

Referee Report on “Uncertainties in climate change projections covered by the ISIMIP and CORDEX model subsets from CMIP5” by Ito et al.

--- General comments

This manuscript presents a valuable study quantifying the spread of CMIP5 projections and biases covered by the subsets of models used in the ISIMIP and CORDEX experiments. This revised manuscript does a better job at explaining the added value of this study compared to previous studies.

Several points are raised in the body of the manuscript, which would benefit from a mention in the discussion section. These include: (i) the evaluation and selection of the high performance models was done independently for the two metrics used, i.e. temperature and precipitation. A next step would be to evaluate T and P jointly. Can the authors suggest whether that would impact the model selection? (ii) Similarly, the randomly selected ensembles can perform well in one variable, but could be evaluated on both variables. E.g. could one set of model perform well for both variables, including ΔT and ΔP ? Adding a few sentences to mention (i) and (ii) as limitations of the study and/or perspectives for future research would strengthen the discussion. (iii) The authors mention using “other variables” in addition to T and P, but do not specify which ones.

This manuscript would benefit from adjustments to the language and sentence structure, to improve the readability of the final manuscript. Most of the minor comments below aim to remove ambiguity and improve the flow of the manuscript.

-- Specific comments

P1 L17-18, “However, the spreads in...”

P1 L21, “with the randomly selected 10,000 arbitrary subset samples” → with 10,000 randomly selected subset samples

P3 L2-4 “the subset covers... randomly sampled” → the subset covers more of the uncertainty in the temperature and precipitation changes projected by 36 CMIP5 GCMs, than other randomly sampled five-GCM subsets.

P3 L30, “how extent” → to what extent

P4 L3-11, this paragraph is difficult to follow due to language and sentence structure. Please rework this paragraph to improve readability.

P4 L31 what do you mean by ‘too dry’ grids? It is unclear whether grid cells or entire models are excluded. If it is entire models (with spatially averaged precipitation < 0.1 mm/day), these models should still be included in the study of the historical bias, and excluding them should only be applied to the study of the projected changes. Currently, it is unclear whether these models are excluded entirely.

P4 L32, then → the ; “mean precipitation” → specify which mean (spatial mean, spatial and time mean?)

P5 L1 and L3, “grid” is ambiguous, please replace with “grid cells” or “models” depending on what you mean.

P5 L7, “multi-precipitation products”; it is unclear here whether you combine all these products (e.g. by averaging them) into a single reference product. According to the caption of Figure 1, all 6 were used in the “obs” ensemble, but GPCC is your reference product. I suggest amending “multi-precipitation products” → “six different precipitation products”, and adding a line saying that all the precipitation biases were calculated using GPCC as the reference.

P5 L17, referred → reference

P5 L19, “we use the model bias to evaluate the quantity itself” → we also evaluate the magnitude of the model bias. Using both metrics enables... the spatial pattern and the bias magnitude.

P5 L28, also arbitrary → also from arbitrary...

P5 L30, “MJ2016 presented...representation.” This sentence is awkward and does not add anything. Remove or adjust → “MJ2016 presented a similar comparison between the original ISIMIP model data and 500 randomly selected five-model ensembles”.

P6 L9, “biases: the other” → “biases; the other”

P6 L24, large → larger. “As the result” → As a result

P6 L32, “four models in ISMIP. Thus...” → four models in ISMIP, indicating that the biases... are almost the same.

P7 L3, “include... observation” → include members showing a spatial pattern of low similarity to that of observations.

P7 L7 “would be related” → could be related

P7 L16 “has not much resembled the observation” → does not resemble that of the observations.

P7 L25 “change of “ → changes in

P7 L 26 “projected increments of the temperature”. Ambiguous. I suggest → average rate of temperature increase per year, calculated from the 20-year period for each model

P7 L30, Any suggestion of why this result is different from other studies? It would be good to include some mention of this in the discussion section.

P8 L2 “with totally covering” → and cover the full range...

P8 L5 “like at least the bias..” → for example whether we use the bias or the skill score.

P8 L7, “capture the full range less...” → capture less than 60% of the full range in all regions.

P8 L8, capture the wider → capture a wider range

P8 L8, “subsets, differing from...” → subsets, a result markedly different than for ΔT , where both CORDEX and ISIMIP have relatively large coverage.

P8 L9, shows the difficulty of → has difficulty capturing

P8 L12, → uncertainty range (maximum-minimum) is ...

P8 L16, Only in... → In Central Asia, the full range of.. remains below the 25th percentile... while the maximum-minimum range of .. adequately covers the IQR... Thus, the three models.. Central Asia underestimate the average tendency...

P8 L19 “despite being... differing from ISIMIP” This is unclear. Please rephrase.

P8 L24 “high coverage for the temperature..” How about precipitation? Please include the precipitation results for ISIMIP too.

P8 L 31 presents distribution → represents the distribution...

P9 L32, "They focused... models" Unclear, please rephrase. In particular, "grid" is ambiguous, I suggest using model (or grid cell) if applicable.

P10 L8, "which are less than nine members" this fragment does not add anything. Please remove or adjust to explain why nine members is relevant.

P10 L24, "high performed models" → high performance models.

P10 L25, "regarding to" → with regard to

P10 L29, "13 models in 25 all high performance models" → 13 high performance model (out of 25) are included...

P10 L32 "Therefore, although... possibility. " Unclear, please rephrase.

P11 L3, the subset → their model ensemble indicates...

P11 L7, the dataset with reasonable for... projections → a dataset with reasonable values for... and with enough coverage of the projection uncertainties.

P11 L8 with based → with values based

P11 L12, Remove "in the case". Also, can you suggest other relevant variables, given that you state temperature and precipitation are not enough?

P11 L19, similar to the above, can you include a suggestion of a variable to characterize "circulation" (e.g. geopotential height, sea level pressure...)?

P11 L23, similar to the above, simply stating "the other ones" is too vague. Make specific suggestions of potentially useful variables.

P12 L3, How about using both precipitation and temperature? Can you comment here on a combined evaluation and whether the results are likely to be different from the ones included in this study?

P12 L20, "with capturing" → while capturing

P12 L22, "the global common subset" → the global common (ISIMIP) subset ; "but large model ensemble is needed" → but performs better when a large number of models is used.

P12 L26, "need to be changed" to what? Specify, e.g. increased to seven, or nine models if possible....

P12 L31, specify which other variable might be relevant.

Figure 4: The red dot is missing for Central Asia. Is it also an overlap? If so, change the caption accordingly. For the 3 regions where the dots overlap, I suggest changing the size of one of the dots (e.g. the red dot slightly larger than the blue dot), so that the outside of the red dot is visible, this would make the plot easier to understand by showing the overlap, rather than having to read the caption.

