Response to GMD-2016-47-RC1 (anonymous)

Introduction

Coding and technical aspects of coupling Earth System Models are often relegated to institutional reports seldom referenced or widely read, and outcomes of work in coupling and load balancing are often blindly used by physical and biogeochemical modeling groups. Therefore I commend the authors for documenting their expansive coupling work, and for submitting it to be reviewed for a journal with a readership that bridges the coupler development and physical modeling communities. I do, however, have reservations about the final results, methods, and one comment about the scope of the cited literature.

Answer: we appreciate that the effort we made to document our technical work could be hosted by the journal. In this revision, we try to answer the referee questions, adding details, without overloading the article too much.

Significance of this paper in the context of other work

The manuscript's introduction could extend the perceived reach of the work if it were to illustrate its international significance. There are several latest-generation regional coupled earth system models in development in the U.S. and Canada, some of which use MCT and takes advantage of the work of Craig et al. (2012) that could have been cited and have appeared in the reviewed literature in recent years. The reason I mention these publications is to say that the introductory argument perhaps could be further enhanced, since work on load balancing high-resolution regional coupled earth system models is taking place in many parts of the Earth System Modeling community. This helps to widen the appeal of the current manuscript, and its significance.

Answer: the referee rightly emphasises that load balancing is not a new issue in our community. Studies based on CESM model, for example, are familiar to the authors (Craig 2012 is cited in the article, line 385). We have also mentioned Dennis et al. 2012 (line 88) and Alexeev at al. 2014 (line 90) in our introduction. We added a reference to Balaprakash et al 2014 (line 809) in chapter "4.3 Strategy for finding an optimum configuration" and we have mentioned that, in our case, "due to the heterogeneity of our coupled systems, a single algorithm cannot be proposed (as in Balaprakash et al, 2014)". Unless the CESM package, the OASIS library allows an unlimited kind of component combination in coupled systems. For the moment, it is rather complicated to propose an automatic load balancing tool that could deliver an optimal solution for all combinations. We hope that the present article will help the OASIS community to develop an ability to better balance their systems and, in a second step, propose solutions that may be gathered in a single tool.

Efficiency versus accuracy

This paper discusses a considerable number (five) of different coupled model configurations using CCLM, however only scant information is provided on each one of these configurations. It would be particularly useful to view maps of model domains to demonstrate the individual configurations for each of the coupled model systems in Table 2.

Answer: Model domains are shown now in Figure 1.

This would help make it clear exactly how much ocean, land and sea ice exist in the respective model domains. Such details can have a large impact on scalability and parallel efficiency, especially in the cryosphere (sea ice and snow). Therefore I suggest providing greater detail on the physical configuration of each of the models chosen, because this, too, has an enormous impact on the model solution.

Answer: We fully agree with the reviewer that details on the physical configuration have an impact on individual performances of components, and consequently, on performances of the whole coupled system. However, the article is not investigating the physical performance of the coupled systems. It is rather focussing on presentation of the coupling method, the computational performance in COSMO-CLM reference configuration and finding of an “optimum configuration”. It is thus out of scope to discuss in detail the amount of ice and snow in the model domain and the impact on computational performance. It is also out of scope to discuss physical and dynamical parameters that could influence the computing performance. This remains for future work. Further below we propose an improved definition of our component characteristics by parameters relevant for computing performances. We also answer the referee’s questions about CICE.

