and GCM model formulation. This would be strictly true only if the RCM were also run in a global mode. The one-way nesting approach introduces a number of potential artifacts which are most acute for large RCM domains and applications that do not use interior (or
spectral) nudging - both of which are the case for the authors' present study (eg Section 2 of Scinocca et al. 2015 J Clim p. 17-35). The authors should acknowledge this issue when introducing the NAVE.

Response: It's a very relevant comment as we missed this point. We agree that the one-way nesting approach is also a source of the difference between a RCM and its driving GCM. From our point of view, without spectral nudging, this source is mostly related to RCM domain configuration. If spectral nudging is not used, as in our NAVE simulations, and a RCM develops its own climatology, the difference between the RCM and GCM is basically defined by RCM physical formulation and domain configuration. If spectral nudging is used, technical aspects of the nudging (e.g. which wavelengths should be nudged and at what altitudes) also contribute to the difference by reducing it. We added explanations in 2.2 Experiment design.

“The difference between a RCM and its driving GCM can, in general, be attributed to three sources, namely: i) different resolution, ii) different physical formulation and iii) artifacts of the one-way nesting approach including size of the RCM domain and application of spectral nudging (e.g. Scinocca et al., 2016). The RCA4 0.88° simulations and the HCLIM-ALADIN 100km one represent a slight upscaling of ERAINT (about 0.7° or about 77km at the Equator) and we refer to them as “no added value experiment” (NAVE). No resolution-dependent added value of the RCMs is expected for these simulations and all differences between the RCMs and their driving ERAINT are attributed to different physical formulations and to the artifacts of the one way nesting. Spectral nudging is not used in our experiment and the one way nesting term is basically reduced to domain configuration. In contrast, if spectral nudging is used, technical aspects of the nudging (e.g. which wavelengths should be nudged and at what altitudes) also contribute to the one way nesting term. In practice, it is not straightforward (if possible at all) to separate the impact of different physical formulation and artifacts of the one-way nesting approach. Hereafter, we use “RCM formulation” as a term that includes both RCM physical formulation and domain-dependent RCM configuration (e.g. size of the full domain).”

3) RCM model tuning:

ll.183-185 "We note that in general, both regional models - RCA and HCLIM-ALADIN were developed to operate at a range of 10-50km resolution and their performance at 100 and 200km may not be optimal." This is a non-trivial point, given the philosophy of the authors' NAVE approach. Where there is systematic improvement of NAVE biases with increased resolution, the authors interpret this as a systematic increase in added value. However, The poorer results at the coarser resolutions may also be related to a lack of model retuning at these non-standard resolutions. Very few physical parameterizations are automatically scale dependent and an adjustment of their free parameters with changing spatial resolution should in principle be performed. Retuning the RCMs at each spatial resolution would represent a significant undertaking and these added degrees of freedom would complicate the main point made in this study. Consequently, I would recommend that this issue be addressed by simply having it raised as a caveat.
Response: We completely agree with this comment and added some explanation on this potential caveat (see also response to Comment 3, 1st reviewer)

“We note that in general, both regional models - RCA and HCLIM-ALADIN were developed to operate at a range of tens of km resolution and their performance at 100 and especially at 200km may not be optimal. A potential caveat here is that very few RCM physical parameterisations are automatically scaled at very coarse resolution. Thus, results at the coarser resolutions may be partly related to the lack of model retuning. However, such coarse-resolution simulations are a useful supplement to simulations at a RCM comfortable resolution zone and help us to understand RCM behaviour without additional, resolution-dependent tuning.”

4) Interior nudging:
In downscaling reanalysis products, the authors chose not to employ any constraints on the interior RCM solution such as spectral nudging (ll.185-186). In focusing on such evaluation experiments, one could argue that it is more appropriate to use spectral nudging to constrain the large scales to obtain the best downscaled results in their study. Any upscale influence produced by the RCM would serve only to degrade the large scale flow as it is well observed and represented in reanalysis produces. By not constraining the RCM in this way, the authors leave open the possibility that locations of large biases in their high-resolution RCM results are due to the downscaling of the wrong large-scale flow rather than a lack of intrinsic added value. For more detail see Section 2 of Scinocca et al. 2015 (JClim p. 7-35).

Response: This is a reasonable comment. The CORDEX-Africa RCMs do not use the spectral nudging (see e.g. Nikulin et al. 2012) and here we follow the same approach for downscaling over Africa. A potential caveat for applying spectral nudging in the tropics was also shortly touched in Nikulin et al. 2012.

“With respect to spectral nudging of an RCM solution toward the driving data at large wavelengths (von Storch et al. 2000), this technique is well established for midlatitude regions, with some theoretical understanding of which wavelengths should be nudged and at what altitudes (Alexandru et al. 2009). This is not the case in the tropics, and it may be more difficult to formulate given the stronger role of surface forcing and multiscale convection in driving large-scale circulations. We therefore chose to preclude spectral nudging from the experimental design, pending further work in this area.”

We reformulated a bit:

“All simulations are conducted without spectral nudging similar to the CORDEX-Africa RCMs (Nikulin et al., 2012) allowing the RCMs to develop its own climatology as much as possible.”
Detailed Minor Comments:

I.26 "Additionally to the two RCMs" perhaps change to "In addition to the two RCMs"
Response: changed

I.31 "the phase of the diurnal cycle is" perhaps change to "the phase of the diurnal cycle in precipitation is"
Response: changed

I.71 "However, added value from RCMs" should be changed to "However, perceived added value from RCMs" for the context of the sentence.
Response: changed

II.141-147. It was unclear whether the difference between v1 and v4 was simply a change in a free parameter for an existing scheme or whether the difference was associated with a change in the equations of the scheme. The former might be considered "tuning" while the latter considered a "formulation" difference.
Response: The difference between v1 and v4 is indeed simply a change in a parameter and we completely agree that such change can not be considered as a formulation difference but a new parameter setting or a new configuration. We added additional explanations on the difference between v1 and v4 in “2.1.1 RCA4” and also made a number of changes in the manuscript (using “two different parameter settings” for example).

“RCA4 has three configurations used for CORDEX simulations that are available through ESGF. They are named (so called RCM version) as v1 (Europe, Arctic, Africa, Southeast Asia, Central and North America, ref), v2 (South Asia, ref) and v3 (South America, ref) and differ in some domain-specific re-tuning. In this study we also include a new configuration - v4. The RCA-v4 is based on RCA4-v1 but with a change in one parameter leading to reduced turbulent mixing in stable situations (especially momentum mixing).”

II.176-178 It would be helpful to show these plots to see if the differences have any correlation with later results (perhaps in an appendix) - particularly the distribution of temperature differences.
Response: We've again looked at this additional sensitivity experiment and found that actually there are also some differences in precipitation, not only in temperature. We reformulated our findings accordingly. At the same time, it's only one simulation at one resolution by one RCM.
We would prefer not focus on this single simulation in detail. A full set of simulations with the same full domain for all RCMs and resolutions is necessary for robust conclusions and we leave more in-depth detailed analysis to forthcoming studies.

We reformulated the respective paragraph:

“As mentioned above, larger size of the computational domain at coarser resolution in our experiment setup may have a potential impact on the results leading to larger IV developed by the RCMs and weaker constraints on the ERAINT forcing. As a simple test for domain-dependant RCM IV we perform an additional experiment with RCA4 at 0.88° resolution taking the full computational domain from the 1.76° RCA4 simulation. Indeed, for the 1981-2010 climatology, seasonal mean precipitation differences between the two experiments can reach up to 1.25 mm/day (up to 25%) at a few individual grid boxes, often at the edges of the tropical rain belt, although in general stay below 0.5 mm/day (not shown). Seasonal mean temperature also differs with up to 1.25°C regionally (not shown). We do not focus on this single additional sensitivity experiment in the study. A full set of simulations with the same full domain for all RCMs and resolutions is necessary for robust conclusions.”

ll.260-262 Fig 2b-p. It was often hard to associate the location of a particular bias with the full field in panel a. Expressing the bias as a percentage difference from the full field would be helpful in the West and central regions. However, where there is weak precipitation in the reference/obs data this may be problematic.

