Towards Multiscale Modeling of Ocean Surface Turbulent Mixing Using Coupled MPAS-Ocean v6.3 and PALM v5.0

Qing Li and Luke Van Roekel

This paper discusses an implementation of super-parameterization (SP) in an ocean model. The embedded small scale model is ported to run on GPUs with openACC and a notable ~10x speed up is reported. A series of simulations is presented to explore the impact of using SP in limited regions.

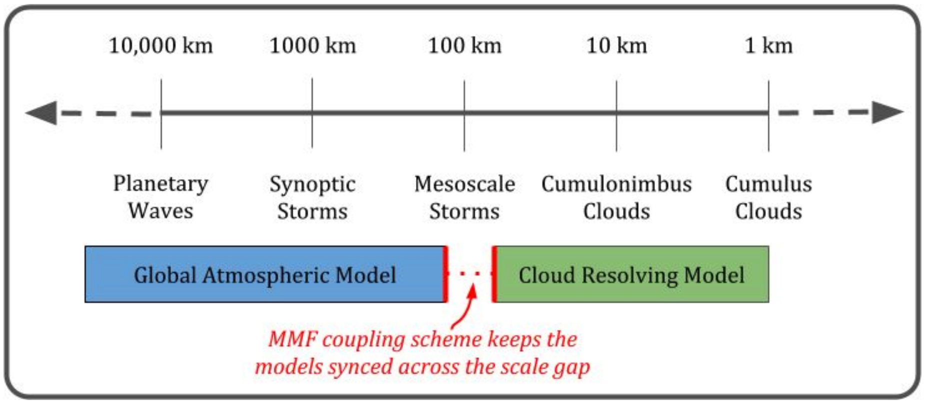
I am not an ocean modeller, so I don't have much to say about the experiments in the latter part of the paper, but they seem reasonably insightful. However, I have worked a lot with super-parameterization in the atmosphere and I take issue with the discussion of the SP implementation presented in this paper. I feel that the authors have erroneously characterized the paradigm. I have outlined my concerns about this below, along with a specific comment about the estimation of GPU speedup. Other than these concerns the paper is very well written.

Mischaracterization of super-parameterization

First off, on a semantic note, I prefer the term "multi-scale modelling framework" (MMF) over "super-parameterization" (SP). I used to consider these as interchangeable terms, but over time I have found that the notion of "replacing parameterizations" with a small-scale model often leads to misunderstandings. I feel "MMF" is the more accurate term since two distinct models that cover different scale ranges are being "coupled" together, which is different than using one model to simply replace the low-order parameterization tendencies of the other.

A key aspect of the MMF/SP idea, going back to the original work by Wojciech Grabowski, is that the large-scale and small-scale models are "tightly" coupled, *which is not the same as nudging*, even though the forcing and feedback terms may resemble nudging tendencies. On line 217 of the current manuscript the authors state that the requirement of mean state equality between the two models is enforced by either nudging the small-scale model or simply replacing the small-scale mean state with the large-scale mean state. I went back to the Khairdinov et al. (2005) paper cited in that sentence, and it makes a single mention of "nudging" when talking about the uncoupled 2D momentum field, but that paper never describes the coupling method as "nudging". The use of "nudging" implies that the small-scale model has no ability to impact the GCM state variables, but this is an erroneous characterization of traditional super-parameterization. As far as I know, nudging has *never* been used for SP in an atmosphere model outside a few special cases like the Q3D model of Jung and Arakawa (2014) and maybe the earliest implementation by Grabowski.

The notion of tight coupling means that the smallest scale of the large-scale model is enforced to be equal to the largest scale of the small-scale model (i.e. domain mean) through the formulation of the forcing tendencies in both directions. This keeps the models synchronized across the scale gap (see diagram).



Another way to think about the coupling strategy is that it is very scale selective. This is why Grabowski (2004) formulates the forcing/feedback tendencies to occur specifically on these scales. Having this mindset, I was very confused when reading section 2.1 of the current manuscript, which seems to formulate the coupling *at the smallest scale of the small-scale model*. The authors ultimately use a nudged version of the classic SP formulation, so I fail to see the relevance of section 2.1 to the manuscript, outside of the discussion of the impact of gradients on the small-scale processes.

On that note, the idea of explicitly including the effect of large-scale gradients in the small-scale model is very interesting. The authors' discussion of this seems to be in the context of equations (19) and (21), but following from my comments above, these equations are not consistent with the traditional SP formulation in which the coupling occurs at the largest scale of the small-scale model. Equations (19) and (20) imply that the coupling is valid on the smallest scale of the small-scale model, which is a very intriguing concept. However, It is difficult to imagine how the concept of "tight coupling" could be applied with this approach, so a nudging framework would probably be needed.

The idea of representing gradients in the small scale model would also require overcoming the periodic boundary conditions. The authors don't really address how this would be possible, except for a mention of perhaps needing to exchange lateral boundary fluxes between the different instances of the small-scale model. It's worth noting that this idea is problematic for performance reasons. It seems that this would require a huge increase in inter-process communication, and would certainly ruin any potential GPU speedup due to the extra GPU/CPU data exchange. I don't think the authors have really thought through these issues based on the discussion in the manuscript.

In summary, I think the authors need to revisit their description of the method used to couple the two models with special attention paid to the scale at which the coupling occurs, as well as a more accurate characterization of previous work.

GPU Speedup

When estimating the GPU speedup for E3SM-MMF we often use an entire Summit node (2 CPUs vs 6 GPUs), but we still have ongoing discussions about how to make the CPU vs GPU comparison "fair". I believe our argument for using (2 CPUS with 42 MPI tasks) vs (2 CPUS + 6 GPUs with 12 MPI tasks) is based on power consumption, along with some subtle aspects of our specific configuration. We also often estimate GPU speedup with standalone versions of the small-scale model to isolate its performance from the large-scale model. Obviously, estimating the model throughput "per watt" would be a much more ideal way to measure speed-up for these different configurations, but that is difficult to obtain.

For the estimate of GPU speedup for PALM, I think mentioning these concerns would be a nice addition to the discussion. Also the number of MPI tasks is important to mention.

- Walter Hannah

References

Grabowski, W. W., 2004: An improved framework for superparameterization. J. Atmos. Sci., 61 , 1940–1952.

Khairoutdinov, M. F., D. A. Randall, and C. DeMott (2005), Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes, J. Atmos. Sci., 62, 2136–2154.

Jung, J.-H., and Arakawa, A. (2014), Modeling the moist-convective atmosphere with a Quasi-3-D Multiscale Modeling Framework (Q3D MMF), J. Adv. Model. Earth Syst., 6, 185– 205, doi:10.1002/2013MS000295.