Point-by-point Reply to Referee 2

I would like to thank the authors for their careful and detailed responses to all of the reviewer comments, as well as for performing numerous additional sensitivity simulations to evaluate the robustness of the presented results and conclusions. The authors' revisions and clarifications have greatly improved the manuscript and the revised figures provide the relevant information in a succinct and compact way. I support its publication in GMD subject to the consideration of the small comments below in the final version.

1. I was curious what the difference between the 'South America' and 'Cumulus3' configurations for the 5km_D02 simulations was, because they appear identical from Table 1. A diff of the provided namelists indicates that they differ in the deep convection parameterization, which was turned on for the 5-km domain in the 'South America' configuration. I suggest that the authors check that the sample namelist is correct or amend the description of which domains employed a cumulus parameterization at line 153 in the revised manuscript/Table 1.

Thank you very much for pointing this out. There was indeed an error in two of the 5km_D02 simulations ("Europe" and "South America"), as they included a cumulus parameterization in the outermost domain. This was a mistake and we have rerun the "Europe" simulation for this setup, now without any cumulus parameterization as describe in the manuscript. As the "South America" and "Cumulus3" are the same for the 5km_D02 setting, we have deleted the "South America" setting from the Figures (3, 4, and 5), as this is then consistent with the simulation for the "No Cumulus" setting (because it is the same as "Cumulus 3 – 1way"). In Table 1 and in the method part we indicate that the two simulations are identical and hence not shown for "South America".

2. Figure R2.1 The results of the version testing are interesting and, in my opinion, relevant for readers interested in using the presented work to inform their own simulations. I suggest adding one sentence to the methods indicating that, compiler differences aside, using the more recent version of WRF essentially degrades the performance.

We agree that this point is interesting for the reader and we have added some lines in the manuscript that highlight this fact. Nevertheless, we do not agree with the word "degrade", as the new version is better in some months and worse in others. We have slightly weakened this statement to:

"In addition, a simulation with the latest version of the model (V4.2.1) was run. However, it showed that the included improvements are not enough to reduce the RMSE and to improve the temporal correlation against the weather station data compared to the other sets of experiments. It further indicates that model versions and compilers can impact the simulations performed with WRF. Consequently, it has not been included in the analysis presented here."

3. Figure S1 & Line 410: "Nevertheless, it must be noted that peak precipitation rates on a sub-daily basis are captured much more realistic with respect to IMERG, compared to the other parameterizations options, but only in domains where the others use a cumulus parameterization (see supplementary Fig. S1)."

It is not clear to me why the authors have included the right column (9-km/D2 in 27km_D4 set up) since the stated goal is to evaluate the impact of different configurations on the kilometer-scale solution. In addition, I don't see the result that the mean diurnal cycle is "much more realistic" in 'No Cumulus.' Although the timing of the afternoon peak is later (sometimes too late), the magnitude is very poorly captured. Can the authors clarify further the added value of showing and emphasizing these data?

We think that this is an interesting result for the reader. The topic on the proper scale to turn off the cumulus parameterization and on the added value of convection resolved model simulations is currently strongly debated. This is why we would like to show also D02 in the Supplementary Material, as it shows resolutions used for example by CORDEX.

Nevertheless, we have adapted the statement and have pointed to the fact that mainly the timing is improved and not necessarily the precipitation amounts.