Response: It is a common problem for showing relative precipitation biases in the tropics when small reference values at the edge of the tropical rain belt or in dry regions lead to artificially excessive relative bias. One solution is to apply a filter, for example, to show only regions where reference precipitation is more than 1 mm/day. However, based on our experience such an approach does not always lead to better visualisation. Showing absolute biases is pretty common for model evaluation studies in Africa and we prefer to keep the absolute bias in Figs. 2 and 3. Additionally, following a comment of the 1st reviewer we mask by white all biases less than 0.5 mm/day and hope the visibility is better now.

ll.350-352 Fig 4. It would be better to use the colour red for the reference GPCC7 curves in this figure. I had difficulty seeing the GPCC7 curves in a number of the model result panels in columns 2-4.

Response: We changed the colour to red for the reference GPCC7 curves in Fig. 4. We also deleted the CRU and UDEL datasets from Fig. 4 as they simply coincide with GPCC7 and do not provide any useful information.
“In general, model formulation related improvements cannot be considered as an added value of downscaling as such improvements are strongly model dependent and cannot be generalised.” Also, such formulations could in principle be used in global models and so obviate the need for the RCM.

Response: We agree and added:

“However, such formulation-related and region-specific improvements from RCMs could in principle be also used in GCMs.”
The impact of RCM regional climate model formulation and resolution on simulated precipitation in Africa

Minchao Wu1,2, Grigory Nikulin1, Erik Kjellström1, Kjellström1,3, Danijel Belušić1, Colin Jones2, Jones4 and David Lindstedt1

Correspondence to: Minchao Wu (minchaowu.acd@gmail.com)

1Swedish Meteorological and Hydrological Institute, Folkborgsvägen 17, 60176 Norrköping, Sweden

2National present affiliation, Department of Earth Sciences, Uppsala University, Uppsala, Sweden

3Department of Meteorology and the Bolin Centre for Climate Research, Stockholm University, 10691, Stockholm, Sweden

4National Centre for Atmospheric Science (NCAS), University of Leeds, Leeds, UK
Abstract

We investigate the impact of model formulation and horizontal resolution on the ability of Regional Climate Models (RCMs) to simulate precipitation in Africa. Two RCMs—(SMHI-RCA4 and HCLIM38-ALADIN)—are utilized for downscaling the ERA-Interim reanalysis over Africa at four different resolutions: 25, 50, 100 and 200 km. Additionally, in addition to the two RCMs, two different configurations parameter settings (configurations) of the same RCA4 are used. By contrasting different RCMs, configurations and resolutions, it is found that model formulation has the primary control over many aspects of the precipitation climatology in Africa. Patterns of spatial biases in seasonal mean precipitation are mostly defined by model formulation while the magnitude of the biases is controlled by resolution. In a similar way, the phase of the diurnal cycle in precipitation is completely controlled by model formulation (convection scheme) while its amplitude is a function of resolution. Although higher resolution in many cases leads to smaller biases, the impact of higher resolution is mixed. An improvement in one region/season (e.g. reduction of dry biases) often corresponds to a deterioration in another region/season (e.g. amplification of wet biases). The experiments confirm a pronounced and well-known impact of higher resolution—leads to a more realistic distribution of daily precipitation. At the same time, higher resolution makes the time-mean climate is not always greatly sensitive to resolution, what the time-mean climate is made up of, higher order statistics, is sensitive. Therefore, the realism of the simulated precipitation increases as resolution increases.
Our results show that improvements in the ability of RCMs to simulate precipitation in Africa compared to their driving reanalysis in many cases are simply related to model formulation and not necessarily to higher resolution. Such model formulation related improvements are strongly model dependent and in general cannot be considered as an added value of downscaling.

Keywords: RCA4, HCLIM, Resolution dependency, Added value, CORDEX-Africa
1 Introduction

Regional climate modeling is a dynamical downscaling method widely used for downscaling coarse-scale global climate models (GCMs) to provide richer regional spatial information for climate assessments and for impact and adaptation studies (Giorgi and Gao, 2018; Giorgi and Mearns, 1991; Laprise, 2008; Rummukainen, 2010). It is well-established that regional climate models (RCMs) are able to provide added value (understood as improved results in climatology) compared to their driving GCMs. This includes better representation of regional and local weather and climate features as a result of better capturing small-scale processes, including those influenced by topography, coast lines and meso-scale atmospheric phenomena (Flato et al., 2013; Prein et al., 2016). However, perceived added value from RCMs may have different causes and it may not always be for the right reason where “right reason” would result from an improved representation of regional process at smaller scales. Such improvement leads to more accurate results in simulations on local scales, and can, to some extent, also reduce large-scale GCM biases (Caron et al., 2011; Diaconescu and Laprise, 2013; Sørland et al., 2018). Contrastingly, added value may be attributed to the “wrong reason”, not directly related to higher resolution in RCMs but to different model formulation in the RCMs and their driving GCMs. It is possible that the physics of a RCM has been targeted for processes specific to the region it is being run for, giving it a local advantage over GCMs that may have had their physics developed for global application. However, it is questionable if improvements of such “downscaling” via physics can be considered as an added value. In general, RCMs can either reduce or amplify GCM biases sometimes even changing their signs (Chan et al., 2013).
Issues as those mentioned above, have raised substantial concerns among regional climate modelers (e.g., Castro, 2005; Xue et al., 2014). It has been pointed out that understanding of the added value remains challenging. It would become even more complicated taking into account the effects of different realizations, such as the size of domain, lateral boundary conditions, geographical location, model resolution and its internal variability (Di Luca et al., 2015; Hong and Kanamitsu, 2014; Rummukainen, 2016). All the above factors potentially influence downscaling results RCM simulations leading to different interpretations of the downscaling effects, thus the robustness of added value. For example, it was shown that over the Alps, downscaling with multiple RCMs at increasing resolutions in general is able to provide a more realistic precipitation pattern than the forcing GCMs, and it is regarded as added values from RCMs (Giorgi et al., 2016; Torma et al., 2015b, 2015). Similarly, Lucas-Picher et al (2017) found added value over the Rocky Mountains, another region with strong topographic influence on hydrological processes. However, the results are not unambiguous and sometimes limited added value is found when comparing to the forcing data, (e.g. Wang and Kotamarthi, 2014) over North America. This implies that the understanding of downscaling effects is context-dependent and one should carefully interpret the downscaled results GCM and RCM simulations in order to detect robust added value.

