

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

Alessandro Cotronei and Thomas Slawig

ts@informatik.uni-kiel.de

Received and published: 17 April 2020

We would like to thank the reviewer for the work and his/her comments. We have the following remarks corresponding to the points stated in section 1 of the interactive comments:

Reviewer's comment: 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

C1

version?

Answer: Here we changed the text to the following: We then tested if the model with the sp version of the subroutine/function compiles, produces no runtime errors, and if its difference to the dp version was in an “acceptable” range. Of course, the latter is a soft criterion, since a bit-identical result cannot be expected. Our criteria are explained below in Subsection 6.1. If the sp output was not acceptable in this sense, we marked the corresponding code part as to be treated separately, as described below in Sections 4 and 5.

Reviewer's comment: – 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.

Answer: This was formulated in a misleading way in our first version. The two versions low_sp and low_dp can be deleted only at the end, when a new “wp” version of “low” is introduced, wp being a variable that could be set to “sp” or “dp”. Then, also the interface becomes redundant.

Reviewer's comment: 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.

Answer: Paragraph was omitted.

C2

Reviewer's comment: 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 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 doubleprecision, the mean could still be zero (meaning zero bias). I recommend either rewriting this paragraph or just deleting it.

Answer: We removed the paragraph.

Reviewer's comment: 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.

Answer: We added an explaining sentence after the formula.

Reviewer's comment: Figures 6, 7, 8: This could simply be GMD's formatting but I can't read the colorbars in these figures. If the authors deliberately chose this size for the Figures, please enlarge the colorbars.

Answer: We enlarged the whole figures to make all detailed better visible.

Reviewer's comment: line 359: so all of the shown results are for "block"? Please clarify this.

Answer: We added this at (new) line 360 (old line 359) and also before in (new) line 328.

Reviewer's comment: 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 Vána et al. use. "Speed-up" sounds like $\tilde{E}G y/x$ to

C3

me i.e. if $x = 5$ seconds and $y = 10$ seconds, the speed-up is 2 because single is twice as fast.

Answer: We followed the suggestion to use "(relative) runtime reduction" all the time and defined the term as $1-x/y = (dp-sp)/dp$, at the beginning of subsection 6.2.

Answer to reviewer's section 2: All mentioned typos were corrected.

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

C4