

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2022-111-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2022-111

Anonymous Referee #2

Referee comment on "Advancing precipitation prediction using a new-generation storm-resolving model framework – SIMA-MPAS (V1.0): a case study over the western United States" by Xingying Huang et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-111-RC2>, 2022

Summary

Huang et al. in "Advancing Precipitation Prediction Using a New Generation Storm-resolving Model Framework - SIMA-MPAS (V1.0) a Case Study over the Western United States" evaluates several versions of the SIMA-MPAS model (e.g., vertical resolution, horizontal resolution, and microphysics schemes) that has been recently implemented into a widely used Earth system model, CESM2. The authors evaluate the five-year, near-term historical model simulations, mostly, over the western United States (save for a global comparison to the "work horse" dycore in CESM2-CAM6, CAM-FV) and compare model performance to a traditional regional climate model and observationally based gridded products.

Overall, I think the paper fits within the scope of GMD and could be, given more work, a valuable contribution to the scientific community. The model advancements/developments (as described in lines 121-141), experimental designs, and data production aspects of this manuscript are clear and robust (save for questions about WUS interannual variability). However, the messaging in the manuscript was very choppy and appeared hastily put together. I have tried to provide constructive feedback to overcome this but highly encourage the entire authorship team to provide a more thorough internal edit during the next revision. As a result, I think there are several major revisions that need to happen prior to this paper being accepted.

Major and Minor Revisions

Line 12-13 – Delete "...for predictions to be useful..." You might also want to bring up the scale-awareness aspects of parameterizations that shape precipitation intensity, duration, and frequency.

Line 17 – Delete “For mean climatology” and add “...with reasonable energy and mass conservation on climatological timescales”

Line 19 – “We mainly investigate...” this was an awkward sentence transition. I think what you’re trying to say is that you wanted to prove comparable energy/mass balance performance between CAM-FV (workhorse dycore) and SIMA-MPAS (newly developed dycore) and that this coarse resolution performance should give confidence to the community that SIMA-MPAS can then be evaluated at finer resolution?

Line 21 – “Effectively” or “Efficiently”?

Line 22 – Change to “...60 km over the rest of the globe...”

Line 23 – Delete “precipitation details”

Line 24 – What are “temperature features”?

Line 26-27 – “We compared and evaluated ...” this whole sentence should be presented earlier, and more specificity is needed. What observations? What traditional model?

Line 29-30 – “The” or “This”?

Line 44-45 – This was another awkward sentence transition. I would delete “Given the recent development of Earth system model frameworks...” and start with “Advances in computer power have now enabled climate models to be run with non-hydrostatic dynamical cores at “storm-resolving” scales, on the order of a few kilometers...”

Line 51 – What are the differences between “structures” and “platforms”?

Line 54 – Delete “one of the leading Earth system models” (this point is proven by the next part of the sentence)

Line 57 – CESM has been used in many other regional climate modeling contexts too. Please thoroughly cite other uniform high-resolution and variable-resolution CESM-SE (spectral element core) studies performed over the years too.

Line 58 – Change “modeling” to “model”

Line 58 – Either use “more recently” or “over the past decade”

Line 64 – If the authors state “number of studies” I think more than one study should be cited.

Line 92-105 – Can the authors, in a paragraph or so, provide hypotheses on how any portions of the new modeling system where there is/isn't scale awareness, particularly when transitioning from ~120km to ~60km to ~3km, could impact regional precipitation statistics (e.g., intensity, duration, frequency, phase, extremes, etc.), with a particular focus over the WUS? Also, would there be any resolution dependent land surface related feedbacks to the atmosphere that might arise?

Line 101 – “Predictions”

Line 103-104 – Change “snowpack statistics features” to “precipitation and snowpack statistics”

Line 114 – Observed SST and sea ice?

Line 144 – The authors alternate between the use of “storm-resolving” and “convection-permitting”, why?

Line 145-147 - I think the authors should provide a justification (a few sentences) for why the WUS is chosen and not some other region of the world.

Line 148 – Change to “...when compared with a traditional regional climate model against best-available observations and observationally-based gridded products at similar...”

Line 149 – Delete “behavior”. Also “heavy” or “extreme”?

Set B, C, and D – I know model simulations can be quite costly, but given the large interannual variability of precipitation in the WUS, are five simulated years enough? Were 1999-2004 consistent with a near-term climatology of the region? If not, could the authors, at least, mention this somewhere as a potential constraint.

Line 163 – Change “western us” to “WUS”

Line 164 and 169 – Was there a reason that the authors did not want to assess summer/fall precipitation (e.g., convective storms, Monsoons, etc.), particularly over the Rocky Mountains? This may be addressed earlier on in the manuscript with an additional paragraph that describes hypotheses of model performance across resolution (or, simply, long-standing biases for particular seasonal precipitation, regardless of resolution).

Line 171-174 – I’d suggest the authors provide more context as to why the addition of graupel could be an important driver of WUS precipitation statistics and should be evaluated as a unique, standalone, and resource intensive simulation (like Set A-C)

Line 166 – Delete “reanalysis” (redundant)

Line 177 – Delete “(i.e., seconds)”

Line 178-179 – Could the authors provide a bit more justification for the physics timestep choice (or point to other studies)?

Line 180-181 – Which simulation cost this much? I imagine each simulation was different in cost. Maybe place this information in Table 1. Which high performance computing system? What was the simulation throughput per node? How large were the output data volumes? (I think this could be useful information to those folks interested in running these sorts of experiments on university clusters or ... in the future when the model configurations are publicly available)

Table 1 caption is a bit colloquial.