Africa is foreseen to be vulnerable to future climate change, which early on inspired efforts to employ RCMs for impact and adaptation studies (e.g. Challinor et al., 2007). Further to previous coordinated downscaling activities over Africa as for example the African Monsoon Multidisciplinary Analyses (AMMA) (Van der Linden and Mitchell, 2009), the Coordinated Regional climate Downscaling Experiment (CORDEX) provides a large ensemble of RCM
projections for Africa (Giorgi et al., 2009; Jones et al., 2011). All CORDEX RCMs follow a common experiment protocol including a predefined domain at 50km resolution and common output variables and format that facilitates assessment of projected climate changes in Africa. Under this framework, RCMs at 50-km horizontal resolution are found to have the capability of providing added value in representing African climatological features compared to their forcing GCMs, which generally have the resolution coarser than 100 km (Dosio et al., 2015; Moufouma-Okia and Jones, 2015; Nikulin et al., 2012). However, a number of common problems with the RCMs are identified, which include, for example, dry biases over convection-dominated regions like the Congo basin, too early onset of the rainy season for the West African Monsoon region and biases in representing the diurnal cycle of precipitation (Kim et al., 2014; Laprise et al., 2013; e.g. Nikulin et al., 2012). So far, it is still not clear if challenges to accurately simulate precipitation climatology in Africa have also been identified for the RCMs. Individual RCMs may exhibit substantial biases in different aspects of the precipitation climatology as seasonal mean (Endris et al., 2013; Kalognomou et al., 2013; Kim et al., 2014; Shongwe et al., 2015; Tamoffo et al., 2019), annual cycle (Favre et al., 2016; Kisembe et al., 2019), onset and cessation of the rainy season (Akinsanola and Ogunjobi, 2017; Gbobaniyi et al., 2014), number of wet days and their intensity (Klutse et al., 2016). At the same time, most of these studies found that such biases often strongly depend on region and season. A RCM with a substantial bias in one region and/or season may accurately simulate precipitation in other regions and seasons. It was also found that the multi-model ensemble usually outperforms individual RCMs but it is a result of the cancelation of opposite-signed biases in different RCMs.
A number of possible explanations for such RCM precipitation-related biases in Africa were suggested as for example: different convection schemes (see discussion in Kalognomou et al., 2013), land-atmosphere coupling (e.g. Sylla et al., 2013b) and biases in moisture transport (Tamoffo et al., 2019). However, most of the CORDEX-Africa studies are still descriptive and process-based evaluation studies like Tamoffo et al. (2019) are mostly lacking. An additional barrier for more process-based evaluation studies is that the CORDEX variable list only defines three pressure levels (850, 500 and 200mb) to be provided that seriously limits evaluation of large-scale and regional circulation (e.g. jet streams) and moisture transport in the troposphere.

Another common problem for almost all RCMs in Africa is the phase of the diurnal cycle of precipitation. The majority of RCMs simulate maximum precipitation intensity around local noon that is too early compared to late afternoon or even late evening maximum evident in observations (Nikulin et al., 2012). This deficiency of the RCMs is related to the convective parameterization used and a specific convection scheme, as for example the Kain–Fritsch (KF), may outperform others, producing a more realistic diurnal cycle (Nikulin et al., 2012).

All the above deficiencies of the RCMs show that higher resolution does not necessarily lead to a better performance of the RCMs in terms of precipitation climatology in Africa. It is also not always clear if differences between the CORDEX Africa RCMs and their driving GCMs are related to higher RCM resolution or to RCM internal formulation, or to the combination of both. A thorough understanding of such differences and of added value of the CORDEX-Africa RCMs is necessary for robust regional assessments of future climate change and its impacts in Africa.
In this study, we aim to separate the impact of model formulation and resolution on the ability of RCMs to simulate precipitation in Africa. We conduct a series of sensitivity, reanalysis-driven experiments by applying two different RCMs, one of them in two different configurations, at four horizontal resolutions. Contrasting the different experiments allow us to separate the impact of model formulation and resolution. We present an overview and the first results of the experiments conducted and leave in-depth detailed process studies for different regions to forthcoming papers.

2 Methods and Data

2.1 The Regional Climate Models

2.1.1 RCA4
The Rossby Centre Atmosphere regional climate model - RCA (Jones et al., 2004; Kjellström et al., 2005; Räisänen et al., 2004; Rummukainen et al., 2001; Samuelsson et al., 2011) is based on the numerical weather prediction model HIRLAM (Undén et al. 2002). To improve model transferability, the latest fourth generation of RCA, RCA4, has a number of modifications for specific physical parameterizations. This includes the modification of convective scheme based on Bechtold-Kain-Fritsch scheme (Bechtold et al., 2001) with revised calculation of convective available potential energy (CAPE) profile according to Jiao and Jones (2008), and the introduction of turbulent kinetic energy (TKE) scheme (Lenderink and Holtslag, 2004). The RCA4 model has been applied in many regions worldwide, among them Europe (Kjellström et al., 2016, 2018; Kotlarski et al., 2015), the Arctic (Berg et al., 2013; Koenigk et al., 2015; Zhang
et al., 2014), Africa (Nikulin et al., 2018; Wu et al., 2016), South America (Collazo et al., 2018; Wu et al., 2017), South-East (Tangang et al., 2018) and South Asia (Iqbal et al., 2017).

In addition to the standard RCA4 configuration, used in CORDEX, in

RCA4 has three configurations used for CORDEX simulations that are available through ESGF. They are named (so called RCM version) as v1 (Europe, Arctic, Africa, Southeast Asia, Central and North America), v2 (South Asia) and v3 (South America) and differ in some domain-specific re-tuning. In this study we also include a new configuration with v4. The RCA-v4 is based on RCA4-v1 but with a change in one parameter leading to reduced turbulent mixing in stable situations (especially momentum mixing). Such change in the model formulation was applied to reduce a prominent dry bias found in the RCA4 CORDEX Africa simulations over Central Africa (Tamoffo et al., 2019; Wu et al., 2016). Using two configurations allows us to examine how sensitive our results are to different formulations such small tuning of the same model. We hereafter denote the original RCA4 configuration as RCA4-v1 and the new one as RCA4-v4 RCM.

2.1.2 HCLIM

HARMONIE-Climate (HCLIM) is a regional climate modelling system designed for a range of horizontal resolutions from tens of kilometers to convection permitting scales of 1-3 km (Belušić et al., 2019; Lindstedt et al., 2015). It is based on the ALADIN-HIRLAM numerical weather prediction system (Belušić et al., 2019; Bengtsson et al., 2017; Termonia et al., 2018). The HCLIM system includes three atmospheric physics packages AROME, ALARO and ALADIN, which are designed for different horizontal resolutions. The ALADIN model
configuration used in this study employs the hydrostatic ARPEGE-ALADIN dynamical core (Temperton et al., 2001), a mass-flux scheme based on moisture convergence closure for parameterizing deep convection (Bougeault, 1985) and SURFEX as the surface scheme (Masson et al., 2013). All details about the version of HCLIM used in this study (HCLIM38), and its applications over different regions can be found in (Belušić et al., 2019). We need to note that HCLIM38-ALADIN used in the study is not the same model as ALADIN-Climate used in CORDEX (Daniel et al., 2019). We refer to HCLIM38-ALADIN as HCLIM-ALADIN hereafter.

2.2 Experimental design

To investigate the response of both RCA4 and HCLIM-ALADIN to horizontal resolution, we conduct a set of sensitivity experiments driven by the ERA-Interim reanalysis (denoted as ERAINT hereafter; Dee et al., 2011) at four different resolutions. These resolutions are 1.76, 0.88, 0.44 and 0.22° for RCA4 with the rotated coordinate system and 200, 100, 50 and 25km for HCLIM-ALADIN with the Lambert Conformal projection. The 0.44° or 50km resolution is recommended by the CORDEX experiment design and used in the CORDEX-Africa ensemble.

Hereafter, the resolution in kilometers is used unless otherwise specified.