To illustrate this point, I focus here on the implementation of CICE Version 5 for CCLM+TRIMNP+CICE. The computational efficiency of the solution in CICE is heavily dependent upon the total number of sea ice thickness categories used, the number of tracers needed, for example, by melt pond and ice-age tracking and biogeochemistry, and most importantly, the sea ice mechanics solution. If CICE 5 has been configured to use anisotropic (Elastic Anisotropic Plastic; EAP) sea ice mechanics, then it will definitely be expensive, and, could take as much as 30% of the total model execution time in pan-arctic fully coupled regional models, if a highly converged plastic sea-ice solution is required (2-second sub-cycling). However, if using the Elastic Viscous Plastic (EVP) sea ice rheology with 10-second sub-cycling, the time to solution of the sea ice model greatly improves, with only slight degradation of the plastic solution. In this configuration, the sea ice model could take only 10% of the total core time of running the model. It is still unknown as to which of the two variants is physically more accurate. This is precisely the same CICE Version 5.1 code, in the same coupled framework, using MCT, but with two different namelist settings yet to be fully explored in the literature. Further issues with the CICE coupling are discussed in the appendix.
Answer: EVP was used (kdyn=1). However, CICE domain covers only the Baltic Sea and Kattegat, not the pan-arctic. The sea ice which appears in a relatively small domain like the Baltic Sea and disappears totally in summer has less complicated features compared to the Arctic. However, we cannot say how much different the calculations would be if EAP was chosen, as no sensitivity tests about these parameters have been conducted. The scope of the paper was to present a strategy of analysis of the computational performance of the coupled system in comparison to stand-alone performance. A deeper analysis is out of scope of the paper and remains for future work. We highlight the relevance and the opportunities of such an analysis in the result section for the CCLM+MPI-ESM coupling (line 1111 ff).

This CICE anecdote drives at my main criticism of this paper as it currently stands: It seems to be a vacant conclusion to discuss model efficiency without discussing model accuracy. The most efficient model one can design is a constant number, but seldom is this model the most accurate. The only way this limitation in the current manuscript can be remedied is to explicitly state the configurations used for each particular model in the tests presented, including graphically representing the domains used. However, due to the number of different models and model configurations used, this may balloon the paper to unmanageable proportions. However, as the paper currently stands, there is too little information available for it to be useful for other groups trying to address coupled model efficiency in their particular configurations.
Answer: The aim of the paper is to analyse the performance of the coupled systems using a configuration common for climate applications. Therefore, the analysis of computational performance was conducted using well tested and recommended climate modelling configurations for each component model without any idealisation, e.g. the I/O is the same as in standard climate applications. This is described in section 4.1, line 746 ff.. We agree with the reviewer that a detailed description of all configuration would balloon the paper and

hope having found an appropriate compromise concentrating on configuration details specific for the couplings described in chapter 2.

However, the computing performances of the coupled system necessarily depends on the performances of each component. We agree with the referee that the choice of an additional component cannot only depend of its computing cost. Obviously, the model accuracy (or model skill) is the most important criterion. The article does not say anything about component accuracy in stand alone mode or, what we consider to be even more important, component accuracy in coupled mode. This article addresses the usability of configurations, which is a prerequisite of scientific analysis described in other papers, as for example Pham et al. 2016 (CCLM+NEMO-NORDIC) or Davin et al. 2016 (CCLM+CLM).

Nevertheless, we agree that more information is usefull to facilitate the comparison of component costs and to estimate the cost of possible other configurations (e. g. with other resolutions). An interesting suggestion is the computing performance metrics described in Balaji et al. 2017, particularly the 2 parameters describing the models: resolution and complexity. “Resolution” -G- is measured as the number of grid points (or more generally, spatial degrees of freedom) NX,NY,NZ per component. “Complexity” -V- is measured as the number of 3D prognostic variables per component (to be able to compare 3D models, like atmosphere, with 2D models, like land models, it is assumed that V of 2D models are equal to 1). These 2 parameters are added in Table 3.

G and V are key parameters to explain why some components are more costly than others (MPI-ESM, with highest G and V, is also the one which induces the highest coupling cost). This information is emphasised in § 4.5 “Extra time and costs”, line 1005 ff. It can also be used for users who would like to estimate the extra cost induced by changes in a coupled component, like a resolution increase (horizontal or vertical) or a complexity increase (additional calculations like biogeochemistry in the ocean or chemistry in the atmosphere ...)

Conclusion

In some respects, the scope of this paper is too large and should be refined. The concluding arguments would be far more compelling, and, I believe, interesting to the modeling community, if it explored individual coupled configurations, and efficiency related to a group of relatively standard model settings in each component model. However, this is probably beyond the scope intended by the authors, and therefore one way to make sure the good work already done is published would be to: 1) Provide greater details of each of the models used to produce the results, including model domain maps, of the model configuration tables, the latter in an appendix; and 2) Provide at least some indication of the accuracy of the solutions. Otherwise, one is left to wonder as to how exactly the results were produced.

