

Interactive comment on “Single precision arithmetic in ECHAM radiation reduces runtime and energy consumption” by Alessandro Cotronei and Thomas Slawig

Anonymous Referee #2

Received and published: 3 March 2020

The authors comprehensively describe the conversion of the radiation code of ECHAM from double-precision to single-precision (strictly speaking, a model version that allows both double and single-precision). They demonstrate that the model performs well both when compared with the double-precision version and a number of observational datasets. They also assess the computational speed-up and the energy savings.

The paper builds on earlier work on the ECMWF model and COSMO and so is not entirely novel. Nevertheless, there are considerably more technical details than earlier papers and so I recommend it for publication as a useful guide for those developing reduced-precision versions of their own models. This recommendation is subject to

C1

the minor revisions below.

1 Minor revisions

- Section 3.3: After reading this section several times, I still don't understand the conversion process. For example:
 - line 167: “If the model output was acceptable” — how do you define “acceptable”? And what do you do if the model output is not acceptable? Do you then delete the sp version?
 - line 167: “low_dp, as well as the interface were redundant” — even if the model output when using low_sp is acceptable, you still need an interface “on top of” low_sp that allows double-precision arguments. This is the purpose of low_dp. If you delete low_dp, how does high_dp call low_sp?

It could be that the authors won't understand my questions because my thinking is so wrong. In any case, I didn't understand it and I recommend that the authors rewrite this section so the procedure is clearer, perhaps including some diagrams.

- line 182: “namely -real-size 32”: this depends on the compiler. I'm not even sure GNU Fortran has an option to set the default REAL precision to 4 bytes, as this is already the language standard (as far as I'm aware). I don't think you need this paragraph at all — you can simply say that you added type declarations to all REAL variables and literals so that the type was always explicit. This is good programming practice anyway.
- lines 297 - 300: I didn't understand this paragraph. For example, “we also took a look at the minimum and maximum over all grid points” — minimum and

C2

maximum error, or minimum and maximum field values? I'm assuming the latter. If so, why does a difference in minimum and maximum indicate a bias? If single-precision has both a larger maximum and smaller minimum than double-precision, the mean could still be zero (meaning zero bias). I recommend either rewriting this paragraph or just deleting it.

- Section 6.1.2: The equation only computes an annual mean if #months in time span = 12. In fact the period is 30 years so I think you mean "temporal mean" not "annual mean". The caption of Figure 6 even uses that name.
- Figures 6, 7, 8: This could simply be GMD's formatting but I can't read the color-bars in these figures. If the authors deliberately chose this size for the Figures, please enlarge the colorbars.
- line 359: so all of the shown results are for "block"? Please clarify this.
- The terms "performance gain", "runtime reduction", "speed-up" and "acceleration" are used interchangeably throughout the manuscript (mainly the first three) but it's not clear what they actually mean. If x and y are the wall-clock times for single and double, respectively, is the performance gain (or whatever) $1 - x/y$ or y/x ? I recommend using the phrase "runtime reduction", meaning $1 - x/y$, as much as possible, as this is what others like Váňa et al. use. "Speed-up" sounds like y/x to me i.e. if $x = 5$ seconds and $y = 10$ seconds, the speed-up is 2 because single is twice as fast.

2 Typographical comments

- line 56, typo: double-precision → double-precision

C3

- line 142, language: not efficient → inefficient
- line 351, typo: below in in Figure → below in Figure
- Figures 9, 10, 11, typo: "performace gain" → "performance gain"
- line 373, language: A cause that no higher speed-up was achieved is that → We could not achieve a significant speed-up in some cases because
- line 391, typo: two version → two versions
- line 393, typo: low and resolution → low resolution

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-3>, 2020.

C4