There are two approaches to setup a RCM experiment with simulations at different resolutions. The first approach is to use the same full domain (including the relaxation zone) for all simulations at different resolutions. Size of the full domain is defined by the coarsest resolution in the experiment (200km in our case). A benefit of such experiment setup is a consistent lateral boundary forcing for all simulations, given the same full domain. However, an unnecessary large full domain for higher resolution simulations is a caveat leading to larger RCM internal
variability (IV) and a higher computational demand at finer resolutions. The second approach is to use different (minimum) full domains for different resolutions defined only by size of the active domain (the same for all resolution) and a necessary relaxation zone (smaller in km for higher resolution). An advantage of this approach is less IV and less computational demand for high resolution simulations while a shortcoming is inconsistent lateral boundary forcing (different size of the full domain). We decided to use the second approach with the minimum size of the full domain (less IV and computational demand), although we should note that a perfect experiment has to include both approaches, if resources allow. The setup of the simulations at the four resolutions is identical apart from the timestep that is adjusted to ensure numerical simulation stability and the size of the full computational domain with the relaxation zone (see Table 1). The relaxation zone has 8 grid-points in all directions and increases (in km) at coarser resolution while the interior CORDEX-Africa domain is the same. Larger

As mentioned above, larger size of the computational domain at coarser resolution in our experiment setup may have a potential impact on the simulated climatology leading to larger internal variability IV developed by the RCMs and weaker constraints on the ERAINT forcing. WeAs a simple test for domain-dependant RCM IV we perform an additional experiment with RCA4 at 0.88° resolution taking the full computational domain from the 1.76° RCA simulation. For precipitation differences between the two experiments are at the noise level while for seasonal mean temperature it can be up to 1°C. Indeed, for the 1981-2010 climatology, seasonal mean precipitation differences between the two experiments can reach up to 1.25 mm/day (up to 25%) at a few individual grid boxes, often at the edges of the tropical rain belt, although in general stay below 0.5 mm/day (not shown). Seasonal mean temperature also
differs with up to 1.25°C regionally (not shown). We do not focus on this single additional
sensitivity experiment in the study. A full set of simulations with the same full domain for all
RCMs and resolutions is necessary for robust conclusions.

Another source of IV in RCMs is related to different initialisation or starting time (e.g.
Lucas-Picher et al., 2008; Sanchez-Gomez and Somot, 2018). We perform two additional
experiments in order to see how different initialisation time impacts the IV in the RCMs. Both
RCA4-v1 and ALADIN at 50km were initialised on 1st January 1980 instead of 1st January
1979 as for all other simulations in the study. It was found that the impact of the different starting
time is much smaller than the impact of the larger domain. For both seasonal mean precipitation
and temperature, differences between the experiments are small over the African continent, in
general, less than 0.5 mm/day for precipitation and 0.25°C for temperature (not shown). Similar
to the domain-dependent sensitivity experiment above, we do not focus on these two additional
initialisation sensitivity experiments in the study. A full investigation of the initialisation-related
RCM IV needs generation of a larger (up to 10 members) ensemble for all RCMs and
resolutions.

We note that in general, both regional models - RCA and HCLIM-ALADIN were developed to
operate at a range of tens of km resolution and their performance at 100 and especially at 200km
may not be optimal. A potential caveat here is that very few RCM physical parameterisations are
automatically scaled at very coarse resolution. Thus, RCM deficiencies at the coarser resolutions
may be partly related to the lack of model retuning. We think that such coarse-resolution
simulations are a useful supplement to simulations at a RCM comfortable resolution zone and
help us to understand RCM behaviour without additional, resolution-dependent tuning. All simulations are conducted without spectral nudging similar to the CORDEX-Africa RCMs ((Nikulin et al., 2012)) allowing the RCMs to develop its own climatology as much as possible. Analysis is done for the CORDEX-Africa domain shown in Fig. 1.

The difference between a RCM and its driving GCM can, in general, be attributed to three sources, namely: i) different resolution, ii) different physical formulation and iii) artifacts of the one-way nesting approach including size of the RCM domain and application of spectral nudging (e.g. Scinocca et al., 2016). The RCA4 0.88° simulations and the HCLIM-ALADIN 100km one represent a slight upscaling of ERAINT (about 0.7° or about 77km at the Equator) and we refer to them as “no added value experiment” (NAVE). No resolution-dependent added value of the RCMs is expected for these NAVE simulations and all differences between the RCMs and their driving ERAINT are attributed to different model physical formulations. We note that in general, both regional models—RCA and HCLIM-ALADIN were developed to operate at a range of 10–50km resolution and their performance at 100 and 200km may not be optimal. All simulations are conducted without spectral nudging and analysis is done for the CORDEX-Africa domain shown in Fig. 1.

Table 1. Details of the RCA4 and HCLIM-ALADIN experiments and to the artifacts of the one way nesting. Spectral nudging is not used in our experiment and the one way nesting term is basically reduced to domain configuration. In contrast, if spectral nudging is used, technical aspects of the nudging (e.g. which wavelengths should be nudged and at what altitudes) also contribute to the
one way nesting term. In practice, it is not straightforward (if possible at all) to separate the impact of different physical formulation and artifacts of the one-way nesting approach. Hereafter, we use “RCM formulation” as a term that includes both RCM physical formulation and domain-dependent RCM configuration (e.g. size of the full domain).

Table 1. The full domain configuration and time step for the RCA4 and HCLIM-ALADIN simulations. The full domain includes the 8 grid point relaxation zone.

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>Horizontal resolution (deg. / km)</th>
<th>Domain size (lon × lat)</th>
<th>Geographical area (deg.)</th>
<th>Time step (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCA4-v* 1.76°</td>
<td>1.76°</td>
<td>66 × 67</td>
<td>-60.5, 55.66</td>
<td>1200</td>
</tr>
<tr>
<td>RCA4-v* 0.88°</td>
<td>0.88°</td>
<td>126 × 121</td>
<td>-54.78, 50.82</td>
<td>1200</td>
</tr>
<tr>
<td>RCA4-v* 0.44°</td>
<td>0.44°</td>
<td>222 × 222</td>
<td>-50.16, 47.08</td>
<td>1200</td>
</tr>
<tr>
<td>RCA4-v* 0.22°</td>
<td>0.22°</td>
<td>406 × 422</td>
<td>-48.07, 44.55</td>
<td>600</td>
</tr>
<tr>
<td>HCLIM-ALADIN</td>
<td>200 km</td>
<td>80 × 90</td>
<td>-58.34, 56.71</td>
<td>1800</td>
</tr>
<tr>
<td>200km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCLIM-ALADIN</td>
<td>100 km</td>
<td>128 × 150</td>
<td>-53.89, 51.70</td>
<td>1800</td>
</tr>
<tr>
<td>100km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCLIM-ALADIN</td>
<td>50 km</td>
<td>240 × 270</td>
<td>-51.56, 48.98</td>
<td>1200</td>
</tr>
<tr>
<td>50km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCLIM-ALADIN</td>
<td>25 km</td>
<td>450 × 512</td>
<td>-50.43, 47.73</td>
<td>600</td>
</tr>
<tr>
<td>25km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 Topography (m) for the CORDEX-Africa domain in RCA4 at 50km resolution. Boxes indicate the four subregions used for spatially averaged analysis: West Africa (WA), East Africa (EA), the southern Central Africa (CA-S), and eastern southern Africa (SA-E).

2.3 Observations and reanalysis

Observational datasets in Africa, in general, agree well for large-scale climate features but can deviate substantially at regional and local scales (Fekete et al., 2004; Gruber et al., 2000; Nikulin et al., 2012). To take into account the observational uncertainties, we utilize a number of gridded precipitation datasets. They include three gauged-based datasets: the Global Precipitation Climatology Centre, GPCC, version 7 (Schneider et al., 2014), the Climate Research Unit Time-Series, CRU TS, version 3.23 (Harris et al., 2014), and University of Delaware, UDEL, version 4.01 (Legates and Willmott, 1990). All these three datasets are at 0.5° horizontal resolution. For the evaluation of precipitation extremes and diurnal cycle simulated by RCMs, we utilize a satellite-based precipitation dataset from the Tropical Rainfall Measuring Mission, TRMM 3B42 version 7 (Huffman et al., 2007), which is at 0.25° horizontal resolution and
3-hourly temporal resolution. The TRMM product starts in 1998 and for evaluation of precipitation extremes and diurnal cycle we use a shorter period (1998-2010) in contrast to 1981-2010 used for evaluation of seasonal means and annual cycle. We also need to note that the TRMM 3B42-v7 precipitation product provides satellite-based precipitation estimates adjusted by the GPCC gauge-based precipitation. This means that monthly mean TRMM 3B42 and GPCC precipitation are almost the same if remapped to the same resolution or averaged over a region. ERAINT as the driving reanalysis is also used for analysis. In contrast to climate models, ERAINT precipitation is a short term forecast product and there are several ways to derive ERAINT precipitation (e.g. different spin-up, base time and forecast steps) that can lead to different precipitation estimates (Dee et al. 2011). ERAINT precipitation is derived by the simplest method, without spinup as in some of the previous studies (Dosio et al., 2015; Moufouma-Okia and Jones, 2015; Nikulin et al., 2012): 3-hourly precipitation uses the base times 00/12 and forecast steps 3/6/9/12 hours, while daily precipitation uses base times 00/12 and forecast steps of 12 hours. The RCMs and ERAINT represent 3-hourly mean precipitation for the 00:00-03:00, 03:00-06:00, ... 21:00-00:00 intervals while TRMM precipitation averages represent approximately the 22:30–01:30, 01:30–04:30, ... 19:30–22:30 UTC intervals.

