Point-by-point Reply to Referee #1

Comment published online on the 20th of December, 2021

In their work, González-Rojí et al. present an assessment of the WRF model skill at simulating precipitation over parts of South America, with a focus on parts of Peru, Brazil and Bolivia. The authors explore different configurations of the WRF model, evaluating the performance of the model by comparing precipitation outputs to observations from weather stations and estimates from observation-based products and ERA5. From their analysis, the authors are able to identify which of the studied model configurations work best for their region of interest. In addition, the authors identify some strengths and weaknesses from observation-based products, like PISCO and CHIRPS.

The manuscript is well written and the topic is very relevant both for the climate science community of South America and the convection-permitting modeling community. I recommend this study for publication after minor revisions. Please see the details below.

Thank you for reading our manuscript so carefully and for your positive and constructive comments.

Major comments:

1. Taking further advantage of the high resolution simulations:

Simulations at such high resolution are very valuable for the region of interest. The analysis of the monthly accumulated precipitation is very interesting. A further analysis of 5-day or daily accumulated precipitation would help to strengthen the paper. For example, the analysis of the statistical distributions of the daily accumulated precipitation would help to identify the value of high resolution simulations at representing extreme precipitation.

As stated in the manuscript, the temporal analysis at finer temporal resolutions such as 15-days, 10-days, pentads or daily was also carried out for our analysis, and the same results are observed for the different intervals. In general, the RMSE increases and the temporal correlations decrease as we increase the temporal resolution. The statement in the manuscript is based on figure R1.1, that is neither shown in the manuscript nor in the supplementary material. Hence, we will add this figure to the supplementary material.

