Interactive comment on “Compact Modeling Framework v3.0 for high-resolution global ocean-ice-atmosphere models” by Vladimir V. Kalmykov et al.

Maxim Kaurkin

maksim.kaurkin@phystech.edu

We thank the reviewer for a detailed analysis of the article and valuable comments. The paper has undergone a major reformulation and revision of English wording, some sections have been extended. Two figures and one section have been added, so the numbering has changed.

The following is the point-by-point response to the reviewers' comments (shown in “italic”).

> ***Referee comment.***
>
> *A.1) The description of CMF2.0, results of Test I and II and related conclusions need to be clarified:*
>
> *- First, I agree with Referee #1 that figure 1 is misleading as it suggests a 1:1 connection between the models processes and the coupler processes. I don’t understand either the “which means data locality” on P.4, L.50, as there is certainly some exchange of data needed between the component processes and the coupler processes.*

Author's response.

Yes, we agree. The figure has been redrawn to show the basic idea that every coupler core communicates with a fixed subset of each component's cores. For example, on the Figure 1 the 1st core of the coupler (c1) sends and receives data from the component cores o1, o4, o7, i1, a1 and a4. Originally, we embedded this meaning in the term “locality”, but indeed it can be misleading. So, the paragraph has been extended and reformulated in “master-slave” terminology without a reference to locality.

> ***Referee comment.***
>
> *- P.5, L.63-66: these two paragraphs are not clear at all. What does “a subset of component’s cores works only with individual master core in the coupler” mean? What does Â n ´ for the two cases, of the source and destination type Â ż mean and why do you put a reference to Craig et al. 2005 here?*

Author's response.

These paragraphs have been reformulated. We are explaining the “master and its subset'” concept in the Section 3.2 (see previous comment), while these paragraphs are devoted to handling interpolation weights. The “source” and “destination” mapping types are briefly explained, while the thorough explanation can be found in (Craig et al., 2005, Section 3.4).

> ***Referee comment.***

- P.5., L. 69-70: You write "All necessary links are initialized at the beginning of run and are used at the calculation stage as persistent (Jacob et al., 2005). Â¨z I suppose that the links you mention here are not the SCRIP links? You should be using another word as this is confusing. Also, why do you mention Jacob et al. 2005 here?

Author's response.

Yes, it is better to say "component-component SCRIP links and intracoupler rearrange routes". The paragraph has been reformulated. It refers to Section 3.3.2 of (Jacob et al. 2005), which can be accessed for a more detailed explanation of the rearranging concept.

Referee comment.

- P.5, L.72-73: I don't understand this sentence: "It is worth noting, that links are not sent directly, but as sorted unique cells vectors which allow one to avoid sending duplicated data. Â¨z. Again what "links" are you talking about here?

Author's response.

It is just the data, which is being interpolated. The paragraph has been corrected.

Referee comment.

In Figure 2, 3 & 4 captions and on P.6, L.88, you should recall what is included in the timing. I suppose this is the time in seconds for the whole 10 model days for the whole ocean-atmosphere and atmosphere-ocean exchanges through the coupler. (Same remark for Figure 6).

Author's response.

That is correct. Figures 2,3,4 and Figure 8 (previously 6) show the timing of 10-day model runs with disabled physics routines (see Section 3.3, 6th paragraph for the Test I description, and 9th paragraph for the description of Test II).

The corresponding remarks have been added to figure captions and to the mentioned place in the text.

Referee comment.

On p.6, L.92, you write: "It is clear that 20-40 coupler cores provide satisfactory speed for such problems, because ∼10 seconds costs . . ." It is not clear how you get these numbers. I see at best, i.e. with the MVS-10p_16, something between ∼15 and ∼20 seconds. Same remark for the number presented on P.7, L.4-5. This re-joins referee#1's comment about the fact that the log scale does not allow one to get the detailed information mentioned in the text.

Author's response.

In our opinion, the logarithmic scale with time vs. number of cores is preferable for the analysis of parallel efficiency, because it allows simultaneously to compare with linear trend (which illustrates "perfect scaling") and to know the time of execution (in contrast to a speed-up plot or a non-

logarithmic time-cores plot). For convenience of estimations, we have added projection lines to the plots on the Figures 2 and 8. For example, on Figure 2, 32 cores of MVS-10p_16 under CMF2.0 give 10.5 seconds. The data of Figure 4 additionally has been presented in a table.