2.4 Methods
The coarsest resolution 200 km is used as a reference resolution for spatial maps. The higher-resolution simulations are aggregated to the 200 km grid by the first-order conservative remapping method (Jones, 1999). In this way we expect that the difference among the aggregated results at common resolution should mainly be caused by the different treatment for fine-scale processes (Di Luca et al., 2012). For the regional analysis, such as the analysis of
annual cycle, diurnal cycle and daily precipitation intensity, we focus on four subregions, presenting different climate zones in Africa: West Africa (10°W~10°E, 7.5°N~15°N), East Africa (30°E~40°E, 15°S~0°S), the southern Central Africa (10°E~25°E, 10°S~0°S), and the eastern South Africa (20°E~36°E, 35°S~22°S) as defined in Fig. 1. The period 1981-2010 is used for the analysis in this study, unless otherwise specified.

3 Results and Discussion

3.1 Seasonal mean

In the boreal summer defined here as July-September (JAS), the tropical rain belt (TRB) associated with the intertropical convergence zone (ITCZ) is positioned to its northernmost location with the maximum precipitation north of the Equator (Fig. 2a). CRU, UDEL and GPCC aggregated to the 200km resolution, generally agree well with each other, with only slight local differences (Fig. 2a-c). ERAINT overestimates precipitation over Central Africa and along the Guinea Coast while underestimates it over West Africa, north of the Guinea Coast (Fig. 2d). All RCA4-v1 simulations have a pronounced dry bias (Fig. 2e-h) that spatially almost coincides with the wet bias in ERAINT and increases at coarser resolution (Fig 1e-f). RCA4-v4 shows a similar bias pattern compared to RCA4-v1 but substantially reduces the dry bias over Central Africa at all four resolutions (Fig. 2i-l). For both configurations of RCA4, the smallest dry bias is found at the highest 25km resolution, although, at the same time, an overestimation of precipitation north of the region with the dry bias becomes more pronounced, especially for RCA4-v4. HCLIM-ALADIN, in general, shows some similarities to RCA4 with a pronounced dry bias in West and Central Africa at 200km that is strongly reduced
with increasing resolution. However, a wet bias emerges on the northern flank of the rain belt at 50 and 25km. For JAS there is a common tendency for both RCMs to generate more precipitation at higher resolution leading to a reduction of the dry biases over Central Africa. Such bias reduction may be considered as an resolution-related improvement. However, the RCM simulations clearly show that the added value of higher resolution can be region dependent. An improvement of the simulated precipitation climatology over one region corresponds to deterioration of the climatology over another region. Moufouma-Okia and Jones (2015) found a mixed response to resolution in simulated seasonal mean precipitation over West Africa. Their RCM simulations at 50 and 12km bear a great deal of similarity with each other while a simulation at 25km shows wetter conditions in the Sahel and drier ones near the coastal area in the south (see their Fig. 8). In contrast, Panitz et al. (2014) found almost no difference in seasonal rainfall over West Africa between two RCM simulations at 50 and 25km. We conclude that for both RCA4 and HCLIM-ALADIN, spatial bias patterns are similar and more related to model formulation while magnitude of biases are more sensitive to resolution. For example, the sign of the bias pattern in our no added value RCM simulations at 100km in JAS (Fig. 2f, j, n) is almost opposite to the sign of the bias pattern in the driving ERAINT (Fig. 2d).
Figure 2. GPCC7 mean JAS precipitation for 1981–2010 and differences compared to GPCC7 in (b-d) the other gridded observations, (e-h) the RCA4-v1, (i-l) RCA4-v4 and (m-p) HCLIM-ALADIN simulations. All data sets are aggregated to the coarsest 200km grid.
In boreal winter (December-February, DJF), the TRB migrates to its most southerly position covering the latitudes from southern to Central Africa, with the maximum over southern tropical Africa and Madagascar (Fig. 3a). Similar to JAS, observational uncertainties are generally small in DJF and there is a pronounced wet bias in ERAINT over Central Africa (Fig. 3d). At 25 and 50km RCA4-v1 has a dipole bias pattern with an underestimation of rainfall over Central Africa and an overestimation over southern Africa. At 200km there is a pronounced deterioration in the simulated rainfall: a strong dry bias appears along the eastern coast and Madagascar while the wet bias is amplified over large parts of southwestern Africa. At 25 and 50km RCA4-v4 shows a large-scale dipole bias pattern similar in some degree to RCA4-v1. The RCA4-v4 biases are smaller than the RCA4-v1 ones showing an impact of the re-tuning (reducing mixing in the boundary layer). The behaviour of RCA4-v4 at coarser resolution is also similar to RCA4-v1. A similar strong dry bias is emerging along the eastern coast at 200km. However, in contrast to RCA4-v1, the dry bias over the Democratic Republic of Congo almost completely disappears at both 100 and 200km. HCLIM-ALADIN simulates almost the same bias pattern at all resolutions, strongly underestimating rainfall over southeastern Africa and overestimating it over the Guinea Coast, parts of central Africa and southern Africa. There is a tendency to an increase in precipitation with higher resolution in HCLIM-ALADIN: the wet biases are amplified and the dry biases are reduced. Both RCA4 and HCLIM-ALADIN show a common feature - intensification of the dry bias along the eastern coast of Africa at 200km. Even if both RCMs have this dry bias in common, there are also differences showing the importance of model formulation. HCLIM-ALADIN has about the same bias pattern at all four resolutions while the RCA4 bias pattern substantially changes across the resolutions. Such resolution dependency in
RCA4 may be related to the fact that RCA4 is based on a limited area model and not developed to operate at 100-200km resolution. Contrastingly, HCLIM-ALADIN that is based on a global model shows more consistent results even at 100-200km resolution. This indicates that HCLIM-ALADIN parameterisations may be better suited to work also at coarser resolution.

Although, we also note that the resolution dependency of the RCA4 bias pattern over southern Africa is similar to that found for the CMIP5 GCMs (Munday and Washington, 2018). They show that the GCMs with the coarsest resolution and respectively the lowest topography have the wettest bias over the Kalahari basin and the driest bias over the southeast Africa coast, the Mozambique Channel and Madagascar. Such a bias pattern is related to a smoother barrier to northeasterly moisture transport from the Indian Ocean that penetrates across the high topography of Tanzania and Malawi into subtropical southern Africa. However, in our analysis, HCLIM-ALADIN does not show such resolution-related dependency. In general, similar to JAS, the added value of higher resolution in DJF is region dependent: with higher resolution biases are reduced over one region but amplified over another.
3.2 Annual cycle

