Review on article gmd-2023-26 with title "Optimization and sophistication of the super-droplet method for ultrahigh resolution cloud simulations"

This paper presents a new implementation of the superdroplet method (SDM) to reach centimeter-to-meter scale resolutions in cloud macrophyiscs simulations. The authors discuss general aspects of how to represent microphyiscal processes in the SDM framework and the corresponding algorithms. In particular, the authors discuss the initialization, the SD movement, and the use of mixed precision in some parts of the implementation. Besides this general aspects, the authors also present the particular implementation, validation and scaling analysis in the supercomputer Fugaku, in particular the use of cache. The test case is shallow cumuli cases, and the new SDM implementation is compared with bin and bulk methods. The new SDM model is comparable and even faster that these methods, and it shows a weak-scaling of 98% up to 36864 nodes, which demonstrates the potential of SDM to advance research and simulation of microphysics processes and its coupling to cloud macrophysics.

I think the paper is very interesting. This work addresses the important topic of what are the current and near-future capabilities to study the interactions between cloud microphysics and cloud macrophysics, and thereby advance the understanding of the role of clouds in climate and weather. The discussion is thoroughly presented, and I think this is a good contribution to the community and this journal. I only have a few minor comments that the authors could consider before publication:

1. Maybe the only general comment is whether the authors could reduce the length of the manuscript, or consider a more clear separation between the discussion of general aspects of the SDM and the particularities of the implementation in Fugaku, as I tried to do in the first paragraph of this letter. I wonder if section 6 could be shorten. As said, this is a minor point and I realize it is difficult, but somehow I think the reader might be better guided through the paper to concentrate on the aspects that might be more interesting for her or for him.

   I found the conclusions, however, very well and they helped me to end the paper with a good idea of the various parts.

2. Title, abstract, line 1, and through the paper, the authors use the term *"ultrahigh-resolution"* and I was wondering if they could substitute that term by another one that is more informative. It seems that, in line 9, the authors refer explicitly to centimeter to meter scale, so why not say "centimeter-to-meter scale resolution" or "submeter resolution"?

   Otherwise, what would come after "ultrahigh resolution" when we reach the following step towards higher resolution?

3. Title, abstract, line 5, and through the paper, the authors use the term *"sophisticated"*. I was not sure what it means. Does it mean that it considers more physical processes, or a better model of them, or does it refer to the technical implementation from a computer science point of view?

4. abstract, line 4, the authors say *"does not make any assumption for the droplet size distribution"*. Since some aspects of the dynamics within the superdroplet or the interaction between superdroplets are still modeled, as discusses in section 2 and 3, I wonder if it might be better to say "makes less assumptions about the droplet size distribution and it is more physically sounded", or something similar.

5. Abstract, line 18: instead of *"perfect weak scaling"*, it might be stronger and clearer to say 98% weak scaling up to the corresponding number of nodes or cores, as it is done in the conclusions.

6. Introduction, line 35, the authors refer to Schulz and Mellado, 2018, but the reference Schulz and Mellado [2019] might be stronger for their case. Schulz and Mellado [2019] studies one micro-physical effect, namely, sedimentation, and shows that sedimentation is more important than previously thought, which strongly supports the efforts presented in this SDM manuscript to better represent the DSD.

7. Introduction, line 34 say *"which sets the eddy viscosity constant"* referring to DNS. I wonder if this sentence is needed. This seems to suggest that DNS is one type of LES, which is not the way DNS is used in turbulence research, where the concept originates from, since there is no eddy viscosity in DNS [Orszag and Patterson, 1972, Moin and Mahesh, 1998, Pope, 2000, Mellado et al., 2018]. DNS rescales the original in terms of size or in terms of physical properties to study Reynolds number effects, and remains accurate in the smallest resolved scales, which LES does not.

8. In several places through the manuscript, the authors refer to particle-in-cell (PIC) methods. What are the differences between the superdroplet method (SDM) and particle-in-cell methods?

9. Section 2, line 118, the authors write *"the anelastic equations assume horizontally uniform mean fields and are not appropriate for computing wider domains"*. I wonder if this sentence is needed. The mean fields need not be the reference fields that are used in the anelastic formulation, and one can have anelastic formulations of statistically inhomogeneous flows (a cloud bubble).

10. Equation (1), what are the assumptions for using this equation? The same applies for instance to the description of collision-coalescence and relates to point 3 before, namely, that there are still some assumptions in the SDM and it might be convenient to explain them as clearly as possible to better interpret the results and the limits of applicability.

11. Section 3.3.1, *"The effective resolution is* $6\Delta - 10\Delta$ *for planetary boundary layer turbulence"* I assume that this is for the second-order methods that the authors use, but they might indicate it explicitly here because the effective resolution might be substantially smaller in pseudo-spectral schemes or similar, which are often used in boundary-layer meteorology [Sullivan and Patton, 2011, Pope, 2000].

12. In section 6.4.1, I was wondering how much computational time or memory save is gained by using FP32 instead of FP64 in that part of the microphysics. I was trying to have a sense of priorities in addressing the standing challenges. For instance, the disk space challenge indicated in section 6.3 seems most important, but I might be wrong.

13. Conclusions, line 1129 *"In contrast, the constant multiplicity method is a natural choice for DNS"*. I think I did not understand this sentenc.

14. I think the conclusions do not refer to section 3.3.2 "Super-droplet movement", which I thought was interesting in general, not only for the particular implementation described here.

References

J. P. Mellado, C. S. Bretherton, B. Stevens, and M. C. Wyant. DNS and LES of stratocumulus: Better together. *J. Adv. Model. Earth Syst.*, 10:1421–1438, 2018.

P. Moin and K. Mahesh. Direct numerical simulation: A tool in turbulence research. *Annu. Rev. Fluid Mech.*, 30:539–578, 1998.

S. A. Orszag and G. S. Patterson. Numerical simulation of three-dimensional homogeneous isotropic turbulence. *Phys. Rev. Lett.*, 28(2):76–79, 1972.

S. B. Pope. *Turbulent Flows*. Cambridge University Press, 2000.

B. Schulz and J. P. Mellado. Competing effects of wind shear and droplet sedimentation at stratocumulus tops. *J. Adv. Model. Earth Syst.*, 2019. doi: doi:10.1029/2019MS001617.

P. P. Sullivan and E. G. Patton. The effect of mesh resolution on convective boundary layer statistics and structures generated by large-eddy simulations. *J. Atmos. Sci.*, 68:2395–2415, 2011.