Reply to Anonymous Referee #2

General comments:

Messmer et al. present sensitivity simulations with a convection-permitting configuration of the atmospheric model WRF over Mt Kenya. The authors evaluate the impact of several parameterization collections and nest configurations (total number, ratios and interactions), with a focus on monthly total precipitation and monthly mean near-surface air temperature, for a study period of the year 2008. The presented work is intended to determine an optimal configuration for climate simulations over this region. Applying WRF at kilometer-scale grid spacing over Mt Kenya is a novel contribution, and an improved understanding of current and future variability in precipitation and water resources in this region is of high societal relevance. However, I have a number of concerns about the presented simulations and the analysis that need to be addressed before I can support the publication of the manuscript.

Thank you for going through our manuscript so carefully and for taking the time to review it. Thank you for providing insightful comments which improve the quality of our paper.

Regarding the simulations:

1. *The authors are using WRF V3.8.1, a version of the model that is four years old. The current version is 4.2.1 and numerous code improvements and bug fixes have been issued in the meantime. A brief look at the reported changes shows a major update to the Grell-Freitas scheme and bug fixes for Noah-MP (V3.9; https://www2.mmm.ucar.edu/wrf/users/wrfv3.9/updates-3.9.html) as well as updates to WSM6 (V4.1; https://github.com/wrf-model/WRF/releases). Although ongoing code developments are relevant for all published WRF studies, the older model version employed here may limit the current applicability of the presented results. I suggest that the authors perform an additional simulation of their best-performing configuration (identified as Cumulus3 1-way) with the latest version of WRF to assess the potential impact of model version on their conclusions.*

   The used version V3.8.1 is a stable and widely used version of WRF. Since we have started first simulations more than two years ago, we have decided to stick to the same version, in order to keep some level of consistency throughout the course of the experiment. Additionally, since the model is updated quite frequently (every year there is a major update, while there are new subversions on an approximately 1 to 3-monthly basis), it is unusual to work with the most up to date version. Especially because this version is also often subject to errors and instability.

   We have, nevertheless, started a simulation with the optimal setting using the most recent WRF version 4.2.1. The currently used WRF version 3.8.1 was compiled with Intel, but as the cumulus parameterization 3 (Grell-Freitas) is not running with an optimized compilation on our high-performance computer, we had to use a different compiler. With PGI, WRF V4.2.1 is now running with our optimal setup, but it must be kept in mind that small changes might be introduced because of differences in the numeric solution of WRF when using another compiler. We will analyse the results of the simulation and decide if it is worthwhile to include additional figures or if a short notice is sufficient in the next version of the manuscript.

2. *It appears that all of the configurations with cumulus parameterizations (CPs) used these schemes in the 3- and 1-km grid spacing domains. Conversely, the No Cumulus configuration did not use a CP in any domain, including the 27- and 25-km grid spacing ones. The established approach is to explicitly resolve convection when grid spacings are less than a few kilometers (e.g., Weisman et al., 1997). Although recent work shows that it might be appropriate to neglect CPs at coarser grid spacings than previously thought (Vergara-Temprado et al., 2020), excluding a CP in D4 could mean that convective processes are inadequately resolved (as the authors propose at line381). Overall, the*

authors need to provide a better justification for the CP settings in their tested configurations and, ideally, perform sensitivity simulations to illustrate the impact of (not) using a CP at 3- & 1-km (27- & 25-km) grid spacing.

It seems that there is some misunderstanding, and we have realized that Table 1 in our manuscript is not precise enough. Hence, we will update Table 1 in the new version of the manuscript, so that it is clear that we are not using CP in domains that have a grid spacing of 5 km or less. This scale is certainly convection permitting and hence, no parameterization is needed here, as the reviewer points out correctly. For the larger domains it is generally suggested to use a cumulus parameterization and here we already perform various experiments with different schemes and also turning it off as suggested by Vergara-Temprado et al. (2020). Since we are not using a cumulus scheme in the fine domains, and because we have done several experiments with turning on or off the CP in the coarser domains, there is no need to perform additional sensitivity experiments.

Nevertheless, we will include a clearer motivation, why we turn off or make use of a cumulus parameterization in the next version of the manuscript.