Answer: As already stated, (1) we use recommended and overall tested model configurations for climate application over Europe. (2) a map is added in Figure 1 showing the model domains and (3) metrics are added in Table 3 to better estimate the model accuracy. Furthermore, the results of computational performance are revised and presented in a more consistent way. Figures 5 and 6 together with table 8 provide consistent results. In table 8 the section 3.3 shows a systematic analysis of extra costs of coupling for all couplings investigated at optimum configuration. The components are described in lines 920 ff.

Currently the paper fails the reproducibility test, because insufficient information is provided to repeat the experiments. This, alone, is grounds for significant revision, which I hope the authors will undertake.

Answer: We thank the reviewer for this comment. The reproducibility of results is an important aspect of community work in the CLM Community and we hope to be able to show it in the following. We added details on how to get the model versions and configurations used for the performance analysis presented in the Appendix under "source code availability", line 1135 ff. At the moment, the model versions used are not official CLM-Community model versions but available from the model developers. An implementation into an official CLM-Community released version is ongoing. Hereby we follow the procedure of source code development introduced in the COSMO and CLM Community. Each experiment can be repeated with the set up information from the article and using the model input files. To get the individual coupled systems, model input files and configuration details the authors have to be contacted as described in the Appendix.
All results presented and the original model output files used are available from the lead author, following the rules of good scientific practice.
However, the machine blizzard is not available anymore. Thus the results are, strictly speaking, not reproducible. This, however, is not the responsibility of the authors and true for each numerical model result after some years. The authors believe that the results highlighted are robust and can be obtained on a similar machine as well.
—-

Appendix – CICE configuration and coupling

This appendix addresses technicalities of the CICE setup that were puzzling to the reviewer. First, the authors may be interested to know that there were important bug fixes in the code between version 5.0 and 5.1 of CICE (update is in Hunke et al., 2015), however these would be unlikely to influence computational performance. Setting this aside, there are further improvements in the computational performance of the model using EAP that are being updated by the University of Reading at the current time. It is impossible to know whether or not this affects the results in this paper, because the CICE configuration used in this paper is never made clear. Also, and perhaps I missed it in the text, whether or not the namelist option "distribution_type" is changed in CICE is not discussed. This affects computational performance.

Answer: Parameters used in CICE and TRIMNP are the same as in real climate simulations for Europe. They are listed and discussed in the following but not included in as much detail in the paper.

<u>CICE:</u>
+ kitd = 1; ktherm = 2; conduct = 'MU71'
+ kdyn = 1 (means EVP is used); ndte = 60; revised_evp = .false.; advection = 'upwind'
+ shortwave = 'dEdd'; albedo_type = 'default'
+ tr_brine = .false.; skl_bgc = .false.; bgc_flux_type = 'Jin2006'
+ formdrag = .false.
+ tr_iage = .true.; tr_FY = .true.; tr_lvl = .true.; tr_pond_cesm = .false.; tr_pond_topo = .false.; tr_pond_lvl = .true.; tr_aero = .false.
+ distribution_type = "cartesian"; processor_shape = "square-pop"; distribution_wght = "latitude"; ew_boundary_type = "open" ; ns_boundary_type = "open"

The original formula of category boundary (kcatbound = 0) with the thickness boundaries for five thickness categories and the linear remapping of the ice thickness distribution (kitd = 1) are configured in this study. The thermodynamics option new "mushy" formulation (ktherm=2) is applied in which salinity evolves (Turner et al., 2013). For each thickness category, CICE computes changes in the ice and snow thickness and vertical temperature profile resulting from radiative, turbulent, and conductive heat fluxes. The ice has a temperature-dependent specific heat to simulate the effect of brine pocket melting and freezing. The standard thermal conductivity option used is 'MU71' following Untersteiner (1964) and

Maykut and Untersteiner (1971). The explicit melt pond parameterisation uses the delta-Eddington radiation scheme with the default (ccsm3) shortwave parameterisation which incorporates melt ponds implicitly by adjusting the albedo based on surface conditions. The revised Elastic Viscous Plastic (EVP) sea ice rheology and the upwind advection algorithm are applied.