The annual cycle of precipitation over the four subregions is shown in Fig. 4. The observed annual cycle of precipitation over West Africa depicts the West African Monsoon (WAM)
rainfall, with maximum precipitation in August (Fig. 4a). All observational datasets (CRU and UDEL are not shown) and ERAINT agree well with each other with only a small underestimation of rainfall by ERAINT in June-August. In contrast to the observations, RCA4-v1 has a bimodal annual cycle with a too early onset of the rainy season (Fig. 4b). The simulated rainfall is overestimated in March-May, underestimated in July-August during the active WAM period and is well in line with the observations during the cessation of the WAM rainfall in September-November. RCA4-v4 shows a similar behaviour but the first rainfall peak in May is reduced and the annual cycle has a more unimodal shape (Fig. 4c). HCLIM-ALADIN, in general, shows similar features as both configurations of RCA4, although has more similarities with RCA4-v4 (Fig. 4d). The too early onset of the rainy season is a common problem for many RCMs reported by Nikulin et al., (2012). Our results show that this is not dependent on resolution but instead related to model formulation. Higher resolution reduces the wet bias during the onset of the rainy season for RCA-v1, has no impact for RCA-v4 and amplifies the wet bias in HCLIM-ALADIN. Nevertheless, the impact of higher resolution is more consistent during the rainy season. Increasing resolution tends to increase monsoon rainfall for both RCMs, resulting in smaller dry biases and a pattern closer to the unimodal one in the observations. Eastern and Central Africa have a bimodal annual cycle of rainfall with two peaks around November and May (Fig. 4e,i). GPCC, CRU and UDEL (both not shown) agree well on the phase and magnitude of the annual cycle for both subregions. ERAINT has a weaker bimodality overestimating precipitation in December-February over Eastern Africa and all year round over Central Africa with the largest wet bias during October-April. Both configurations of RCA4 fail to reproduce the bimodal annual cycle in Eastern Africa at 200km underestimating
precipitation all year round and showing a single rainfall peak in December (Fig. 4j,k).

Increasing resolution reduces the dry bias and leads to an improvement in the shape of the annual cycle. The bimodal shape begins to appear at 100km and becomes much closer to the observation at 50 and 25km. Despite some mixed dry and wet biases in different seasons, the 25 and 50km RCA4 simulations show the best agreement with the observations. In contrast to RCA4, HCLIM-ALADIN simulates the unimodal annual cycle at all four resolutions and some sign of bimodality only appears at 25km (Fig. 4h). Similar to RCA4, increasing resolution leads to an increase in precipitation in HCLIM-ALADIN, although a dry bias is a prominent feature from November to May in all HCLIM-ALADIN simulations. For Central Africa, the bimodality of the annual cycle is well reproduced by both RCMs at all resolutions (Fig. 4j-l). An interesting feature is that RCA4 shows completely opposite behavior in Central Africa compared to Eastern Africa. Increasing resolution leads to decreasing precipitation for both configurations of RCA4 during the rainy seasons and especially in January. HCLIM-ALADIN maintains similar behavior to that in Eastern Africa, although difference in precipitation across the resolutions is small (Fig. 4l). HCLIM-ALADIN maintains similar behavior to that in Eastern Africa, although the difference in precipitation across the resolutions is small (Fig. 4l). On the other hand, for both configurations of RCA4 in Central Africa, increasing resolution leads to decreasing precipitation during the rainy seasons, especially in January. Both RCMs strongly reduce the ERAINT wet bias even in the no-added-value experiment NAVE at 100km. Such improvement indicates that model formulation plays a more important role than resolution over Central Africa. For the eastern Southern Africa, the annual cycle of precipitation is unimodal with its maximum in austral summer (Fig. 4m). Similar to West Africa, uncertainties between observational datasets
and reanalysis are small. RCA4 in general overestimates rainfall during the rainy season with the largest wet bias at 200km. Surprisingly, the simulated rainfall is almost the same at 25 and 100km while the smallest bias is found at 50km for both RCA4 configurations.

HCLIM-ALADIN also overestimates precipitation during the rainy season at all four resolutions (Fig. 4p). However, the smallest wet bias in the HCLIM-ALADIN simulations is found at 50 and 100km.
Figure 4. Annual cycle of precipitation over the four subregions for 1981-2010 in observations/ERAINT and as simulated by RCA4 and HCLIM-ALADIN at the four different resolutions. Only land grid boxes are used for averaging over the subregions. Units are mm/day.
3.3 Diurnal cycle

The diurnal cycle is a prominent feature of forced atmospheric variability with a strong impact on regional- and local-scale thermal and hydrological regimes. The diurnal cycle of precipitation in the tropics is well documented and includes a late afternoon/evening maximum over land (Dai et al., 2007). However, it is still a common challenge for GCMs (Dai, 2006; e.g. Dai and Trenberth, 2004; Dirmeyer et al., 2012), RCMs (e.g. Da Rocha et al., 2009; Jeong et al., 2011; Nikulin et al., 2012) and reanalyses (Nikulin et al., 2012) to accurately represent the diurnal cycle of precipitation.

The TRMM diurnal cycle of precipitation generally shows an increase of rainfall starting around noon with maximum reached at around 18:00 local solar time (LST) (Fig. 5). The ERAINT diurnal cycle is completely out of phase over all subregions with the occurrence of maximum precipitation intensity around local noon. A common feature of ERAINT is an overestimation of precipitation around local noon and an underestimation during the rest of the day.

HCLIM-ALADIN shows exactly the same behaviour as ERAINT. Both configurations of RCA4 simulate the diurnal cycle of precipitation more accurately compared to ERAINT and HCLIM-ALADIN. The phase of the diurnal cycle, in general, is pretty well captured over all four subregions. In terms of precipitation intensity RCA4 underestimates rainfall from afternoon to morning over West (Fig. 5b,c) and Central Africa (Fig. 5j,k). Reducing mixing in the boundary layer results in flattening of the diurnal cycle over West Africa (Fig. 5b, c) while there are almost no changes over Central Africa (Fig. 5j, k). RCA4-v1 very well simulates the diurnal cycle over Eastern Africa with only some underestimation in early morning and afternoon (Fig. 5f). RCA4-v4 improves rainfall intensity in early morning but at the same time shows a slightly
larger underestimation in afternoon than RCA4-v1 (Fig. 5g). Over Southern Africa the RCA4 simulations at 200km are the closest to the observation (Fig. 5n,o) while the simulations at higher resolutions underestimate the amplitude of the diurnal cycle in the afternoon. Figure 5 clearly shows that the phase of the diurnal cycle of precipitation in Africa does not depend on resolution but instead depends on model formulation. Both ERAINT, with the Tiedtke convection scheme (Tiedtke, 1989), and HCLIM-ALADIN with the Bougeault scheme (Bougeault, 1985) trigger precipitation too early during the diurnal cycle while both configurations of RCA4 with the same Kain–Fritsch (KF) scheme (Bechtold et al., 2001) simulate much more realistic diurnal cycle. It has previously been shown that the KF scheme is able to reproduce late afternoon rainfall peaks for the regions where moist convection is governed by the local forcing, for example in the southeast US (Liang, 2004) and in the tropical South America and Africa (e.g. Bechtold et al., 2004; Da Rocha et al., 2009). Nikulin et al., (2012) also found that a subset of RCMs that employ the KF scheme show an improved representation of the phase of the diurnal cycle in Africa. Our results indicate that the impact of resolution is only seen in the amplitude of the diurnal cycle. However, such impact is not homogeneous across the subregions and the RCMs. For HCLIM-ALADIN, increasing resolution leads to increasing rainfall intensity in all regions but southern Africa. RCA4 shows a similar behaviour over West Africa, while there is a mixed response over Eastern and Central Africa. These findings are in line with previous studies investigating resolution effects for GCMs (Covey et al., 2016; Dirmeyer et al., 2012) and for RCMs (Walther et al., 2013). In coarser-scale models (e.g >10km), increasing resolution only leads to changes in the magnitude, but not in the phase of the diurnal cycle of precipitation over land.
Nevertheless, studies conducting sensitivity experiments using resolutions finer than 10 km do find improvements in the representation of the phase (Dirmeyer et al., 2012; Sato et al., 2009; Walther et al., 2013).

