Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC2
https://doi.org/10.5194/gmd-2021-364-RC2, 2022

Comment on gmd-2021-364

Anonymous Referee #2

Referee comment on "Atmospheric river representation in the Energy Exascale Earth System Model (E3SM) version 1.0" by Sol Kim et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-364-RC2, 2022

In this manuscript, the authors compare historical simulations of global AR frequency in an Earth System Model (E3SM) to AR frequency in a reanalysis dataset (MERRA2). The authors highlight differences in model depictions of global AR frequency and provide some physical insights into these biases. The manuscript is detailed, and I appreciate the authors' attempt to put the E3SM biases into context. However, I feel that the manuscript could benefit from several major changes prior to publication. I base my recommendation on the general comments listed below.

General Comments:

The methodology section is rather sparse on details, especially those related to E3SM, MERRA2, and the AR detection algorithm. For E3SM and reanalysis, which output fields are obtained? Is the daily data instantaneous (if so, at what time?) or daily-averaged? For the algorithm, a few additional details (e.g., regarding geometric criteria) would be helpful.

Regarding the choice of AR detection algorithm, recent intercomparison studies have shown that climate model simulations of AR activity vary based on choice of AR detection algorithm. While the authors are justified in choosing the Guan and Waliser algorithm, I think the authors should acknowledge possible uncertainty in results owing to the choice of algorithm.

I agree with Referee #1 that the characterization of differences in AR frequency should be in a relative sense vs. an absolute sense. Given that ARs only occur for at most 15-20% of the time steps, these absolute differences are considerable. I agree that this section should be overhauled to reflect the relative differences.

Regarding the choice of domain, while I agree with Referee #1 that a regional domain would be useful for more targeted studies, I think that the global domain is appropriate for the scope of this paper, which I interpret as a first step in understanding depictions of AR frequency in E3SM.

While the authors provide a good initial overview of the biases seen in E3SM, the authors do not provide much insight into potential pathways toward improving the model. I am specifically thinking of the parameterization of convection, PBL turbulence, microphysics, etc., which could have an influence on the depiction of water vapor transport. At the very least, the authors should note that the model physics could play a role in depicting ARs and precipitation, even if this is left to future research.

Specific Comments:

Lines 40-54: The authors may consider adding a recent article, O'Brien et al. (2021), which analyzes changes in AR counts and size in a future climate within the CMIP5/6 models:

O'Brien, Travis Allen and Wehner, Michael F and Payne, Ashley E. and Shields, Christine A and Rutz, Jonathan J. and Leung, L. Ruby and Ralph, F. Martin and Marquardt Collow, Allison B. and Guan, Bin and Lora, Juan Manuel and et al., (2021) Increases in Future AR Count and Size: Overview of the ARTMIP Tier 2 CMIP5/6 Experiment. *JGR A.* https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021JD036013

Lines 78-95: Which fields are analyzed for the model and reanalysis datasets? At the bare minimum, the specific humidity and wind fields would be required to calculate IVT. Are there additional fields downloaded?

Lines 96-106: Given the importance of the choice of AR detection algorithm to this study, I suggest that the authors provide some additional details about the detection algorithm (even though the details have already been published elsewhere).

Lines 109-110: How are the ensemble mean AR frequencies calculated? Are ARs detected for each ensemble member and the frequencies averaged? Or are the ARs detected using the ensemble-average IVT?

Figure 1: While the absolute differences are important, relative differences are likely more interesting given the relative infrequency of ARs, even over the midlatitude storm track.

Figure 2: This is a very interesting plot, but visually, the data points are difficult to discern. There is a lot of "empty" space in these diagrams. Would it be possible to either a) make the data points larger or b) only show a fraction of each diagram? Zooming in somehow would be very helpful for the reader to help discern the slight differences in correlations between ensemble members.

Lines 165-168: Do features like the Indian Monsoon, large-scale tropical convection, or TCs get detected as ARs?

Lines 171-173: Are the feature-averaged values (namely, feature-averaged IVT) weighted by latitude?

Line 178: Is the feature's centroid based on area alone, or is there a weighting based on IVT intensity?

Lines 184-185: I don't agree with this statement. While this may be the case here, tropical ARs could have weaker moisture transport due to a lack of strong winds aloft. I suggest the authors clarify this statement.

Figure 7: Although the details are provided in the figure caption, I suggest adding labels to these panels for quicker/easier interpretation.

Lines 322-323: Are these individual AMIP and fully-coupled simulations or an ensemble average of each?

Figures 8-10: As with Figure 7, it would be helpful if the panels were labeled. Also, is AMIP subtracted from fully-coupled, or vice-versa?

Technical Corrections:

Line 51: "but only *a* few"

Lines 54, 60, 78: Parentheses around citation

Line 87: Spell out the AMIP abbreviation

Line 179: Use “fewer” rather than “less” here

Line 321: Is this “southern Africa”, as opposed to the country?

Line 344: Use “occur” rather than "occurs" here