Review of "Parameterization adaption needed to unlock the benefits of increased resolution for the ITCZ in ICON"

Manuscript authors: Kroll et al

Summary

The authors use simulations at different resolutions and with the gustiness factor perturbed in the ICON model to examine mean precip biases. They argue that the double ITCZ bias persists at all resolutions but that increasing the minimum wind speed for the evaporation improves this, which however does not fix the larger underlying humidity biases across the tropics and subtropics.

There are some results of interest here, but I had a hard time following this manuscript and left it not especially convinced. In part this stems from issues with the writing, in part from the experimental design, and in part from the lack of uncertainty quantification.

Addressing all this adequately would take a serious amount of work, and as such I recommend rejection for eventual resubmission by the authors.

Major comments

Experimental design

It would be much cleaner if the highest resolution run didn't also have the convection and gravity wave schemes disabled as well, c.f. L120-122. As it stands, going from the 40-km model to the 5-km model, you're both increasing resolution **and** changing the model formulation. And even a 5-km grid is surely not fine enough to resolve the large number of convective updrafts that are smaller than 5x5km.

A relevant paper here is *Clark et al.* (2024), and references therein. In the GFDL AM4 model in an aquaplanet context, at high resolution the model behavior still changes dramatically depending on whether the convective parameterization is enabled or not.

This is especially concerning given that some fields such as the humidity and temperature biases are non-monotonic in resolution, with the change occurring going from the 40-km version in which the deep convective parameterization are activated to the 5-km version in which they are disabled.

Uncertainty quantification

C.f. L250-254, The 4, 5, and 6 experiments precip RMSE values are all within 0.03 of each other. Is that even a statistically significant separation? I worry about sampling uncertainty given the short durations of

the runs. I have the same concern about other results; apart from Fig. A1 there is very little discussion of uncertainty quantification and its implications for interpreting the results. It seems plausible that you're overinterpreting differences across simulations that aren't statistically well separated.

Experiment names

The results would be much easier to follow if the experiments had more descriptive names. So, instead of "R2B4," call it for example "160-km." The "PTB-X" simulation names are similarly unintuitive.

Unclear arguments regarding model tuning

I find the discussion of tuning peppered throughout the introduction to be frustrating. It feels speculative and almost conspiratorial, seeming to imply that the modeling groups are somehow, in the case of focusing on global-mean TOA radiative fluxes rather than regional circulation fidelity, doing something obviously wrong. In the discussion of tuning high-resolution models, L65-67, "misconception that all relevant processes are now resolved" is not justifiable, nor is it appropriate in tone. Model developers are fully aware of the physical scales of the various processes involved and how those compare to the scales resolved by their model.

The paper should incorporate *Zhao et al.* (2018a,b), who discuss how a tuning strategy targeting TOA fluxes was used to improve ITCZ simulation.

English writing

My impression is that English is not the lead author's native language for writing. There are quite a lot of sentences where the grammar and/or word choice are difficult to follow. In aggregate, these make for a somewhat jarring reading. At least one of the coauthors is a native English speaker, and so I know it is within the authors' collective ability to, in the revision, significantly tighten up the English writing. Here is a non-exhaustive list of sentences that I struggled with:

- L36-37
- L49
- L111-114
- L132-134
- L134-135
- L138-141
- L150-151
- L235-236
- L237-238 "vicious cycle" over the top
- L238-244 (break up this sentence)

Separate from the English usage, there were far too many typos. Please carefully proofread the revision carefully as a final step before submitting.

Line-by-line comments

L13 "this could endanger the representation of the global circulation, energetic balance and teleconnections" confusing, due to the "could." Does it degrade these fields in your simulations or not?

L15 what does "non-discardable" mean?

L25 "bias has been central to the precipitation bias discussions" this reads funny to me; consider rephrasing

L27-29 Correct and you should cite one or more papers that document these transient double ITCZ states, e.g. Magnusdottir and Wan 2008, `https://journals.ametsoc.org/view/journals/atsc/65/7/2007jas2518.1.xml`. And this preprint is particularly relevant: `https://essopenarchive.org/doi/full/10.22541/essoar.174017095.57302520`

L42 Philander et al emphasize ocean-atmosphere coupling and continental geometry, making it an odd choice to cite regarding this claim about net energy imbalance; *Frierson et al.* (2013) would be more appropriate.

L50-57 I don't find this discussion of the model tuning especially compelling, in large part because I'm struggling to follow it. Can you make your argument more precise and clear?