3. *The finest resolution domain in the four-domain configuration is placed upstream of the main regional circulation systems, close to the lateral boundaries of its parent domain. It also appears to be more limited in extent than in the three-domain configuration. Both of these differences will impact the development of small-scale features. Furthermore, the experimental set up does not isolate the impact of the number of nests from the effects of the 1-km domain extent and proximity to boundaries.*

   The domain is more than 50 grid points away from the east boundary of D03 and more than 30 grid points from the northern edge of D03. This is much further away than the 5 to 10 grid points around the boundaries that are needed for the relaxation of the outer boundary conditions (Rummukainen, 2010). Hence, we do not expect any problems from the boundaries. Furthermore, as the grid spacing is already very fine, it is expected that the two domains have a rather similar solution as several processes can be explicitly described. We will provide a more zoomed in figure for the reviewer in the point-to-point response, as in Fig. 1 of the manuscript the domain borders are rather thick (and well visible), so that it seems that the domains are very close to each other.

   Rummukainen, M. (2010), State-of-the-art with regional climate models. WIREs Clim Change, 1: 82-96. https://doi.org/10.1002/wcc.8

A few additional comments on the numerical simulations are provided in the specific comments below.

Regarding the manuscript:

4. *The introduction is missing some literature (see specific comments). The analysis would benefit from considering higher temporal resolution data (only monthly totals or means are examined) as well as an enhanced focus on process understanding to provide more insight into the differences between configurations. For example, the differences in near-surface air temperature are attributed to precipitation differences without presenting any supporting analysis of, for example, simulated soil moisture, the latent heat flux, cloudiness, or net radiation. Finally, there is minimal discussion, including of any caveats that might impact the reliability of the results and conclusions (e.g., the sparsity of observations to the southeast of Mt Kenya and above 3048m, the choice of simulation year, etc).*

   We will expand the introduction with the suggested literature. Since reviewer 1 also requests a sub-daily analysis, we will try to include such an analysis, even if we are only able to perform this with IMERG, as all the other datasets and observations are available on a daily basis only. If this does not provide insightful results because of the restriction to IMERG, we can further investigate daily values or pentads.

We will further investigate the link between temperature and precipitation, by investigating LH, soil moisture, clouds etc as suggested by the reviewer. We will decide afterwards, if we will include further figures directly into the manuscript or the supplementary, or if a short notice in the text is sufficient.

Further, we will include a more extended discussion on the reliability of the data in the next version of the manuscript.

Specific comments:

1. *Line 27-28: The discussion of the impact of the Walker circulation changes on in-terannual precipitation variability should cite at least one of Stefan Hastenrath's papers on this topic (for example, Hastenrath and Polzin, 2004, 2005)*.

   Thank you for pointing this out, we will include this in the next version of the manuscript.

2. *Line 38: the long rains are also associated with flooding and drought events (e.g., Kilavi et al., 2018)*.

   We will be more precise with this statement in the next version of the manuscript.

3. *Lines 39-42: the introduction makes no mention of the Indian Ocean Zonal Mode and its significant impact on moisture variability in East Africa (e.g., Behera et al., 2006; Nicholson, 2015; Saji et al., 1999; Ummenhofer et al., 2009).*

   Thank you for pointing this out, it will be included in the next version of the manuscript.

4. *Lines 46-48: Wainwright et al., (2019) is relevant to the discussion of the long rains. Please, clarify what is meant by "downward trend".*

   We will clarify this in the next version of the manuscript.

5. *Line 55: Are the authors referring specifically to climate simulations?*

   Yes, we do. We will be more specific here in the next version of the manuscript.

6. *Line 70: Collier et al., (2018) also performed a decadal simulation with WRF at convection-permitting resolutions in East Africa.*

   We will include this literature in our introduction as well.

7. *Line 71: The authors should clarify in this sentence already that it is not only the scale but the ability to neglect a cumulus parameterization that has an impact*.

   We will make it clear in the new version of the manuscript.

