

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

Bastian Kern and Patrick Jöckel

bastian.kern@dlr.de

Received and published: 2 September 2016

We thank Ilja Honkonen (**Referee #1**) for his comments on the manuscript and for the insight provided by his point of view. The clarifications we include in the revised manuscript really help to improve it.

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.

C1

The authors implement a module for ICON modeling framework that allows runtime 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 runtime (<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.

We are a bit reluctant to use the term post-processing here because:

- 1) We want to clearly distinguish between the “on-line post-processing” and classical post-processing of output data subsequent to a simulation, because in our community, post-processing is mostly associated with “off-line post-processing”.
- 2) It is correct that there is no feedback from the on-line diagnostic tools to the model variables. But as the MESSy framework was introduced to make exactly this feedback possible (see also below), we want to stress that the MESSy submodels can be divided into *process submodels* influencing the results, the purely *diagnostic* ones, and the submodels providing the underlying infrastructure.
- 3) Moreover, ICON itself contains already a specific facility to apply post-processing to data just before the output. With our naming (“diagnostic”) we want to make clear that the data produced by the *diagnostic* submodels is available in the memory (also for other submodels, see below) and not a concluding post-processing step.

C2

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.

What you refer to as item 1) is an already existing infrastructure submodel of MESSy, namely CHANNEL (Jöckel et al., 2010), which is used for memory and meta-data management, output control, and checkpointing. This submodel has been extended to meet the needs of the new basemodel ICON (as described in Section 2.3 and Appendix A1 of the manuscript).

Your item 2) is the novel *diagnostic* submodel GRAGG (GRid AGGregation), which utilises CHANNEL (cf. item 1) for meta-data management, memory access and output, i.e. for the definition of new variables (so called CHANNEL *objects*).

Indeed, we never thought of the GRAGG submodel as a transparent re-gridding tool, but as all submodels in the MESSy framework can access data in the memory via the CHANNEL submodel, it could be seen this way. Thank you for this insight. As there is already a generic MESSy submodel for re-gridding available, we should think of a further generalisation.

We include the idea of a more generalised re-gridding in Section 5 “*Ongoing developments*”, which we rename to “*Ongoing and future developments*”:

C3

(p. 16, l. 6)

“Last, but not least, as there is already a re-gridding tool available in the MESSy framework, an integration into the GRID submodel (Kerkweg and Jöckel, 2015) with further generalisation of the aggregation developed in this study seems beneficial.”

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?

The MESSy submodel CHANNEL incorporates the concept of *representations* (see Jöckel et al., 2010), which define the underlying geometric structure of *objects* (i.e. variables). Those can be defined as required. Thus, basemodel variables are naturally defined on the native grid geometry of the basemodel. Nevertheless, any MESSy submodel can define its own geometries.

Can several PP modules run simultaneously on different sets of processes?

MESSy is not a classical external coupler, although it contains also an infrastructure submodel for external coupling (Multi Model Driver, MMD, Kerkweg and Jöckel, 2012b). For the *diagnostic* submodels discussed here, several of those can be applied internally coupled, i.e., in sequence on the same task set.

Can a PP module use data from other PP modules or only simulation results i.e. non-PP or basemodel data?

C4

CHANNEL *objects* can be accessed from *all* submodels, implying also that submodel A can access *objects* from submodel B.

Is two-way coupling of PP modules possible, i.e. can two PP modules use each others' output data as their own input?

Yes, because CHANNEL *objects* can be accessed from everywhere.

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?

For technical reasons, this is currently not possible in GRAGG, because the intention of this submodel is data reduction.

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.

We were not aware of the SWMF, which was initiated at about the same time as the MESSy framework, but we know some frameworks and couplers used in climate science. We think the framework and purpose of MESSy was exhaustively described in various articles (Jöckel et al. 2005, 2010, 2016). However, we add a short paragraph to our manuscript, which briefly relates MESSy to similar frameworks and couplers and

C5

gives some examples of recent applications. References to comparable frameworks in climate science and the SWMF (See *References* below) are included in Section 2.2 of the revised manuscript.

