Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-182-RC2, 2020

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Advanced parallel implementation of the coupled ocean-ice model FEMAO with load balancing" *by* Pavel Perezhogin et al.

Anonymous Referee #2

Received and published: 5 October 2020

This paper describes the parallelization of the Finite Element Model of the Arctic Ocean (FEMAO). The focus of the paper is primarily on the technical aspects of domain decomposition and load balancing. While none of the techniques the authors present are new, they have done a couple things a little differently and it's a useful documentation of the model infrastructure so will recommend acceptance with minor revisions.

The main subject of the paper is the MPI implementation and load balancing, but I suspect some aspects of the model are limiting on-node performance. For example, they describe their choice of conditional masking vs multiplicative masking for land points (pg 2 line 42), their non-optimal combination of loop and index ordering (pg

2 L90), and the potential advantages of unstructured meshes (p2, L41). They have included at least some discussion of these in the paper so I'm not suggesting any changes now, but may have some implications on later comments below and would encourage them to explore these as they continue their optimizations in the future.

The focus of the paper is a domain decomposition that subdivides the domain into small blocks that are distributed across MPI ranks. The blocks can be made small and oversubscribed to ranks for better load balancing. The actual partitioning of the mesh uses a space-filling curve approach devised by Dennis. These approaches are typical in the ocean/sea-ice modeling community and appropriate.

A big part of the paper is an exploration of various load balancing techniques. Not surprisingly, they find the load balance is different for the 3-d ocean model vs the quasi-2d sea-ice and find that a weighted blending of the approaches works best. While the current practice in many coupled climate models is to run the two components in different partitions, there is a growing trend to recombine the ice and ocean components, so this weighted load balancing may be of interest to those efforts.

The main difference in the authors' approach is that rather than store these small sub-blocks as individual contiguous haloed arrays as other models have done, they create a larger rectangular domain that geographically covers the area of all the local sub-blocks and creates local index ranges for each of the sub-blocks. This eliminates the need for local halo copies and probably allows them to push toward very tiny sub-blocks (4x4 in many cases presented), but at the expense of contiguous data access and the ability to independently thread over sub-blocks.

Most of the paper is fine, but I remain confused over some of the discussion in section 6.1 and the scatter plots in figures 5,6.

First, the authors show speedups in figure 4 with significantly super-linear speedups in the 2-d case. They attribute this to cache performance without additional evidence (eg from hardware counters or other performance tools). That may be the case, but I think this super-linearity is large enough to warrant further exploration into the cause.

Second, the computational time as a function of wet points seems a bit counter-intuitive (Fig. 5). The authors have shown percentage of wet points rather than total wet points to emphasize their diagnosis again of memory access. But without also seeing the total number of points (computational load), it's a little hard to get a more complete picture. Again, this effect seems too big to attribute solely to cache effects and it seems like more might be going on here.

Third, the large variation in work load at high core counts (fig 5,6) also seems higher than one might expect. As you get fewer points/blocks per core, there will naturally be a little higher variability, but this seems larger than expected and might point to additional problems.

I suspect some further analysis of the edge cases in the above would help to illustrate what is going on. The issues may go beyond cache performance and may be partly due to the on-node choices mentioned earlier.

A few other minor edits:

The journal editor will probably mention this, but most references should be changed so that the parentheses are around both author and date unless an integral part of sentence. So for example p1L22-23, should have (FEMAO; Iakolev, 1996, 2012) and (Chernov, 2013; Chernov et al. 2018). And so on throughout the manuscript.

Fig 2 has cropped the bottom of figures

P9L170-180; In this bulleted list, move the text "These three bullets…" and "This reduces…" after the bulleted list as "The first three bullets…" and "The final bullet…" Mixing these comments in with the bulleted list was confusing.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-182, 2020.