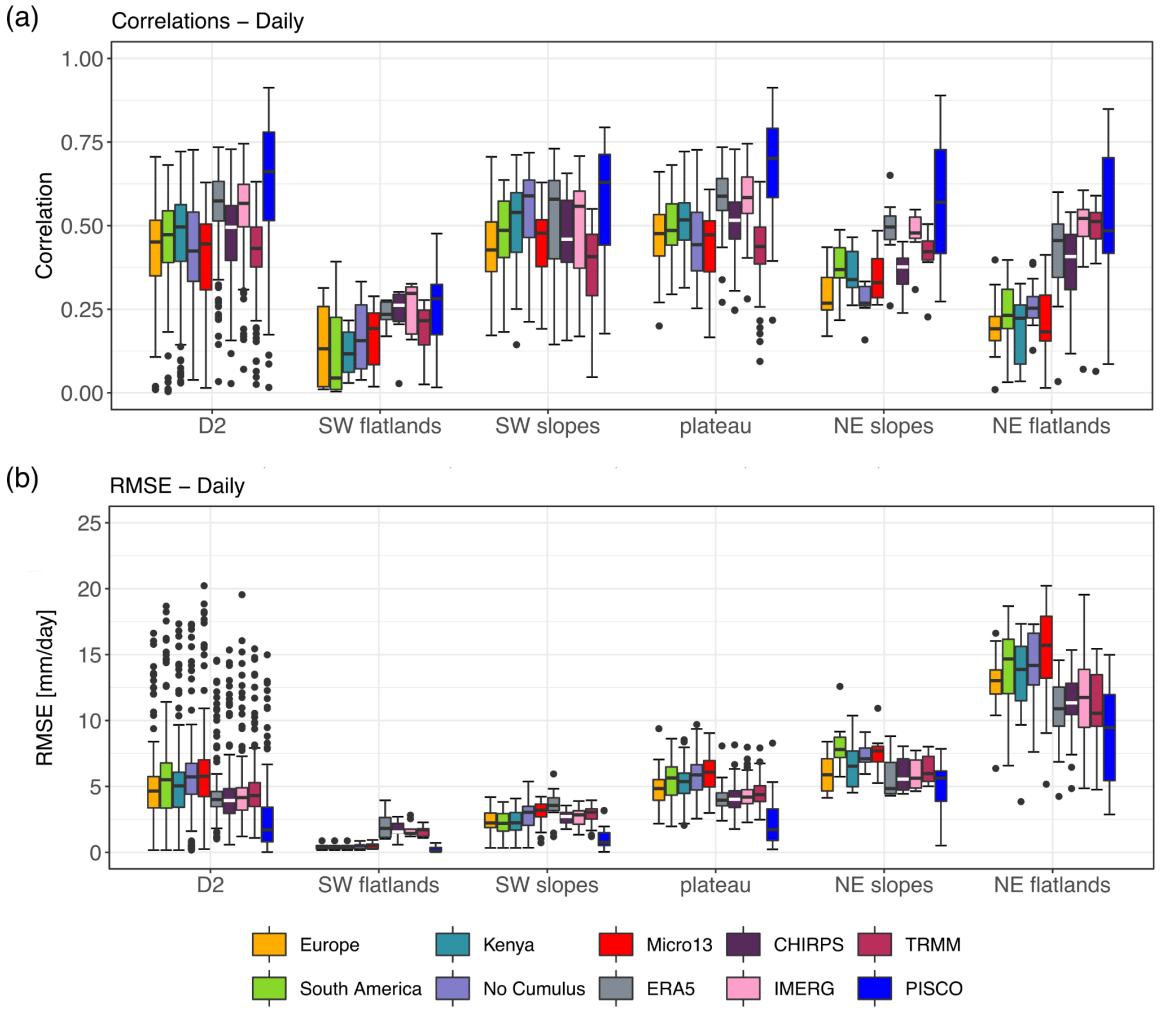

Figure R1.1: (a) The temporal correlation and (b) root-mean-square error (RMSE) between the annual cycle for the year 2008 of measured and simulated daily precipitation sums at the nearest grid point to the station's location shown for the different parameterization options and gridded observational datasets. The whiskers extend to the value that is no more than 1.5 times the inter-quartile range away from the box. The values outside this range are defined as outliers and are plotted with dots.

In addition, in the last section about the diurnal cycle, it would be interesting to add the observations, if hourly data is available.

We agree that it would be interesting to include the observations to the plot, but the weather station data obtained for the validation of the sensitivity experiments is only available with daily temporal resolution. For precipitation we have added the data of IMERG (Fig. R1.2), as it is available with 30 min temporal resolution, but it must be noted that IMERG is not the best gridded data set available in any of the five regions studied. For all the other variables we are not aware of any data set with higher than daily temporal resolution. Figure R1.2 shows that Micro13, South America and Kenya are able to capture relatively well the precipitation of the first half of the day, but they all miss the peak in precipitation in the afternoon. Conversely, NoCumulus is able to simulate a peak of precipitation during the afternoon. We will replace this panel in the new version of the manuscript.

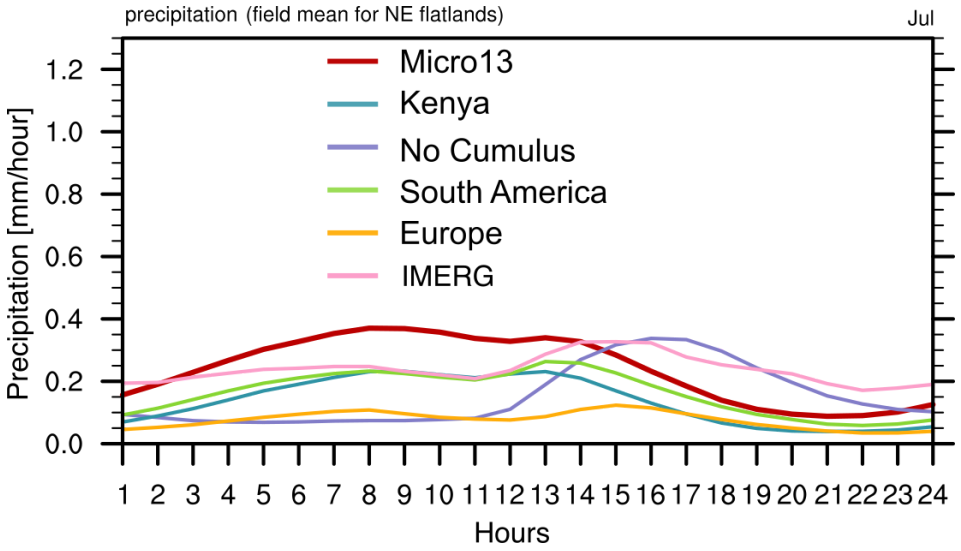

Figure R1.2: Monthly mean daily cycle for July of a field mean over the northeastern flatlands for precipitation (mm) including also IMERG (pink line).