Addressing the points above, we add the following paragraphs to the revised Section 2.2:

(p. 4, l. 9)

“In geoscientific modelling, coupling of multi-institutional codes with generally different domain decompositions is a widely used approach for building model systems. In general, either external couplers or frameworks for internal coupling are applied. An extended classification of coupling methods can be found in Appendix A of Kerkweg and Jöckel (2012b) and in Jöckel (2012). An overview of different coupling techniques in Earth System Modelling is presented by Valcke et al. (2012). Common external couplers in the Earth System Model community are, e.g., OASIS3 (Valcke et al., 2006; Valcke, 2013), OASIS4 (Redler et al., 2010), and CPL6 (Craig et al., 2005), as used in the Community Climate System Model version 3 (CCSM3, Collins et al., 2006). Widely used examples for internal coupling are the Earth System Modeling Framework (ESMF, Collins et al., 2005) and the Community Climate Model version 4 (CCSM4, Gent et al. 2011). This approach is also used in space weather modelling with the Space Weather Modeling Framework (SWMF, Tóth et al., 2005). Recently, Hanke et al. (2016) developed the C-library YAC (Yet Another Coupler), which provides parallelised and efficient algorithms for grid transformation, interpolation, and data exchange.

In contrast to the coupling of “domains”, MESSy was originally developed to work on the same spatial domain and parallel domain decomposition as the basemodel, applying a formalised process based operator splitting (Jöckel et al., 2005). The original imple-

C6

mentation was intended to equip the atmospheric general circulation model ECHAM5 (Roeckner et al., 2006) with additional processes for atmospheric chemistry (EMAC, ECHAM5/MESSy Atmospheric Chemistry model, Jöckel et al., 2006, 2010). Operator splitting as internal coupling method is implemented (implicitly and less formalised) in the numerical model codes anyway, to integrate the different processes. However, the operator splitting approach of MESSy proves more powerful, also allowing for coupling of different domains, e.g., demonstrated by the integration of an ocean subsystem (Pozzer et al., 2011). An extension by Kerkweg and Jöckel (2012b) allows for one-way coupling of different spatially nested domains using a server-client approach with point-to-point communication. Extension into a two-way nested atmospheric chemistry model system is currently under development. ”

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.

In GRAGG, a temporal analysis is not included yet. However, the generic submodel CHANNEL already supports temporal operations (MIN, MAX, AVE, ...) over the output interval (see Section 2 and the Supplement of Jöckel et al. 2010). This facility can also be used for the variables calculated in GRAGG.

In Section 3 we add:

(p. 4, l. 9)

C7

“Note that a temporal aggregation is not envisaged for GRAGG, because the generic CHANNEL submodel already allows time aggregation operations, such as minimum, maximum, average and standard deviation, etc. over the output time interval (see Section 2 and the Supplement of Jöckel et al. 2010). This facility can also be used for the variables calculated in GRAGG.”

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.

The calculations of GRAGG and VISOP are currently triggered as requested by the user via namelist, either **every** time-step, or only every regular output time-step. For the calculation of correct time average, minimum, maximum, standard deviation, etc. w.r.t time, the trigger needs to be set to “every time-step”. This provides the flexibility to reduce the expensive MPI operations to the minimum required. We plan to automate this in that way that an output request of temporal statistics (MIN, MAX, AVE, ...) via CHANNEL automatically triggers the corresponding calculations in every time-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.

C8

Since we currently do not use adaptive meshes, the weights for the grid transformations are calculated only once during the initialisation phase. An extension for time varying adaptive meshes is currently not planned.

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 example. 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.

The grid really seems structured, but the neighbour cells can not be deducted from the memory layout. As the grid is constructed from an icosahedron, resulting in vertices surrounded by five triangles and vertices surrounded by six triangles. The neighbourhood structure has to be explicitly stored, and hence the grid is indeed unstructured (https://en.wikipedia.org/wiki/Types_of_mesh).

Technical suggestions/corrections

Page 4, line 16: und -> and

Oh, the German tongue is breaking trough ;-) . Is corrected.

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

C9

The referee is right that the grid is not only used for the calculation of averages, we update the sentence in the manuscript as follows:

“Overlaid is the user-defined regular coarse grid (0.5×0.5) used to perform the on-line diagnostics in this study.”

References

Collins, N., Theurich, G., DeLuca, C., Suarez, M., Trayanov, A., Balaji, V., Li, P., Yang, W., Hill, C., and da Silva, A.: Design and Implementation of Components in the Earth System Modeling Framework, *International Journal of High Performance Computing Applications*, 19, 341–350, doi:10.1177/1094342005056120, <http://hpc.sagepub.com/cgi/doi/10.1177/1094342005056120>, 2005.

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. D., and Smith, R. D.: The Community Climate System Model Version 3 (CCSM3), *Journal of Climate*, 19, 2122–2143, doi:10.1175/JCLI3761.1, 2006.