Figure 1 – I think the authors should be more specific in the caption, “SIMA-MPAS mesh

configuration for the 60-3km experiments”, to make the figure standalone.

Section 2.2 – “Other datasets” is a bit broad. How about “Observations and observationally-based gridded products used to evaluate model performance”

Line 199-200 – PRISM and Daymet are “observationally-based”. Also there are known issues with PRISM’s ability to represent extreme precipitation events (which might be important when comparing a smaller subset of years) - <https://doi.org/10.1175/JHM-D-15-0019.1> - “...In general, over the entire period, the gridded datasets performed reasonably well, with over 50% of median errors on individual days falling between –37% and 44% and water-year total errors within $\pm 10\%$ (Table 2). However, errors in individual storm events sometimes exceeded 50% for the median difference across all stations, and in some years, these underpredicted storms led to 20% error in water-year total median statewide snowfall (e.g., water-year 2008, Fig. 9)...” A path forward may be to use a recently released dataset that better accounts for extreme precipitation - <https://doi.org/10.1175/JHM-D-20-0212.1> (http://cirrus.ucsd.edu/~pierce/nonsplit_precip/)

Line 199-200 - Daymet is not usually used for snowpack evaluation in the WUS. From what I’ve gathered, Daymet treats snowpack as more of an “energy budget fixer” term in the Daymet dataset generation process (discontinuities can be seen between Dec 31 and Jan 1). UofA (<https://nsidc.org/data/nsidc-0719/versions/1>) and the UCLA (https://nsidc.org/data/WUS_UCLA_SR/versions/1) datasets are more widely trusted/used and I’d encourage the authors to consider one (or both) of them for this evaluation.

Line 196-209 – did the authors remap any of the datasets before comparison? Could the authors provide explicit details on the type of remapping and which “direction” was used (i.e., remap observations to models or models to observations or all datasets to a common resolution or ...)

Line 205 – Delete “as a”

Line 224-25 – This sentence is very confusing.

Figure 3 – is the caption correct (e.g., CESM2-CAM6)? Is CAM6 used as an umbrella term for both SIMA-MPAS and CAM-FV?

Line 238-248 – Can the authors discuss why the largest differences (e.g., ice water path and cloud fraction) between CAM-FV and SIMA-MPAS occur in the midlatitudes (e.g., WUS)?

Line 250-260 – to me the difference maps in Figure S1 highlight that CAM-FV and MPAS are dissimilar, particularly at high altitudes in the tropics and poles (they even have different signs in bias).

Line 265 – Delete “coastal ranges and” (redundant)

Line 273-275 – Why is “for about” used? What are spatial textures?

Figure 3 – what aren’t a) and b) on the same page?

Line 280 – Delete one of the “)”...

Line 286-287 – I’m not following the reasoning about the “smoother topography”. The innermost circle (representing the refinement patch of SIMA-MPAS) in Figure 1a appears to properly encapsulate the entirety of the WUS. Can you provide a difference plot of Figure 2 (and from an observed DEM) for easier comparison (place in Supplemental)? Also, average precipitation rate is low biased (compared with PRISM) throughout the Klamath, Great Basin, and Sierra Nevada too (which is well within the innermost domain of SIMA-MPAS). Could this also be an interannual variability issue with only 5-years of simulation (i.e., atmospheric rivers, etc. may not have made landfall as frequently or in the same location or at the same magnitude as what was observed)?

Line 284 – Again, what is “precipitation texture”?

Figure 5 – Change “frequency” to “intensity”. Also, I think it will be important to compare/contrast with, at least, one other observationally based gridded precipitation product (see earlier comment).

Line 320-332 – Do the authors have any idea as to why precipitation intensity (particularly in the extremes) is so much lower in SIMA-MPAS and so much higher in WRF? Providing some context for the readers could be useful (and, if possible, a path forward to fixing this).

Line 344-346 – Rather than simply stating MG3 “could be a better option”, could the authors spell out in more detail the physical intuition as to why this might be the case (over a few sentences)? Do the additional ice phase hydrometeors alter droplet

coalescence/droplet size and the intensity of precipitation, seeder-feeder effects in complex terrain can begin to arise, or ...?

Figure 6 – rather than changing the color of MG3, could the authors change the stippling of the line and keep the same line color? I do think interannual variability is likely shaping the California results (e.g., not enough atmospheric river landfalls as was observed).

Line 369 – What are “snowpack statistics features”?

Line 371-375 – This is a confusing set of sentences. Please revise.

Line 381 – Change “retrieving” (too colloquial and not sure what is meant)

Line 371-384 – this is a basic analysis of snowpack skill. Could the authors use a similar figure structure as Figure 6/7 to daily dSWE (mm/day) magnitudes across the WUS? This would be a proxy for snowfall rate while still being able to leverage the observationally-based gridded SWE products that, likely, don't provide it.

Line 408-410 – What is meant by this statement? Are they authors trying to setup that future climate change analyses would be justified based on these results?

Line 419 – Again, please proofread before submitting a manuscript, “For further investigation, we have investigated’...

Line 440-448 – This entire paragraph needs to be overhauled. In particular, the first and last sentences.

Line 500-502 – Not sure what is meant by this sentence. Please revise.

Line 505-510 – Not sure what is meant by this sentence. Please revise.

Line 454-514 – To provide more added value from this study, I think the authors need to compare some of the biases in SIMA-MPAS to previous studies using MPAS, VR-CESM, etc. (a paragraph or two) to contextualize model performance, especially if those studies also

used PRISM, Daymet, etc. The authors frequently use the term "better", "pretty good", "good", etc. and I think these qualitative statements need to be quantified and contrasted a bit more.