Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-126-RC1, 2016

Geoscientific Model Development Discussions
Open Access

Interactive comment on "A diagnostic interface for the ICOsahedral Non-hydrostatic (ICON) modelling framework based on the Modular Earth Submodel System (MESSy, 2.50)" *by* Bastian Kern and Patrick Jöckel

I. Honkonen (Referee)

ilja.j.honkonen@nasa.gov

Received and published: 2 August 2016

Overall the paper is quite good but a some clarifications would allow the presented work to be compared with existing literature in space sciences where similar work already exist. My answers to GMD review criteria are at the end.

The authors implement a module for ICON modeling framework that allows run-time post-processing (PP), for which they use the term online diagnostics, of simulation results. Supported PP operations are the calculation of spatial average, spatial sum, discrete probability density function (PDF) and a two variable joint PDF. These can be

calculated over a rectangular (in latitude and longitude) grid defined by the user. In a parallel run, if a cell of the PP grid spans more than one simulation cell, data required by PP operations is communicated between all processes that have simulation cells within the PP cell.

The term post-processing seems applicable here as the diagnostic module does not affect simulation results. The same term is also used e.g. in OpenFOAM documentation which shows an example of streamlines calculated at run-time (http://www.openfoam.com/version-v3.0+/post-processing.php retrieved on 2016-08-01). On the other hand the presented module could also be described as a one-way coupled model configurable with Fortran namelists.

Overall it seems that the presented implementation can be divided into two distinct parts: 1) A new one-way coupled model for the MESSy framework that allows new variable(s) to be added to the framework relatively easily and allows calculations performed by the model to be chosen at the start of simulation. 2) A method for moving data from simulation grid to the PP module grid in a transparent way that supports different grid geometries, cell sizes and aggregation of all data from simulation cells into overlapping PP cells.

Major comments

Item 1) above looks like a regular MESSy component which uses the same grid as the basemodel, would that be a fair statement? Can several PP modules run simultaneously on different sets of processes? Can a PP module use data from other PP modules or only simulation results i.e. non-PP or basemodel data? Is two-way coupling of PP modules possible, i.e. can two PP modules use each others' output data as their own input?

Item 2) above allows a PP module to use a different grid from the one used by basemodel. In the presented work, cells of PP grid are larger than in simulation grid, is it possible to make PP grid cells smaller than in simulation grid?

Taken together the functionality of items 1) and 2) are quite similar to e.g. the Space Weather Modeling Framework (SWMF, doi:10.1029/2005JA011126) which handles two-way data exchange between separate models running in parallel on different grids, although SWMF seems to lack the ability to collect all data from multiple smaller cells of the source grid into one cell of the destination grid. It would be helpful to clarify how the functionality presented here relates to that available in SWMF.

The support, or whether there is any, for temporal analysis within the PP module was unclear to me. Is it possible to process simulation data over several time steps while calculating one diagnostic output, e.g. can the minimum and maximum values of a variable over a certain time range in each cell of the PP grid be determined? This would determine the min and max values more reliably than calculating them from saved model output which presumably would not be feasible to do for every simulation time step.

If the above is possible, then are MPI operations within each PP grid cell available at each simulation time step, or only at the end of a PP step? I think MPI operations at each simulation step would be needed to e.g. reliably calculate a time series of the variance of a simulation variable within each PP cell because the average in each PP cell would have to be communicated at each simulation step so processes can calculate the variance of their simulation cells for that step.

Minor comments

The authors did not show initialization times for different test cases but would initialization of the PP system be feasible to perform e.g. every 10th simulation time step? The use case would be e.g. adaptive mesh refinement of simulation grid which would require the cell/process information to be updated between simulation and PP grids.

The authors, as well as the original Zängl et al. reference, use the term unstructured grid for the simulation grid. I would perhaps hesitate to call it unstructured because it seems that cells cannot be divided into an arbitrary number of smaller cells, for exam-

ple. While the number of neighbors of a cell can vary due to mesh refinement the grid seems to be structured e.g. in the sense that every possible cell of the grid, i.e. its relative size, location and relative position to other cells, can be uniquely represented by a single integer. This was done for a cartesian grid in doi:10.1016/j.cpc.2012.12.017 where each cell can be divided into 2N cells along each edge (instead of N here), where N is a positive integer.

Technical suggestions/corrections

Page 4, line 16: und -> and

Page 6, figure caption: used for the calculation of averages. -> used for diagnostics in figures 2 and 3.

Answers to question at geoscientific-model-development.net/peer_review/review_criteria.htm

Does the paper address relevant scientific modelling questions within the scope of GMD? yes

Does the paper present a model, advances in modelling science, or a modelling protocol that is suitable for addressing relevant scientific questions within the scope of EGU? yes

Does the paper present novel concepts, ideas, tools, or data? yes

Does the paper represent a sufficiently substantial advance in modelling science? yes

Are the methods and assumptions valid and clearly outlined? yes

Are the results sufficient to support the interpretations and conclusions? yes

Is the description sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? with a few additions, yes

Do the authors give proper credit to related work and clearly indicate their own new/original contribution? yes

Does the title clearly reflect the contents of the paper? yes

Does the abstract provide a concise and complete summary? yes

Is the overall presentation well structured and clear? yes

Is the language fluent and precise? yes

Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? yes

Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? after traceability is improved, no

Are the number and quality of references appropriate? with SWMF reference, yes

Is the amount and quality of supplementary material appropriate? yes

---

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-126, 2016.