2. More details about the configuration and domains:

Include some other standard details about the simulations, like the number of vertical levels, model top, type of nudging (if used). In addition, please be more explicit about which of the domains is used in each part of the paper (for example state explicitly whether results from D02 or D03 are used in section 3.4, and so forth).

All the sensitivity simulations include 49 vertical eta levels until the model top at 50 hPa, and the adaptive time step was employed while running the simulations. No nudging was applied to the input data. These details will be added to section 2.1 as suggested. Additionally, we will be more specific about the fact that we only show results from the second domain (D02).

3. A small extension of the relationship between variables:

Both figures 8 and 9 are very interesting, as they allow to talk about possible relations between variables looking for an explanation about the behaviour of the simulated precipitation. It would be very interesting to see this analysis extended to at least one of the other regions studied in Figure 3. The authors could select for comparison, for example, the regions in flatlands vs. regions over the plateau, or over the slopes, where the simulated cloud field (both in terms of magnitude of cloud fraction and periodicities -e.g. annual and diurnal cycle-) might be qualitatively different.

The main reason why we only evaluated the seasonal and daily cycles of different variables over the northeastern flatlands was because our region of interest (the department of Madre de Dios) is located there. It is true that the cloud cover is different depending on the region, and we can repeat the analysis for the plateau. Depending on the outcome of this analysis we will consider including or not an additional figure to the supplementary material, as the manuscript is already quite long.

4. A comment about the order of the sections:

The manuscript is well written, and the sections are clear. However, it seems to me more clear to start with section 3.3, where the mean biases are presented, and then go to sections 3.1 (temporal correlations) and 3.2 (spatial correlations) where second-order metrics (correlations and RMSEs) are studied.

We started analyzing the RMSEs and the temporal and spatial correlations as they provide a quantitative way to evaluate the performance of the simulations compared against weather stations and gridded observational data sets. However, it is true that we can also start section 3 by showing the accumulated precipitation maps first, and then continue with the more quantitative analysis of the results. We will consider adapting the structure of the paper if the consistency or the story line are not affected by this change.

Minor comments:

L147-149. The authors write:

"Based on previous studies by the authors, the "Europe" experiment includes the updated parameterizations used over that region (Messmer et al., 2017), i.e., Noah-MP instead of Noah land surface scheme".

One could interpret that only the "Europe" run uses the Noah-MP scheme. But in previous lines it is stated that the Noah-MP LSM is used in all runs. Please clarify.

As pointed out by the reviewer, the sentence was not clear enough. In the new version of the manuscript, we will change it to:

"*Based on previous studies by the authors, the "Europe" experiment includes the typical parameterizations used over that region (Messmer et al., 2017), but with the updated version of the Noah land surface scheme."*

L152-153:

"The "South America" experiment takes as a reference the parameterizations used to simulate storms over the central Andes (Zamuriano et al., 2019)."

The reference to Zamuriano et al. 2019 in https://nhess.copernicus.org/preprints/nhess-2019-286/ (search on December 20, 2021) appears as:

"Review status: this preprint was under review for the journal NHESS. A final paper is not foreseen. "

"This preprint has been withdrawn."

Even though the manuscript is available at: https://nhess.copernicus.org/preprints/nhess-2019-286/nhess-2019-286.pdf, the authors should re-consider (o better justify) the citation of this reference.

Thank you for pointing this out. We will rephrase this sentence to justify this selection in a different way.

Throughout the paper the authors use "the parameterization options" when referring to WRF simulations. It would be easier and more standard to read simply "the model" or "WRF", or "run".

Throughout the paper the authors write "monthly precipitation sums". It would be more standard to write "monthly accumulated precipitation".

