Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-83-RC2, 2019

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Machine dependence as a source of uncertainty in climate models: The HadGEM3-GC3.1 CMIP Preindustrial simulation" *by* Maria-Vittoria Guarino et al.

Anonymous Referee #2

Received and published: 13 June 2019

This paper examines the difference in two control simulations with a state-of-the-art climate model using nominally the same code but on different machines (with different compilers, chips and libraries). The rate at which the simulations diverge from each other with the same initial coniditions is consistent with the expected divergence associated with a sensitive dependence to initial conditions, but over the longer term, there are no detected differences in the climatology.

Historically, it has been the case that some climate model simulations have had a climatolgogy that varied as a function of machine platform, though this has rarely been discussed in the literature (though see a very recent example here: https://www.geosci-

Printer-friendly version

Discussion paper

model-dev-discuss.net/gmd-2019-91/). Generally this is an indication of bugs in the code that manifest themselves differently and systematically so, on a different architecture, and thus would have been problematic had this been discovered here.

However, the results presented here demonstrate that the climatology of the two simulations is the same - given a long enough averaging period the simulations are indistinguishable. This is a good result, however, it is not the conclusion that the authors come to.

Part of the problem I think is conceptual. In an ideal simulation of a chaotic system on two different architectures with identical initial conditions, the differences in the machines will manifest themselves as machine-level rounding differences spread throughout the calculation (as shown in section 2). Since the GCM is a simulation with sensitive dependendce on inititial conditions, this will place the simulations on slightly adjacent trajectories which would then be expected to diverge with whatever Lyapunov exponent is relevant. Subsequent time-steps will simply repeat the exercise (i.e. perturbing the initial coniditons for the next time step by machine precision). I do not see how this will produce anything fundamentally different from a standard initial condition ensemble. Therefore the question to be asked of the two simulations discussed here is whether the simulations are distinguishable from an IC ensemble on a single machine, not whether they diverge at all. I note that this is the standard used in Hong et al, 2013 in a slightly different context. I am a little puzzled that the authors are not seeing this (especially given the statement on p11. line 24). Given this, the analysis in sections 3 and 4 are of little interest.

Thus I do not think the current paper is publishable. A re-conceptulaised analysis of these runs and this issue might be acceptable, but that would be quite a different paper.

Minor points:

p2 line 31. The climate modeling community spends an enormous amount of effort to ensure bitwise reproducibility for testing and development purposes. The point made

in the Liu et al 2015 paper is whether more effort should be spent to ensure it in a broader context (i.e. over years and across platforms), not whether it's worth doing at all.

p2 line 35. "the uncertainty attributable to machine-dependent processes" - I disagree. The authors have not attributed this at all.

p8. line 34. No. It suggests only that there are adjustment times longer than 100 years in the climate system.

p11. line 10. The authors are merely assuming that the differences between the two runs are due to 'machine dependence' - this is begging the question.

p11. line 16. There is no 'machine-induced bias' in these runs (bias is a difference in the long-term means).

p11. line 30-34. I am all for greater ensemble sizes (which actually, many groups are doing - i.e. the CESM Large ensemble, or the CCCM large ensemble), but this is again related to standard IC-related dependence, not machine dependency.

p12. line 8. "suspicious"??? This is a very odd term to apply.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-83, 2019.

Printer-friendly version

Discussion paper