8. *Line 107: To be pedantic, WRF is an atmospheric model that can be used for many applications, including regional climate modelling. Please rephrase.*

   We will rephrase it to: We adopt the numerical weather prediction model WRF …

9. *Section 2.1: the details provided on the WRF configurations are insufficient to re-produce the study results. Please provide additional details, in Table 1 or elsewhere, including the grid dimensions, selected surface layer scheme, the moisture trigger used with the KF parameterization, diffusion option, and upper boundary condition. Ideally, sample namelists would be made available to interested readers.*

   We will add our namelists to Zenodo, so that they are available for anybody.

10. *Line 116: Please provide information on the length and computational expense of the simulations.*

    All the experiments used up around 750'000 CPU hours and they fill up around 140 TB of storage space.

11. Line 127: The model top of 50 hPa is low for climate simulations (as per WRF developer recommendations and previous studies), especially over mountainous terrain.

50 hPa is the standard best practice p_top value for the WRF simulation (https://www2.mmm.ucar.edu/wrf/users/namelist_best_prac_wrf.html). Additionally, it is one of the most widely used value also in recent studies over complex terrain, e.g., over Europe (e.g., Dörenkämper et al., 2020) and this level also depends on the available levels from the input data, i.e., ERA5 or CESM for our later studies.

Dörenkämper, M., Olsen, B. T., Witha, B., Hahmann, A. N., Davis, N. N., Barcons, J., ... & Mann, J. (2020). The Making of the New European Wind Atlas–Part 2: Production and Evaluation. *Geoscientific Model Development*, *13*(10), 5079-5102. https://doi.org/10.5194/gmd-13-5079-2020

12. *Line 130: Please provide the permitted timestep range for each outer grid resolution (or a general relation, e.g., from 3* to 8*dx). Also, did the simulations employ restarts? If yes, were they reproducible with the adaptive timestep?*

    The range of the dt for the outermost domains is given in line 129-131 in the current version of the manuscript. If the reader needs more details about the exact setting, she/he can refer to the namelist on Zenodo, which we will upload before the next submission.

    We are using restart files and we will test if the restart files are reproducible for the upcoming revision. The result of this test will be added as a short discussion to the next version of the manuscript.

13. *Line 145: Whether or not using the lake model improves regional precipitation in the presented simulations could be tested explicitly.*

    This is correct. We have already started a new simulation, where the lake model is turned off. We will add the results to the next version of the manuscript.

14. *Are the authors using the default land use and terrain datasets as input? How representative are these datasets of conditions on and around Mt. Kenya, including the forest belts and grasslands on the slopes? Mölg and Kaser (2011) reported improved high-elevation simulations with WRF over Kilimanjaro using updated land use datasets, and both updated land use and terrain datasets have been employed in recent WRF studies to better represent surface conditions (e.g., Collier et al., 2018, 2019)*.

    It is true that there are other published studies where updated land use data is used. However, we decided to use the default data provided within the model, because in a next step of our study, we plan to run different climatologies for the region with different land use options. In order to identify the effect of changed land use, we first need a base state to compare it to.

15. *Line 142: Can the authors also provide which dveg option they are using with Noah-MP? There are issues with certain options for domains containing both hemispheres that could have a significant impact on the results: https://github.com/wrf-model/WRF/issues/707.*

    We have used the default option and we are also aware of the problem along the equator. However, when we started the first experiments, no solution for this issue was available except changing to another land surface model, something that we did not want to do. We have started a new simulation with the optimal setting and option dveg = 9, to estimate how significant the impact on the results of the simulation actually is.

16. *Line 168: Why were these pressure levels (and not all available levels) between1000 and 700 hPa selected? Does this choice significantly impact the simulations?*

    The plan is to use this model setup with input from a climate model, where we also have to interpolate the levels and it is not meaningful to just interpolate various levels from a limited number of sigma-pressure levels. Our intention is to have the model run as freely as possible, as we will need this for the climate simulations. Also, other climate models from, e.g., CMIP5 and 6 do not provide this many pressure levels. Hence, the selected levels are sufficient to drive a regional climate model.