For example, the sentence:

"The Europe parameterization option simulates especially low monthly precipitation sums in the wet seasons"

could be written:

"The Europe run simulates especially low monthly accumulated precipitation in the wet seasons"

Thank you for these two suggestions, we will change these terminologies as suggested by the referee in the new version of the manuscript.

L269-271:

"For 5-day and daily intervals the values drop for the correlations and rise for RMSEs. The increase in the RMSEs and the reduction in the correlations are expected due to the fact that capturing the exact amounts of precipitation at the same time as the observations is rather challenging for the model"

I would not say that this is challenging for the model, but a consequence of sensitivity to initial conditions (internal variability, present even if the model were perfect). These are not weather forecasts, but a continuous climate run.

As we increase the temporal resolution of the analysis, the parameterization schemes play more and more an important role in the simulation, i.e., the exact point in time when a process such as cloud nucleation is invoked. As the simulation of this timing is a challenge for the regional climate model and the corresponding parameterizations, we argue that capturing the representation of the variables as observed with small temporal increments is more difficult for the model than capturing the coarser monthly values. From our perspective the internal variability should not play a major role, as the simulations are driven by the same reanalysis product, so the internal variability between the observations and the WRF runs should somehow align, as it is imposed through the boundary conditions of the model. We agree that this might play a role in case a real climate simulation is investigated, where the year 2008 in the model and in the observations would not be the same.

We will add parts of this explanation to the sentence pointed out by the referee to indicate which process is more challenging for the model to capture.

In Figure 5 you say "The bold numbers indicate the best option for each region". The figure would be easier to read if the corresponding names of the experiments (on the left) were also in bold face.

This is a valid suggestion and we will also change the names of the experiments to bold face with a larger number of months above the reference threshold.

In addition, usually the symbols like the asterisks used in Figure 5 are used when a correlation is statistical significant.

We agree with the referee that it is more common to use the asterisks to highlight the statistical significance of the results in the literature. However, as Fig. 5 is already complicated and full of details, in order to make it cleaner and clearer to the readers, we decided to use the asterisks to cross the statistically insignificant pattern correlations. The second column (the one related to the gridded datasets) is always significant, and otherwise, each pixel of that column would have been covered with the asterisks, and it would have complicated the interpretability of the figure.

L336: change from "patters" to "patterns"

Thank you for pointing out this typo, we will replace it as suggested.

Is it possible to add Obs in Figure 8a?

We will consider including the observations as well, but as we only have 15 stations compared to several thousands of grid points in the NE flatlands, we will check if and how such a comparison can be accomplished.

It is not clear to me which domain it is being used for section 3.4 and for the previous sections. For sections 3.1 to 3.3 where the authors using results from D02? Are results of section 3.4 from the D03 domain? Please clarify.

As already stated in one of the major comments we will be clearer about the fact that only results from D02 are shown in the analysis.

L384. Which is the area for the computation of the field means? The entire D03 domain? Please clarify.

Figure 8 was created by computing the field means of the different variables for the NE flatlands included in D02. This means that only the grid points with an elevation between 0 and 1000 metres were included in the analysis. We will clarify this in the new version of the manuscript and we will add figure R1.3 below to the supplementary material.