The distribution type option is the standard Cartesian distribution of blocks which allows redistribution via a 'rake' algorithm for improved load balancing across processors, and redistribution based on space-filling curves. The processor shape is square-pop. The 'latitude' option weights the blocks based on latitude and the number of ocean grid cells they contain. The Neumann boundary conditions are set up for both east-west and north-south boundary type.

TRIMNP:
hdif_u=50., hdif_v=50., hdif_w=0., hdif_s=25., hdif_t=25., hdif_q=0.,

The dynamics of the free surface are discretised semi-implicitely, and the resulting linear equation system is solved with a pre-conditioned conjugate gradient method. The vertical mixing and friction including non-linear bottom friction and surface wind stress are also solved with a semi-implicit method. The vertical mixing and friction coefficients are parameterised using prognostic equations for turbulent kinetic energy and dissipation (Umlauf and Burchard 2005). For horizontal diffusion, harmonic terms are used with scale dependent constants. The lateral diffusion and the viscosity constants are 25 m2/s and 50 m2/s, respectively. Advection for all time-dependent variables is done with a Semi-Lagrangian method, where at the end of each time step the values of the variables at the corresponding grid points (the arrival points) are determined by following a trajectory backwards in time for one time step interval to the departure points. The values of the variables at the departure points are determined by trilinear interpolation. For details see Cheng et al. (1993).

Most importantly, however, is the information within Table 5 on how CICE is coupled to CCLM. My understanding is that the U symbol indicates fluxes being passed from CCLM to CICE. If this is the case, there is only one feedback from CICE to CCLM in Table 5 (SST), which draws into question the physical consistency of the coupling. If this were to be a fully coupled model, then there must be more feedbacks that just surface temperature to the atmosphere. For sea ice, the most important feedback is either albedo or reflected shortwave radiation, passing back from the sea ice model to the atmosphere, but neither is listed, which leads one to assume that albedo is being calculated in the atmospheric model independently. Given the sophistication of the Delta-Eddington albedo parameterization in CICE, this seems odd. This inconsistency should be addressed before publication.

It is also odd that the atmosphere is calculating sensible and latent heat fluxes, given that the CICE configuration has five sea ice thickness categories each calculating an independent surface temperature upon which turbulent fluxes are based. Hence the turbulent heat fluxes must be inconsistent with the surface stress term, which is being calculated internally in CICE in the configuration given. When this calculation is done within CICE, assuming Monin-Obukhov stability calculations are being performed, the drag coefficient accounts for the individual surface temperature of each of the five sea ice thickness categories. If this calculation is not being performed in CICE, then the only alternative would be for the sea ice model to use only neutral drag, which would also be inconsistent with the sensible and latent heat flux components of turbulent transfer being passed from the atmosphere.

The only way to remedy this is either to specify surface stress from the atmospheric model, or to fully use the turbulent transfer calculations in CICE, and pass the sensible and latent heat fluxes back to the atmosphere from the sea ice model. This is the reverse of what is currently being done, or at least described in this manuscript. This inconsistency should also be addressed before publication.

Answer: We agree that the inconsistency exists and needs to be improved in the future. We explain this inconsistency in the paper now (chapter 3.4, line 630 ff). In the experiment CCLM+TRIMNP+CICE, only SSTs are passed to the atmosphere as in the version of CCLM used at the time when the experiment was conducted for this study the partial sea ice cover, snow on sea ice and water on sea ice are not considered. In a water grid box of CCLM, the albedo parameterisation switches from ocean to sea ice if surface temperature is below a freezing temperature threshold of -1.7°C. We would have passed sea ice fraction to CCLM as it was done for NEMO-Nordic. However, we think that careful checks e.g. for reflected shortwave radiation should be made for the coupled system model CCLM+-TRIMNP+CICE if sea ice fraction and albedo from CICE are sent to CCLM. These checks remain for future work.

In the current study, no sea ice information from CICE was passed to CCLM. But they were sent to TRIMNP. In TRIMNP the surface temperature is calculated as a combination of SSTs from TRIMNP and the sea ice skin temperatures from CICE, weighted by the sea ice concentration before the combined surface temperature is passed to CCLM. In Table 5, “surface temperature over sea/ocean” is used instead of SST to avoid a potential misunderstanding in case of sea ice existence.