Craig, A. P., Jacob, R., Kauffman, B., Bettge, T., Larson, J., Ong, E., Ding, C., and He, Y.: CPL6: The New Extensible, High Performance Parallel Coupler for the Community Climate System Model, *International Journal of High Performance Computing Applications*, 19, 309–327, doi:10.1177/1094342005056117, 2005.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang,

C10

Z.-L., and Zhang, M.: The Community Climate System Model Version 4, *Journal of Climate*, 24, 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

Hanke, M., Redler, R., Holfeld, T., and Yastremsky, M.: YAC 1.2.0: new aspects for coupling software in Earth system modelling, *Geoscientific Model Development*, 9, 2755–2769, doi:10.5194/gmd-9-2755-2016, 2016.

Jöckel, P.: Earth System Modeling, in: *Atmospheric Physics*, edited by Schumann, U., chap. 35, pp. 577–590, Springer, Berlin Heidelberg, Germany, doi:10.1007/978-3-642-30183-4_35, 2012.

Jöckel, P., Sander, R., Kerkweg, A., Tost, H., and Lelieveld, J.: Technical Note: The Modular Earth Submodel System (MESSy) - a new approach towards Earth System Modeling, *Atmos. Chem. Phys.*, 5, 433–444, doi:10.5194/acp-5-433-2005, 2005.

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A. J. G., Gromov, S., and Kern, B.: Development cycle 2 of the Modular Earth Submodel System (MESSy2), *Geoscientific Model Development*, 3, 717–752, doi:10.5194/gmd-3-717-2010, 2010.

Kerkweg, A. and Jöckel, P.: The 1-way on-line coupled atmospheric chemistry model system MECO(n) – Part 1: Description of the limited-area atmospheric chemistry model COSMO/MESSy, *Geoscientific Model Development*, 5, 87–110, doi:10.5194/gmd-5-87-2012, 2012a.

Kerkweg, A. and Jöckel, P.: The 1-way on-line coupled atmospheric chemistry model system MECO(n) – Part 2: On-line coupling with the Multi-Model-Driver (MMD), *Geoscientific Model Development*, 5, 111–128, doi:10.5194/gmd-5-111-2012, 2012b.

C11

Kerkweg, A. and Jöckel, P.: The infrastructure MESSy submodels GRID (v1.0) and IMPORT (v1.0), *Geoscientific Model Development Discussions*, 8, 8607–8633, doi:10.5194/gmdd-8-8607-2015, 2015.

Pozzer, A., Jöckel, P., Kern, B., and Haak, H.: The Atmosphere-Ocean General Circulation Model EMAC-MPIOM, *Geoscientific Model Development*, 4, 771–784, doi:10.5194/gmd-4-771-2011, 2011.

Redler, R., Valcke, S., and Ritzdorf, H.: OASIS4 – a coupling software for next generation earth system modelling, *Geoscientific Model Development*, 3, 87–104, doi:10.5194/gmd-3-87-2010, 2010.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model, *Journal of Climate*, 19, 3771–3791, doi:10.1175/JCLI3824.1, 2006.

Tóth, G., Sokolov, I. V., Gombosi, T. I., Chesney, D. R., Clauer, C. R., De Zeeuw, D. L., Hansen, K. C., Kane, K. J., Manchester, W. B., Oehmke, R. C., Powell, K. G., Ridley, A. J., Roussev, I. I., Stout, Q. F., Volberg, O., Wolf, R. A., Sazykin, S., Chan, A., Yu, B., and Kóta, J.: Space Weather Modeling Framework: A new tool for the space science community, *Journal of Geophysical Research: Space Physics*, 110, A12 226, doi:10.1029/2005JA011126, 2005.

Valcke, S.: The OASIS3 coupler: a European climate modelling community software, *Geoscientific Model Development*, 6, 373–388, doi:10.5194/gmd-6-373-2013, 2013.

C12

Valcke, S., Guilyardi, E., and Larsson, C.: PRISM and ENES: a European approach to Earth system modelling, *Concurrency and Computation: Practice and Experience*, 18, 247–262, doi:10.1002/cpe.915, 2006.

Valcke, S., Balaji, V., Craig, A., DeLuca, C., Dunlap, R., Ford, R. W., Jacob, R., Larson, J., O’Kuinghtons, R., Riley, G. D., and Vertenstein, M.: Coupling technologies for Earth System Modelling, *Geoscientific Model Development*, 5, 1589–1596, doi:10.5194/gmd-5-1589-2012, 2012.

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