17. *Line 170 & Figure 2: the data need to be detrended, if significant trends are present, to examine anomalies. The months of April and May 2008 look very anomalously dry compared with the whole period, and the impact of choosing [only] 2008 as the study period on the results and conclusions needs to be discussed. Also, since the precipitation field in ERA5 is highlighted as being unreliable, would it be preferable to consider CHIRPS data for precipitation in Figure 2?*

    Thank you for this comment. We will detrend the data and redo the figure, also with CHIRPS to see how this changes the results. Furthermore, a more detailed discussion on the selected year will be added.

18. Line 178: Data should be interpolated from higher to lower spatial resolution, as interpolation does not add physically meaningful detail.

    It is correct that no detail is added to the coarser resolution if we interpolate from the lower to the higher resolution. However, if we interpolate from lower to higher resolution, we lose the fine scale information of WRF, while the other way round we do not change the physical detail. Additionally, the grid of WRF is the common thread, while all the gridded data sets have different grid resolutions. Furthermore, we expect that if we interpolate the WRF grid points to a coarser grid, the spatial correlations would be higher, and the test is less restrictive. For these reasons we prefer to keep it the way it is right now. We will add a justification for this approach in the next version of the manuscript.

19. *Section 2.3.3: CHIRPS merges satellite and rain gauge data. Do the authors know if any of the weather station data been assimilated into this product?*

    As far as we know the observations gathered by CETRAD are independent from any other used station data.

20. *Lines 241-243: Please move this information to the methods and provide an estimate of statistical significance where applicable.*

    We will consider moving those lines to Section 2.1 as suggested, but we are not sure if that change improves the readability of the paper. The statistical significance of the results was not included in the current version as only 12 monthly values were used to calculate the temporal correlations. If the (sub-)daily analysis is performed in the next version, the statistical significance will be evaluated and included.

21. *With regards to the statistical methods, the authors do not state which correlation they use and if it is appropriate to have a sample size of 12. In addition, the data have not been de-seasonalized, which means that a relatively high correlation is to be expected unless WRF fails to capture the seasonal cycle, which should be acknowledged. For their analysis, I suggest that the authors consider a finer temporal resolution, such as pentads, to provide a more detailed and robust assessment of model skill.*

    As we said before, also reviewer 1 suggests to perform the same analysis but with a finer temporal resolution. This will be evaluated in the next steps. Additionally, we think that a deseasonalization with only one year is not possible, as you would need a longer period to capture a typical seasonal cycle. We will include a short note on this fact in the new version of the manuscript.

22. *Lines 327-328: where is this result visible?*

    This result is observable in the insets of Figure 5. There, it can be seen that the annual precipitation sums for the Cumulus 3 options is below the one from ERA5. We will add a label towards Figure 5 in the text, to make it clearer.

23. *Lines 331-334: this is the first time that a justification is provided for using an outer domain with a grid spacing of ~25 km. I suggest moving this information to the methods, so the reader understands why the authors include this aspect earlier.*

    This is a valid point. We will do so in the next version of the manuscript.

24. *Figures 6 to 11: The choice of months could be more robustly justified. In addition, both June and November show low pattern correlations between the observations and the gridded datasets, which undermines using CHIRPS as the main comparison (lines 335-337). Overall, these figures take up a lot of space in the manuscript without providing a great deal of new information. Some suggestions would be to remove the panels for ERA5 and Europe (the authors have already established the poor representation of precipitation) or to replace some map figures with elevational profiles, for compactness and so that the reader can more clearly see some of the reported findings (e.g., lines 403 to 405).*

    We will consider your suggestions regarding those figures and the selection of the months.

25. *Lines 393-396: Please move this information to the methods.*

    We will move this information to Section 2.3.5 as suggested.

26. *Line 459: The authors repeatedly mention that the underestimation of precipitation in the Europe configuration stems from the LW and PBL parameterizations – could they discuss what difference in process representation might be underlying the difference?*

    We will try to include more information regarding the differences in the next version of the manuscript.

27. *Line 464-465: Can the authors please clarify what they mean about one-way vs two-way nesting?*

    In the two-way nesting option, the results from the nest overwrite parent domain results. In the one-way nesting, the results from the nests do not overwrite the data in the parent domain and independent results are obtained in each domain. We will make this difference clear in the new version of the manuscript.