**Figure 5.** Diurnal cycle of 3-hourly mean precipitation over the four subregions for 1998-2010 in observations/ERAINT and as simulated by RCA4 and HCLIM-ALADIN at the four different resolutions. Only land grid boxes are used for averaging over the subregions and only wet days with more than 1mm/day are taken for estimations of the diurnal cycle.
3.4 Frequency and intensity of daily precipitation

Figure 6 shows the empirical probability density function (PDF) of daily precipitation intensities over the four subregions. The TRMM7-0.25 dataset, aggregated to the common 1.76° resolution (TRMM7-1.76), as expected has a shorter right tail with no precipitation intensities larger than 100 mm day-1 and higher frequency for lower intensities less than 25 mm day-1 (Fig. 6a,e,i,m). The two TRMM7 PDFs provide reference bounds for datasets with resolution between 0.25° and 1.76°. However, uncertainties in gridded daily precipitation products in Africa are large (Sylla et al., 2013) and we take the TRMM bounds as an observational approximation focusing more on differences in the simulated PDFs across the four resolutions. Over West, East and central Africa ERAINT overestimates the frequency of low (< 10 mm day-1) and extremely high (>150 mm day-1) intensities while it underestimates the frequency of precipitation intensities in between (Fig. 6a,e,i), especially over West Africa (Fig. 6a). In southern Africa (Fig. 6m) ERAINT represents the frequency of daily mean precipitation more accurately compared to the other three regions but shows almost no events with more than 150 mm day-1 in contrast to the observations. Both RCMs, in general, have the same tendency to generate more higher-intensity precipitation events with increasing resolution over all four subregions. In West Africa RCA4-v1 strongly underestimates the frequency of intensities with more than 20 mm day-1 at 200, 100 and 50km (Fig. 6b). A substantial improvement appears only at 25km where the right tail of the PDF extends up to 250 mm day-1, although the frequency of precipitation events from about 50 to 150 mm day-1 is still underestimated.
The RCA4-v4 configuration markedly reduces the RCA4-v1 biases and shows more realistic PDFs at all four resolutions (Fig. 6c). The RCA4-v4 50km simulation generates precipitation events up to 250 mm day\(^{-1}\) strongly contrasting to the RCA4-v1 simulation at the same resolution (no events more than 100 mm day\(^{-1}\)). However, RCA4-v4 overestimates frequencies of high intensities at 25km. Such sharp difference between two configurations of RCA4 at the same resolution shows that model formulation also plays an important role for accurately reproducing daily precipitation. Over West Africa all HCLIM-ALADIN simulations overestimates the frequency of low precipitation intensities (less than 10 mm day\(^{-1}\)) and underestimates the frequency of intensities in the range of 10-150 mm day\(^{-1}\) (Fig. 6d). Similar to RCA4, higher resolution leads to more high-intensity precipitation events in the HCLIM-ALADIN simulations.

However, RCA4 and HCLIM-ALADIN behave in a different way with increasing resolution. Both RCMs change the PDFs by adding more higher-intensity precipitation events extending the right-hand tail towards higher intensities. In addition, RCA4 also increases the frequency of medium- and high-intensity events especially going from 50 to 25km. In eastern Africa both RCA4 configurations reproduce the observed PDFs almost perfectly (Fig. 6f, g). All four resolutions are located within the TRMM-1.76 and TRMM-0.25 boundaries and the coarsest and finest resolutions coincides with the respective TRMM PDFs. Contrastingly, HCLIM-ALADIN strongly underestimates the frequency of precipitation events with more than 20 mm day\(^{-1}\) (Fig. 6h) over eastern Africa and even the highest 25km resolution is located below the coarse TRMM-1.76 dataset. In central Africa both RCMs overestimate the occurrence of intensities less than 20 mm day\(^{-1}\) (Fig. 6j,k,l), especially HCLIM-ALADIN (Fig. 6l) and strongly underestimate
the frequency of higher-intensity events. The PDFs at all four resolutions for both RCMs are located below the coarsest TRMM-1.76 PDF. We note that observational uncertainties in precipitation are very large over central Africa and we should be careful in the interpretation of Fig. 6j-l. Seasonal mean precipitation, for example, can differ by more than 50% across different observational datasets (Washington et al., 2013). Additionally, the TRMM dataset is scaled by the gauge-based GPCC precipitation product while almost no long-term gauges are available in the region (Nikulin et al., 2012). In southern Africa RCA4 and HCLIM-ALADIN simulate the precipitation PDFs pretty accurately (Fig. 6n-p). An interesting detail is that the 50km HCLIM-ALADIN simulation shows higher frequency for intensities with more than 150 mm day$^{-1}$ in the range of 50 to about 200 mm/day than the 25km simulation.

In general, we see the improvement of simulated daily rainfall intensities with increasing resolution across the African continent. There are many studies showing a similar resolution-dependent improvement over both complex terrains and flat regions (e.g. Chan et al., 2013; Huang et al., 2016; Lindstedt et al., 2015; Olsson et al., 2015; Prein et al., 2016; Torma et al., 2015a, 2015; Walther et al., 2013). Our results are in agreement with the above studies and confirm increasing fidelity of simulated daily rainfall intensities with increasing resolution.
Figure 6. Probability distribution function of daily precipitation intensities pooled over the four subregions for 1998-2010 in observations/ERAINT and as simulated by RCA4 and HCLIM-ALADIN at the four different resolutions. TRMM7-1.76 represents TRMM7-0.25 aggregated from its native 0.25° resolution to 1.76°. A base-10 log scale is used for the frequency axis and the first bin (0-1 mm day⁻¹) is divided by 10. Only land grid boxes are used for pooling over the subregions and the season is different for the different regions.
Summary and Conclusion

In this study we have investigated the impact of model formulation and spatial resolution on simulated precipitation in Africa. A series of sensitivity, ERA-Interim reanalysis-driven experiments, were conducted by applying two different RCMs (RCA4 and HCLIM-ALADIN) at four resolutions (about 25, 50, 100 and 200 km). The 100km experiment, at resolution a bit coarser than the driving ERA-Interim reanalysis, by default does not provide any resolution-dependent added value while such added value is expected for the 50 and 25km experiments. The 200km experiment is about 3 times upscaling of ERAINT to resolution of many CMIP5 GCMs and should only be considered as a supplementary experiment since RCMs do not aim to operate at such coarse resolution. In addition, to the two different RCMs, the standard CORDEX configuration of RCA4 is supplemented by another configuration with reduced mixing in the boundary layer. Such configuration was developed to deal with a strong dry bias of RCA4 in Central Africa. Contrasting the two different RCMs and the two different configurations of the same RCM at the four different resolutions allow us to separate the impact of model formulation and resolution on simulated rainfall in Africa.

Even if the results often depend on region and season and a clear separation of the impact of model formulation and resolution is not always straightforward, we found that model formulation has the primary control over many aspects of the precipitation climatology in Africa. The 100km no added value experiment (NAVE) shows that patterns of spatial biases in seasonal mean precipitation are mostly defined by model formulation. These patterns are very different between the driving ERAINT and RCMs, sometimes even with opposite signs, exemplified
by the two configurations of RCA4 in JAS (Fig. 1e-l). Resolution in general controls the magnitude of biases and for both RCA4 and HCLIM-ALADIN higher resolution usually leads to an increase in precipitation amount while preserving large-scale bias patterns. A side effect of such an increase in precipitation amount is that an improvement in one region (e.g. reduction of dry biases) often corresponds to a deterioration in another region (amplification of wet biases) as for HCLIM-ALADIN in JAS (Fig. 1m-p). Nevertheless, on average the smallest biases in seasonal means are found for the simulations at 50 and 25km resolution.

