Review of "Evaluation of regional climate models ALARO-0 and REMO2015 at 0.22 resolution over the CORDEX Central Asia domain" by Top et al. 2020

General Comments

The authors surely put some efforts in trying to answer my comments on their previous version of the manuscript. Nevertheless I do not think that all these comments were exhaustively considered. In my opinion the paper still suffers from a series of major issues that need to be carefully addressed before it may be considered for publication for Geoscientific Model Development.

- The quality of the text and the structure of the manuscript are surely the points that have received more attention by the authors, but still some parts need revision. In particular you should check for consistency among the different subsections. For example in the subsections of the methods you specified what you used ERAInterim for, but you did not do the same for other data-sets such as GPCC. Another similar example can be found in the results subsection about DTR. Here you discuss the table with spatial means for maximum temperatures but not for minimum temperatures. Check for such inconsistencies throughout the text and correct them.

- You are not very accurate in the specification of the model behaviour and when you discuss the maps of the bias. I found a lot of inaccuracies in the text and I invite you to review it accordingly. Here a couple of examples:

  - **At the annual scale, the bias of the minimum temperature ranges mostly between -3C and 3C for REMO and between 0C and 5C for ALARO-0 (Fig. 5).**
    For REMo annual biases exceed the absolute value of 3 C over several areas such as Mongolia and the Himalayas. For Alaro a large part of the domain has a negative bias exceeding -5C in some case.

  - l. 258-261 **Based on Fig. 3, both RCMs perform best during autumn and the spatial correlation is lowest during summer for ALARO-0 while, the biases during summer are smaller than during winter and spring for both RCMs (Table 2 and Fig. 2)**
    From figure 2 you cannot say that the biases in summer are smaller than in winter and spring for all the points of the domain.

  - l. 452-454 **Fig. 2 and 4 show that for most parts of the domain the mean temperatures of ALARO-0 and REMO are lying within the range of spread between the reference datasets during autumn. From this we conclude that both**

RCMs simulate temperatures in autumn within the range of observational uncertainty.

This is not true for the entire points of the domain. When you propose such conclusions I invite you to first compare directly the map of the spread with the one of the bias (for example plotting their differences).

- My main concern is that despite my previous comments, even though you added an analyses of the seasonal cycle for sub-region in the new version of the manuscript, you did not conduct the same analysis for the mean bias, spatial correlation, standard deviation and RMSE. The information of the mean bias calculated over the entire domain intuitively makes no sense, as already highlighted in my previous report. You should conduct the analyses of the bias per sub-domain too. One interesting thing that I would suggest you to do is to consider mean absolute bias instead of the bias (for both the entire domain and sub-regions, since spatial biases might compensate each other). Additionally, also Taylor diagrams should be calculated for every sub-region. For this I also think it would be important for you to independently check the values of the spatial correlations you obtained, since they seem to be too high given the spatial patterns of the bias. Since you use pre-defined functions (in R) for calculating the Taylor diagrams, I think it would make sense to double check the correctness of the results, independently. Finally, the plots of the seasonal cycle should be improved. In particular it was impossible to understand the ones drawn for temperature.

- The discussion part does not always result very clear and I would suggest you to carefully revise it while modifying it in consideration of my new comments.

- As a final remark, it seems that the authors are a bit too positive about the models performance for the region. I would suggest them to try to be more objective in their conclusions. Maybe new analyses might help in this sense. In my personal opinion the models results cannot be considered reliable over a large part of the domain. It is true that over some areas there is an issue with the poor reliability of observations, but over some regions the main issue is still the model. Having a bias above 10C does not make the model reliable. This is for example the case of temperatures simulated by ALARO over the north-western and north-central part of the domain, both in winter and in summer. Appropriate bias correction methods could be used to make the model more in-line with "reality", but this is not inherent to model evaluation and should be made clear in the text.