Review of egusphere-2023-493

Summary:
This is an interesting and generally well-written study about atmospheric predictability. The authors investigate the predictability limits that arise due to uncertainty in the diabatic heating of MJO convection. The experiment is based on the IFS model and uses a “perfect model” approach. The results suggest that the predictability limits (as defined by error saturation) for zonal wavenumbers 0-3 in the tropics are >60 days, 30-40 days for wavenumbers 4-10, and 20-30 days wavenumbers 11-21. The error, which is initially confined to the tropics, contaminates the extratropics via the mid-latitude storm tracks at around day 9-10. There is some discussion about Rossby wave source and downward propagation of error in the stratosphere, but it is less clear how those results fit into the narrative of the manuscript.

Recommendation:
Major revision

Paper Strengths:
The experiment set up is clean and so are the results, especially the timeseries plots of error growth. A noteworthy result is that the large scales in the tropics have longer predictability than often assumed, in other words there is untapped predictability potential in the tropics.

Major Comments:

1. While it is true that only the model physics over the Indian Ocean are perturbed, the effect of chaos seeding (Ancell et al. 2018) quickly spreads the error over the whole globe in non-physical ways. This likely means that the results are indistinguishable from results that would have been obtained if SPPT perturbations were added to the entire tropical belt or to another location that is convectively active, such as the Amazon. I suggest testing this out.
2. It would be interesting to see how different the result would be when “butterfly seeding” were used, i.e., tiny initial perturbations of the initial conditions everywhere on the globe, as in Judt (2018) or Zhang et al (2019). I recommend running an additional ensemble with this kind of perturbation and comparing the results with the ones obtained so far (this additional experiment would also be more in line with *intrinsic predictability*, which usually addresses predictability limits arising due to miniscule initial condition uncertainty).
3. Is it necessary to discuss the Nov and Jan initialization experiments separately? In my opinion no, as the differences between Nov and Jan events are not large enough to warrant the extra work for the reader to keep track of two sets of results. I therefore recommend combining all experiments into one “grand experiment”. This shouldn’t affect the conclusions.
4. It looks like Fig. 3 is not referenced in the text. Furthermore, I am not sure why the “Rossby wave source” is analyzed at all. I suggest removing this analysis or better motivating it.
5. Section 3.4 seems to be lacking a conclusion, or at least I’m left with this impression.

Minor Comments:

1. L.45: How does the presence of baroclinic instability limit the predictability of MJO teleconnections? Through error growth associated with baroclinic instability?
2. Figs. 1-3 and 6-9 imply that the runs are 30 days long, while the other figures show the entire 60 day time period. Why are only the first 30 days shown in the Hovmöller plots? Maybe nothing interesting happens after 30 days, but then it should be indicated somewhere so the reader doesn't end up confused whether or not the experiments are 30 or 60 days long.
3. L.102: Just curious, why are you not using ERA5 to initialize the ensembles?
4. L. 155-177: I don't think the description of the figures is necessarily wrong, but I do see a lot of noise in the Hovmöllers and not so much of the described
5. Fig. 1 (and Fig. 2): The evolution of the standard deviation in panels (d) shows very little propagation with the maximum being anchored between 70 and 100 deg E, unlike the heating amplitude . Is this because of the continuous perturbation in this region?
6. L. 188: "In the Indian Ocean the two fields are comparable." I disagree, there seems to be more red in Fig. 1d than in Fig. A1 over the Indian Ocean.
7. L. 211: Where does the 0.5 threshold come from? It seems arbitrary.
8. L. 298: "small scales" is relative, wavenumber 21 is "large scale" from the view of a synoptic/mesoscale meteorologist.

Editorial Suggestions:

1. L. 186: two "the"

References:

Ancell, B. C., A. Bogusz, M. J. Lauridsen, and C. J. Nauert, 2018: Seeding Chaos: The Dire Consequences of Numerical Noise in NWP Perturbation Experiments. Bull. Amer. Meteor. Soc., 99(3), 615-628, https://doi.org/10.1175/BAMS-D-17-0129.1.

Judt, F., 2018: Insights into Atmospheric Predictability through Global Convection-Permitting Model Simulations. J. Atmos. Sci., 75, 1477–1497, https://doi.org/10.1175/JAS-D-17-0343.1.

Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S. Lin, J. Chen, and K. Emanuel, 2019: What Is the Predictability Limit of Midlatitude Weather?. J. Atmos. Sci., 76, 1077–1091, https://doi.org/10.1175/JAS-D-18-0269.1.