

***Interactive comment on* “The ESCAPE project: Energy-efficient Scalable Algorithms for Weather Prediction at Exascale” by Andreas Müller et al.**

Anonymous Referee #2

Received and published: 22 March 2019

The paper presents a review of the work done in the Escape project. This reviewer is familiar with some of this work and reporting the work done as part of the project is certainly of interest to the community. Unfortunately, the paper is not well written, it is full of mistakes, informal language and confusing or unclear explanations. I have read and documented changes as far as page 17, but this has taken a long time as the paper has not been properly proof read before submission. Referee #1 calls for a substantial revision, and a possible change of paper type, therefore there doesn't seem much point in fully detailing necessary changes beyond this point. A shorter, more focused article as a review and perspective paper would improve the readability and is probably more appropriate for the content.

I include the detailed points below, which need to be addressed.

[Printer-friendly version](#)

[Discussion paper](#)



Page 2 line 7 Weather prediction (models?)

Line 20, sentence reads as if heavy precipitation patterns could lead to tropical cyclones, modify

Line 21 being satisfying “being”

Intro 1st paragraph is rather clumsy, there are plenty of reasons improved forecasts in general would have economic and societal benefit besides heavy precipitation. The need to improve resolution is given as the main motivation for improved forecasts but then Climate is thrown into the following sentence. Improved resolution versus complexity for improved climate is a matter of debate. There is no mention of current resolutions for the reader to compare 1km. What does global resolution range mean? The last sentence is also confusingly written. This paragraph needs re-writing with proper thought on what is the motivation for improving resolution of weather and climate simulations. There are plenty of justifications.

Line 27 “guarantee the continued efficiency” is probably a bit strong. “Enable efficient implementations of “ or similar is probably more realistic.

Line 33 “The authors there” is rather informal language.

Page 3. Line 7 Citation needed to back authors claim that ECMWF is world leading in terms of track forecast. It would be quite odd for a higher resolution forecast to “compromise the accuracy” of a lower resolution, especially as this is the motivation for escape.

The paragraph on supercomputers is also rather clumsily written and there should be a reference. An obvious one would be “Crossing the Chasm : how to develop weather and climate models for next generation computers?”. Some of the authors of this paper are also authors of the chasm paper. Other references would also be beneficial.

Section 2, line 18, The use of “at once” suggests either “immediately” or “at the same time” but “too big” implies neither. This sentence needs re-writing.

Page 4 The text in figure 1 is far too small to read.

Figure 2a “halo’s” has no apostrophe.

Page 5 figure 3 What does the vertical line denote? Line 5 what does “released dwarf” mean?

Page 7 line 5 “use 3D var like the “ is too informal language, use for example or e.g.
Page 7 line 23 “need” -> “needs”

Page 8 Figure 8. The figures are not very well explained. A careful reading of a technical appendix is required and even then it is not clear how they illustrate the point that overlapping the comms of data transfer is necessary. The main constraint for NWP is time to completion, it is not obvious that the assumptions (necessarily) made to express the scaling in terms of energy are sufficiently universal to infer the point the authors wish to make. Whilst using less energy reduces the cost, the run-time constraint is the primary motivation for NWP. This section needs to be re-worked to describe the performance model and explain how any figures used enable the authors draw the conclusions that they do.

Page 9 section 3 line 23 Last sentence. What does this mean?

Page 10, refers to figure 5 which is on page 11. How are the profiles produced? What machine are they produced from, especially the node count etc would change the profile? The “Anticipated” future are these profiled from data or some performance model? The 1.25km run may well have a different profile on a different machine?

Page 11 refers to figure 7 which appears on page 13, the text “like on the left” is rather informal and inappropriate for a scientific paper.

Page 12 refers to figure 8 which appears on page 14. In the figure, there is no open diamond referred to in the legend and the caption, the open rectangle referred to in the caption doesn’t appear but a dash or line does which isn’t referred to in the caption.

[Printer-friendly version](#)[Discussion paper](#)

Pages 12 and 13 there is a complicated discussion of code changes. This would be illuminated by some code fragments as examples.

Figure 10. Page 16 The data points are connected with a line (something spreadsheet application does readily), however, the horizontal axis is Number of GPUs, which is a discrete variable, so a line graph is wrong. Whilst it is not unusual to see such a plot, the authors are not predicting the speed up on 16.5 GPUs so why the line? The plot should be re-drawn appropriately.

Page 17 lines 6-15 the paragraph discusses the pack-unpack operation. The sentence sender and receiver share their memory layout as they may differ is confusing. How can they share a memory layout if they are different? What was the change that made the performance improvement? Were the pack and unpack scanning memory unnecessarily?

The final paragraph on page 17 is a discussion of implementing some of the GPUs on CPUs. Again some code fragments here would be helpful. If the GPU optimisations are too intensive, how can they be used elsewhere for other architectures? The stated goal of Escape is to re-assemble the models from optimised dwarfs. How is this managed? If there are conflicting optimisations how are they resolved? Is single source code possible?

Figure 11, the “barrier” optimisation is not clearly explained. Again, line plots of non-continuous variables should be changed.

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-304>, 2019.

[Printer-friendly version](#)[Discussion paper](#)