We also think that even if sea ice fraction from CICE is sent to CCLM, the latent and sensible heat fluxes in CCLM are still different to those in CICE due to different turbulent schemes of the two models CCLM and CICE. The inconsistency can be removed only if all models use the same energy fluxes, calculated in one model at the highest resolution, for example in CICE model, as the reviewer suggested. This strategy could be applied in future studies considering the result of this performance study, that exchanging much more fields has a small impact on cost.

References:

- Alexeev, Y., Mickelson, S., Leyffer, S., Jacob, R., and Craig, A., 2014: The Heuristic Static Load-Balancing Algorithm Applied to the Community Earth System Model, in: 28th IEEE International Parallel and Distributed Processing Symposium, no. 28 in Parallel & Distributed Processing Symposium Workshops, pp. 1581–1590, IEEE, doi:10.1109/IPDPSW.2014.177
- Balaji, V., Maisonnave, E., Zadeh, N., Lawrence, B. N., Biercamp, J., Fladrich, U., Aloisio, G., Benson, R., Caubel, A., Durachta, J., Foujols, M.-A., Lister, G., Mocavero, S., Underwood, S., and Wright, G., 2017: CPMIP: Measurements of Real Computational Performance of Earth System Models in CMIP6 , *Geosci. Model Dev.*, **46**, 19-34, doi:10.5194/gmd-10-19-2017
- Balaprakash, P., Alexeev, Y., Mickelson, S. A., Leyer, S., Jacob, R. L. and Craig, A. P., 2014: Machine learning based load-balancing for the CESM climate modeling package, in Proceedings for 11th International Meeting on High-Performance Computing for Computational Science (VECPAR 2014)
- Davin, E. L., Maisonnave E. and Seneviratne, S. I.,2016: Is land surface processes representation a possible weak link in current Regional Climate Models ? , *Environ. Res. Lett.*, **11** 074027
- Dennis, J. M., Vertenstein, M., Worley, P. H., Mirin, A. A., Craig, A. P., Jacob, R., and

Mickelson, S., 2012: Computational performance of ultra-high-resolution capability in the Community Earth System Model, Int. J. High Perf. Comp. Appl., 26, 5–16, doi:10.1177/1094342012436965

- Cheng, R., Casulli, V., Gartner, J., 1993. Tidal, Residual, Intertidal Mudflat (TRIM) Model and its Applications to San Francisco Bay, California. Estuarine, Coastal and Shelf Science, 36, 235–280
- Jin, M., Deal, C. J., Wang, J., Shin, K. H., Tanaka, N., Whitledge, T. E., Lee, S. H., and Gradinger, R. R., 2006. Controls of the landfast ice-ocean ecosystem offshore Barrow, Alaska. Ann. Glaciol., 44:63–72.
- Maykut G. A. and Untersteiner N., 1971. Some results from a time dependent thermodynamic model of sea ice. J. Geophys. Res., 76:1550–1575.
- Pham, Trang Van, J. Brauch, B. Früh, B. Ahrens, 2016: Simulation of snowbands in the Baltic Sea area with the coupled atmosphere-ocean-ice model COSMO-CLM/NEMO.*Met. Z*., DOI: 10.1127/metz/2016/0775
- Souverijns, N., Gossart, A., Demuzere, M., Lhermitte, S., Gorodetskaya, I., Van Lipzig, N., 2016: Evaluation of a default COSMO-CLM simulation over Antarctica with a focus on accumulation and the surface mass balance, Cosmo Assembly, Lüneberg, 20-23 September 2016
- Turner, A. K., Hunke, E. C., and Bitz, C. M., 2013. Two modes of sea-ice gravity drainage: a parameterization for large-scale modeling. J. Geophys. Res., 118:2279–2294, doi:10.1002/jgrc.20171.
- Umlauf, L., Burchard, H., 2005. Second–order turbulence models for geophysical boundary layers. A review of recent work. Cont. Shelf Res., 25, 795–827.
- Untersteiner, N., 1964. Calculations of temperature regime and heat budget of sea ice in the Central Arctic. J. Geophys. Res., 69:4755–4766.