Geoscientific
Model Development
Discussions

Interactive comment on "Optimization of the WRFV3.7 adjoint model" *by* Qiang Cheng et al.

Anonymous Referee #1

Received and published: 1 January 2019

This paper proposes code optimizations and strategies to improve the performance of adjoint models in WRF. The code has been already merged as part of WRFPLUS.

1 General comments:

- Overall, from a computational perspective, I see little evidence or data supporting your conclusions. I would suggest to add more information about your analysis and verification methods. (Please see my Specific Comments).

- You mention that your code is already part of WRFPLUS (dyn_em). Can you please add any link or URL for a pull request or discussion of the code? This is valuable information for future work and other scientists.

C1

- English: Can you please make a general review, readability could be improved and some phrases and expressions are hard to understand.

2 Specific Comments:

- P2, l16: "Apparently, different ways have different costs in runtime or memory." Can you please be more specific (examples, including data about impact on memory usage).

- P2, l21: You make a leap from WRF to WRFPLUS that can make it confusing for readers that are not familiar with the version (package) differences. Can you please add some text in here (1/2 lines) talking about their differences thus, increasing the readability?

- P3, l5: "This implementation requires comparatively less memory consumption but much more computational cost, although there are always productive ways available for reducing the latter in practice." Can you please be more specific on about 'less memory' and 'more computational cost'? Can you please give an example of other ways of resources usage optimization?

- P3, l31: "As a typical strategy of the reverse accumulations..." Can you please add a reference in here?

- P6, l25: "allocating/ deallocating them outside of the running cycles of the procedures." Can you please specify the cost of this? What it is the benefit of this for the overall run (intuitively this will improve performance on your part of the code but move the cost somewhere else).

- P7, l23: "Through careful IO analysis". During all the paper you did not indicate any technique (e.g. tracing, or data dumping) that lead you to all these

conclusions more than reading the code and finding IO calls. Can you please elaborate this more and provide more insights/evidence of this analysis and its conclusions?

- P8, l7: "on a cluster system with 250 nodes, each of which has 20 processors" Can you please be more specific (node type, processor type, memory, network interconnection...)?

C3