L75-77 This sentence is meaningless to the reader, like me, who doesn't know what ICON XPP and ICON Sapphire are. I think you can omit this entirely, or if you want to keep it consider moving to the methods section or revising to provide more context

L88 I would omit "the fuel for the hydrological cycle"; unneeded and too imprecise

L94 "their implementation should receive more attention" this feels like too much of an editorializing statement to me in the context. I don't really know what 'minor-looking treatments' are beyond your summary having not read the Kawai et al paper, but based on your summary it's not obvious to me why they should indeed receive more attention.

L111 I would omit the footnote; just include it in the parenthetical

L118 calling 5 km convection "resolving" is a stretch…very few convective updrafts span 5x5 square km. "convection permitting" is a widely used and I think more appropriate choice.

L165 Doesn't ERA5 directly output specific humidity?

L214 fix the citations

L215 A lot of typos through the end of this paragraph; feels sloppy.

L229 What feedback loop?

L229-232 Is this proposed feedback your idea? If yes, it feels rather speculative. If not, it needs citations.

Fig. 6 What does ”normalized” probability density function mean?

Fig. 7 Is OAFlux ultimately a better product than ERA5 for the surface LH fluxes? If so, then why show the biases of the simulations against both? Why not just use OAFlux in the context of this whole discussion? Perhaps I missed something here.

L342-344 I don't understand this. Why do the signs of the respective biases lead to this inference about "symptoms" vs. "root cause"?

Fig. 8 symbols for control run are too faint

References

Clark, J. P., P. Lin, and S. A. Hill (2024), ITCZ Response to Disabling Parameterized Convection in Global Fixed-SST GFDL-AM4 Aquaplanet Simulations at 50 and 6 km Resolutions, *Journal of Advances in Modeling Earth Systems*, *16*(6), e2023MS003,968, doi:10.1029/2023MS003968.

Frierson, D. M. W., Y.-T. Hwang, N. S. Fučkar, R. Seager, S. M. Kang, A. Donohoe, E. A. Maroon, X. Liu, and D. S. Battisti (2013), Contribution of ocean overturning circulation to tropical rainfall peak in the Northern Hemisphere, *Nature Geoscience*, *6*(11), 940–944, doi:10.1038/ngeo1987.

Zhao, M., J.-C. Golaz, I. M. Held, H. Guo, V. Balaji, R. Benson, J.-H. Chen, X. Chen, L. J. Donner, J. P. Dunne, K. Dunne, J. Durachta, S.-M. Fan, S. M. Freidenreich, S. T. Garner, P. Ginoux, L. M. Harris, L. W. Horowitz, J. P. Krasting, A. R. Langenhorst, Z. Liang, P. Lin, S.-J. Lin, S. L. Malyshev, E. Mason, P. C. D. Milly, Y. Ming, V. Naik, F. Paulot, D. Paynter, P. Phillipps, A. Radhakrishnan, V. Ramaswamy, T. Robinson, D. Schwarzkopf, C. J. Seman, E. Shevliakova, Z. Shen, H. Shin, L. G. Silvers, J. R. Wilson, M. Winton, A. T. Wittenberg, B. Wyman, and B. Xiang (2018a), The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 1. Simulation Characteristics With Prescribed SSTs, *Journal of Advances in Modeling Earth Systems*, *10*(3), 691–734, doi:10.1002/2017MS001208.

Zhao, M., J.-C. Golaz, I. M. Held, H. Guo, V. Balaji, R. Benson, J.-H. Chen, X. Chen, L. J. Donner, J. P. Dunne, K. Dunne, J. Durachta, S.-M. Fan, S. M. Freidenreich, S. T. Garner, P. Ginoux, L. M. Harris, L. W. Horowitz, J. P. Krasting, A. R. Langenhorst, Z. Liang, P. Lin, S.-J. Lin, S. L. Malyshev, E. Mason, P. C. D. Milly, Y. Ming, V. Naik, F. Paulot, D. Paynter, P. Phillipps, A. Radhakrishnan, V. Ramaswamy, T. Robinson, D. Schwarzkopf, C. J. Seman, E. Shevliakova, Z. Shen, H. Shin, L. G. Silvers, J. R. Wilson, M. Winton, A. T. Wittenberg, B. Wyman, and B. Xiang (2018b), The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 2. Model Description, Sensitivity Studies, and Tuning Strategies, *Journal of Advances in Modeling Earth Systems*, *10*(3), 735–769, doi:10.1002/2017MS001209.