The impact of model formulation and resolution on the annual cycle of precipitation is mixed and strongly depends on region and season. For example, in both West and Central Africa the shape of the annual cycle for the 100km no added value experiment NAVE is different from ERAINT. However, the impact of model formulation is opposite between these two regions. In West Africa both RCMs deteriorate the ERAINT annual cycle by simulating a too early onset of the rainy season. In contrast, over Central Africa, both models improve the ERAINT annual cycle by reducing a strong wet bias and changing the unimodal annual cycle to a bimodal one similar to the observations. The impact of resolution can also be different. In West and East Africa, higher resolution (50 and 25km) leads to an improvement in the annual cycle (more realistic shape and smaller biases). In contrast, over Central Africa, the 25km RCA4 simulations show the largest biases while the HCLIM-ALADIN simulations at all four resolutions are almost similar. In general, it is difficult to conclude on a common impact of model formulation and resolution on the annual cycle.

The phase of the diurnal cycle in Africa is completely controlled by model formulation (convection scheme) while its amplitude is a function of resolution. Both ERAINT and
HCLIM-ALADIN shows a too early precipitation maximum around noon while RCA4 simulates a much more realistic diurnal cycle with an evening maximum. Higher resolution does not change the phase of the diurnal cycle but its amplitude, although the impact of resolution on the amplitude is mixed across the four subregions and time of the day.

A pronounced and well-known impact of higher resolution on daily precipitation intensities is a more realistic distribution of daily precipitation. Our results also show that higher resolution, in general, improves the distribution of daily precipitation. This includes reduced overestimation of the number of days with low precipitation intensities and reduced underestimation of the number of days with high intensities. The latter results in extending the right-hand tail of the distribution towards higher intensities similar to observations. This also means that at higher resolutions the time-mean climate (e.g., seasonal mean and annual cycle) is made up of more realistic underpinning daily precipitation than at lower resolutions. It is also worth emphasizing that if low resolution models are not able to simulate high rainfall days then it will be difficult for them to say anything robust about projected climate changes in high rainfall events. However, regionally, model formulation can also play an important role in the distribution of daily precipitation. For example, in West Africa the 50km RCA4-v4 configuration with reduced mixing in the boundary layer shows a remarkable improvement in the shape of the PDF (Fig. 1c) compared to the standard RCA4-v1 configuration at the same resolution (Fig. 1b). Moreover, the RCA4-v4 configuration at 50 km shows almost the same PDF as RCA4-v1 at 25km. Such contrast indicates that for daily precipitation intensities model formulation can have the same impact as doubled resolution.
Improvements in simulated precipitation in high resolution RCMs relative to coarse-scale GCMs are often attributed as being an resolution-dependent added value of downscaling. Our results show that for Africa improvements are not always related to higher resolution but also to different model formulation between the RCMs and their driving reanalysis. A common framework for quantifying added value of downscaling is to evaluate some aspect of the climate in high-resolution RCM simulations and in their coarse-resolution driving reanalysis or GCMs over a historical period (Di Luca et al., 2015; e.g. Hong and Kanamitsu, 2014; Rummukainen, 2016). If the RCM simulations show smaller biases compared to reference observations than the driving GCMs, one can conclude that RCMs provide an added value and vice versa. However, such a framework does not separate the impact of different model formulation between RCMs and their driving GCMs and higher resolution in the RCM simulations. Our results indicate that improvements in RCM simulations may simply be related to different model formulation and not necessarily to higher resolution. In general, model formulation related improvements cannot be considered as an added value of downscaling as such improvements are strongly model dependent and cannot be generalised.

However, such formulation-related and region-specific improvements from RCMs could in principle be also used in GCMs.

Within commonly used RCM evaluation framework, e.g. the CORDEX evaluation experiment, it is not straightforward, if possible at all, to isolate the impact of model formulation and resolution in RCM simulations. We show that running RCMs at about the same resolution as a driving reanalysis (e.g. ERAINT at about 80km or ERA5 at about 30km) helps to separate the impacts of model formulation and higher resolution in dynamical downscaling. We propose that such a
simple additional experiment can be an integral part of the RCM evaluation framework in order
to elucidate the added value of downscaling. In our study, as the first step, we focus only on
precipitation that has large relevance for climate change impact studies. As the next step, we
foresee similar studies looking also at other variables and especially at processes and drivers
relevant for regional climate.
Moreover, the same NAVE framework can be used for quantifying the added value in
RCM-based future climate projections. For this, one needs to downscale GCMs at their native
resolution in addition to the standard CORDEX resolutions (25 or 50km). The RCM projections
at the native GCM resolution serve as the NAVE in the climate change context. A potential
caveat, already mentioned in our study, is that RCMs are generally developed and tuned to
operate at resolution of tens of km. “Downscaling” a GCM at its native resolution, for example
150 or 200km, may lead to artefacts related to a lack of RCM retuning for coarser resolution.
Nerveless, more and more GCMs, for example in CMIP6, have resolution finer than 100km that
allows application of the NAVE.
Code availability

The analysis is done in MATLAB and IDL and codes can be provided by request as they are but without support on implementing them in another computing environment.

Data availability

The ERA-Interim reanalysis is available at https://apps.ecmwf.int/datasets/, the GPCC dataset is available at https://www.dwd.de/EN/ourservices/gpcc/gpcc.html, the CRU dataset is available at https://catalogue.ceda.ac.uk/uuid/5dca9487dc614711a3a933e44a933ad3 , the UDEL dataset is available at http://climate.geog.udel.edu/~climate/html_pages/download.html, the TRMM dataset is available at https://pmm.nasa.gov/data-access/downloads/trmm. The RCA4 and HCLIM-ALADIN data can be provided by request.

Author contribution

MW performed RCA4 simulations and all the analysis and wrote the initial draft. GN developed the experiment design and provided guidance for the analysis. EK and GN revised the initial draft. CJ is responsible for setting up the new RCA4 configuration (v4). DB and DL are responsible for performing the HCLIM-ALADIN simulations over Africa. All the authors contributed with discussions and revisions.

Conflict of interest

There is no conflict of interest in this study.

Acknowledgements

We thank one anonymous reviewer and John Scinocca for their very helpful comments. This work was done with support from the FRACTAL (www.fractal.org.za) and AfriCultuReS (http://africultures.eu/) projects. FRACTAL is part of the multi-consortia Future Climate for Africa (FCFA) programme - jointly funded by the UK’s Department for International Development (DFID) and the Natural Environment Research Council (NERC). AfriCultuReS has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 774652. The authors thank the European Centre for Medium-Range Weather Forecasts (ECMWF), the Global Precipitation Climatology Centre (GPCC), the British Atmospheric Data Centre (BADC), the University of East Anglia (UEA),
the University of Delaware and the Goddard Space Flight Center (GSFC) for providing data. All simulations were conducted on the supercomputer in the National Supercomputer Centre, Linköping, Sweden.
Reference


Covey, C., Gleckler, P. J., Doutriaux, C., Williams, D. N., Dai, A., Fasullo, J., Trenberth, K. and Berg, A.: Metrics for the Diurnal Cycle of Precipitation: Toward Routine Benchmarks for Climate


Favre, A., Philippon, N., Pohl, B., Kalognomou, E.-A., Lennard, C., Hewitson, B., Nikulin, G.,


Kjellström, E., Nikulin, G., Strandberg, G., Christensen, O. B., Jacob, D., Keuler, K., Lenderink, G., van Meijgaard, E., Schär, C., Somot, S., Sørland, S. L., Teichmann, C. and Vautard, R.: European climate change at global mean temperature increases of 1.5 and 2 °C above pre-industrial conditions as simulated by the EURO-CORDEX regional climate models, Earth


Moufouma-Okia, W. and Jones, R.: Resolution dependence in simulating the African


model configurations AROME CY41T1 and ALARO CY40T1, Geoscientific Model Development, 11, 257–281, 2018.


Van der Linden, P. and Mitchell, E., JFB: ENSEMBLES: Climate change and its impacts—Summary of research and results from the ENSEMBLES project, 2009.