Technical comments:

1. *Line 129: The sentence "This is, because..." is unnecessary, please remove.*
2. *Line 132: Please change "in order to improve" to "to optimize."*
3. *Line 164 to 167: The forcing variables are well known and documented, so I suggest deleting these sentences.*
4. *Lines 190 and 201: It is clear that data are only compared for the study period, please remove.*
5. *Lines 234 to 237 are repetitive and could be removed.*
6. *Line 341: CHIRPS is not a model, please rephrase.*
7. *Line 344: "Bearing"*
8. *Line 436-437: Please clarify that these parameterizations are not varied independently.*

We will address all these technical comments in the new version of the manuscript as suggested by the reviewer.

Figures:

Figure 1: I suggest adding the weather station locations to the plots of D3/D4, as the reader does not see where they are located before encountering Figure 6. Also, is Figure 1 referenced in the text?

We will add the observations to Figure 1 and add the missing reference in the text. Thank you for pointing this out.

References:

Behera, S. K., Luo, J. J., Masson, S., Delecluse, P., Gualdi, S., Navarra, A. and Yamagata, T.: Impact of the Indian Ocean Dipole on the East African Short Rains: A CGCM Study* Swadhin, J. Clim., 19(7), 1361, doi:10.1175/JCLI9018.1, 2006.

Collier, E., Mölg, T. and Sauter, T.: Recent atmospheric variability at Kibo summit, Kilimanjaro, and its relation to climate mode activity, J. Clim., 31(10), 3875–3891, doi:10.1175/JCLI-D-17-0551.1, 2018.

Collier, E., Sauter, T., Mölg, T. and Hardy, D.: The influence of tropical cyclones on circulation, moisture transport, and snow accumulation at Kilimanjaro during the 2006 - 2007 season, J. Geophys. Res. Atmos., 124(13), 6919–6928, doi:10.1029/2019JD030682, 2019.

Hastenrath, S. and Polzin, D.: Exploring the predictability of the “short rains” at the coast of East Africa, Int. J. Climatol., 2004.

Hastenrath, S. and Polzin, D.: Mechanisms of climate anomalies in the equatorial Indian Ocean, J. Geophys. Res. Earth Surf., 110(D8), D08113, 2005.

Kilavi, M., MacLeod, D., Ambani, M., Robbins, J., Dankers, R., Graham, R., Helen, T., Salih, A. A. M. and Todd, M. C.: Extreme rainfall and flooding over Central Kenya Including Nairobi City during the long-rains season 2018: Causes, predictability, and potential for early warning and actions, Atmosphere (Basel)., 9(12),472, doi:10.3390/atmos9120472, 2018.

Mölg, T. and Kaser, G.: A new approach to resolving climate-cryosphere relations: Downscaling climate dynamics to glacier-scale mass and energy balance without sta-tistical scale linking, J. Geophys. Res. Atmos., 116(16), doi:10.1029/2011JD015669, 2011.

Nicholson, S. E.: Long-term variability of the East African ‘short rains’ and its links to large-scale factors, Int. J. Climatol., 35(13), 3979–3990, 2015.

Saji, N. H., Goswami, B. N. and Vinayachandran, P. N.: A dipole mode in the tropical Indian Ocean, Nature, 401, 360–363, 1999.

Ummenhofer, C. C., Sen Gupta, A. and England, M. H.: Contributions of Indian Ocean sea surface temperatures to enhanced East African rainfall, J. Clim., 22, 993–1013, doi:10.1175/2008JCLI2493.1, 2009.

Vergara-Temprado, J., Ban, N., Panosetti, D., Schlemmer, L. and Schär, C.: Climate models permit convection at much coarser resolutions than previously considered, J.Clim., doi:10.1175/JCLI-D-19-0286.1, 2020.

Wainwright, C. M., Marsham, J. H., Keane, R. J., Rowell, D. P., Finney, D. L., Black, E. and Allan, R. P.: ‘Eastern African Paradox’ rainfall decline due to shorter not less intense Long Rains, npj Clim. Atmos. Sci., doi:10.1038/s41612-019-0091-7, 2019.

Weisman, M. L., Skamarock, W. C. and Klemp, J. B.: The Resolution Dependence of Explicitly Modeled Convective Systems, Mon. Weather Rev., 125(4), 527–548, 1997