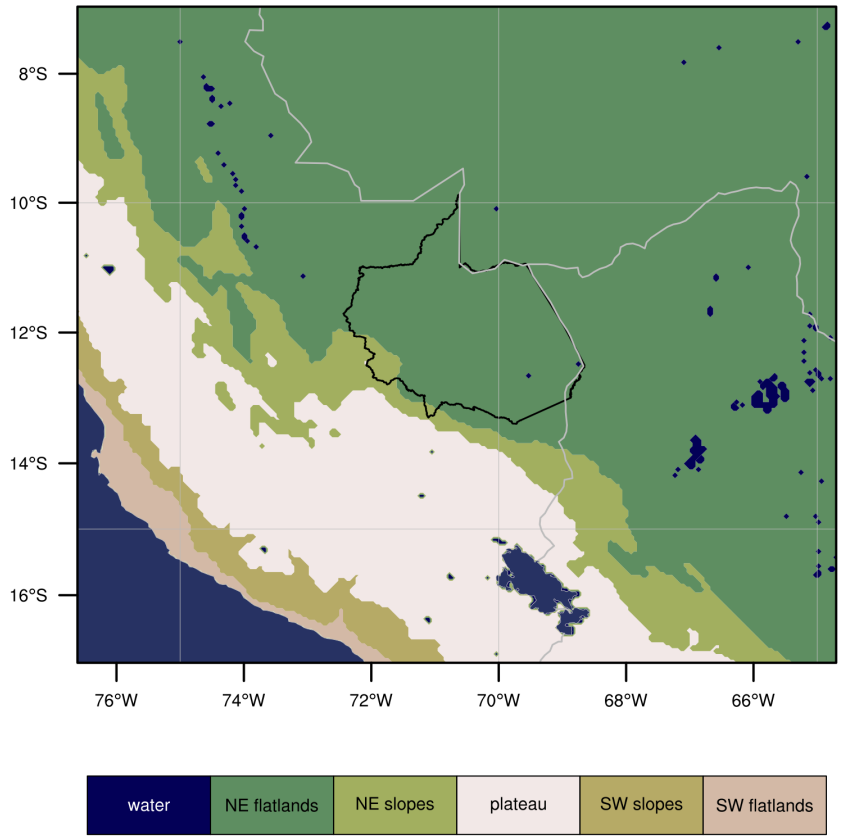

Figure R1.3: Map of the second domain highlighting the five regions included in the analysis with different shading.

Figs 8 and 9. According to their captions, these figures refer to the NE slopes, but section 3.4 is devoted to flatlands. Please clarify.

Thank you for pointing out this error. This is simply a typo in the titles of the panels. The lines correspond to the field means over the NE flatlands, and the title should state that. We will correct this error in the new version of the manuscript.

In addition, Figure 8 is very interesting. Would it be possible to include a similar figure for some of the other(s) regions studied in Figure 3?

As already stated in the corresponding major comment, we only analyzed the seasonal and daily cycles of the NE flatlands because our focus region is located in that area. We will consider including a new figure for the analysis over the plateau if the results are interesting and different enough, as the paper is already quite long.

L396-399. The authors write

"The relative humidity and the precipitable water of the No Cumulus parameterization option is especially low, even though the precipitation is comparable to the other options, which might indicate that this parameterization option has an efficient process to remove moisture from the atmosphere,

i.e., convective processes. This is also supported by the fact that precipitation occurs mainly in the afternoon, while the other options have precipitation distributed over the whole day (Fig. 9)."

Do the results in Figs. 8 and 9 come from domains D01, D02 or D03?

In case results in Figs. 8 and 9 are from D02 or D03, please explain how the use of No-Cumulus in D01 is affecting the results in D02 or D03. In particular, one would expect that since D02 and D03 do not use a cumulus scheme in any of the simulations, the diurnal cycle would be the same, even in the No-Cumulus run. This is an interesting point that the authors could explain a bit better.

As already stated before, we only analyzed the results from D02. We believe that the difference in the distribution of precipitation of the experiments is caused by the fact that only for NoCumulus the cumulus parameterization is switched off in D01, and not in the remaining runs. However, we will further investigate this by analyzing the daily cycle of the different variables over the same area (D02) in the first domain (D01), as the results of D01 are used as boundary conditions for D02. Hence, if moisture availability in D01 is shifted into the afternoon in the NoCumulus runs compared to in the morning, as in all the other runs, this will certainly affect the results in D02 and D03.

L428-430. The authors write

"The region of interest is the entire department of Madre de Dios, but because of the lack of a dense network of weather stations in that area we evaluate the performance of the model over a broader area including the tri-national border of Peru, Bolivia and Brazil."

I find this comment rather unnecessary. The authors do a fine job at assessing the WRF simulations with the available data for both the broader region in D02 and for the smaller region in D03 (which they say is the region of interest). In this sense, both domains D02 and D03 are the region of interest according to the results and analyses presented in the paper. Maybe the authors could just write something like:

"The region of interest includes parts of the tri-national border of Peru, Bolivia and Brazil, with a focus on the region of Madre de Dios. The analysis of the latter is challenging given the lack of a dense network of weather stations in the area".

Thank you for your suggestion, we will change it accordingly.