Referee comment.

P.7, L.6: again here I don't understand why you write "perform only local communication"

Author's response.

Since the master-slave algorithm was described above, this paragraph has gained a minor reformulation, which removes the reference to locality:

"Since every coupler core communicates only with a subset of

component cores, increasing of the coupler communicator size leads both to decomposing of the interpolation computations

and to decreasing of the component-coupler communication overhead, though slightly increasing intracoupler rearrangement

communications."

Referee comment.

Section 3.4 on I/O should be completely revised. The long theoretical section on P.8 with detailed formula is not useful here, especially as you finally simply state "Asynchronous scheme was incorporated in the latest version our framework Â˙z without giving finally any results! The theoretical section should be cut and numbers obtained for the grid sizes you list for INMIO World ocean model should be provided.

Author's response.

We have reduced and rewritten this section, so that it is now devoted to the CMF asynchronous scheme only. Data writing speed test results have been added for the CMF2.0 on the MVS-10P and BlueGene/P systems (new Figure 5). In Section 4.2 we have also added and discussed results of the actual asynchrony test of the CMF3.0 I/O system in conditions of the 0.1-degree global ocean model (new Figure 7).

Referee comment.

The different utility modules available should be briefly described or a reference to a documentation or User Guide should be provided.

Author's response.

We have provided the User Guide, which includes the description of the interface, utility modules and capabilities of the CMF3.0 system, as well as the process of installing and configuring the coupled model, as part of which the CMF is distributed. This manual is available at http://model.ocean.ru:6623/VITIM-manual-eng.pdf

Referee comment.

A.2) The description of CMF3.0 needs to be extended

This paper is supposed to mainly describe CMF3.0, or at least this is what the title implies but very little is written about it. It looks like the author was in a hurry to finish the paper. More details should certainly be given on how the I/O service work (1st paragraph on P.11) and on the Data Assimilation part (currently only 2 lines, P.11, L8-9). These improved descriptions should be backed up with performance results (as is done for the interpolation).

Author's response.

We have added a new section, which describes the mechanism and opportunities of working with global arrays in CMF3.0 by means of the class Communicator_GA (currently, Section 4.3). The I/O service description also has been extended. Its performance test has been presented in the Section 4.2.

As for Data Assimilation, it is a large area of our research, which requires a separate publication. We suppose that the graph of parallel efficiency for the DAS service (as an important result for this work) and references to our published papers about the data assimilation problem are sufficient.

Referee comment.

Also the discussion of the interpolation results should be extended and detailed. How do the author get to 2-3 seconds per modelling day on 20-50 CPL cores? (This is mentioned also in the conclusions P.15, L.28.) Why is it expected that results would be worse than for CMF2.0 (and "worse" should not be used here because it implies that results for CMF2.0 are bad and that results for CMF3.0 are even worse)? Is it because of the shift from MPI to PGAS? Or because the tests were performed on a different platform?

Author's response.

For convenience, we have added projection lines to the Figure 8, thus they show 20-30 seconds per 10 model days. The decline in performance is expected due to the overhead of using GA-library (as an intermediate send/recv data representation) and due to deprecated MPI_SEND_INIT procedures in the CMF3.0. This is a sacrifice for the compact code representation and for convenience of adding new features (like Data assimilation or Nesting technology). We have extended the Section 4 in order to clarify these features of the CMF3.0 system.

We have used "less strong" instead of "worse".

Referee comment.

Also, in the conclusions, you write: "The key part of it, coupler, has a sufficiently small code size for such programs (about 5000 lines of code with unit tests) and is able to manage the main parallel problems of the coupled modeling - synchronization, regridding and I/O." I don't understand why you write that the coupler manages the I/O as this is not the case in CMF3.0.

Author's response.

This phrase must refer to the CMF2.0. It has been corrected.

Referee comment.

A.3) The whole text needs reviewing by a native English speaker. The style and wording needs revision as some sentences are simply not understandable (at least by me),

e.g.:

- P.3, L.16: "Unquestionable advantage of non-coupler design is the absence of interference in the user code": why "non-coupler design"? Also, with OASIS, there is some interference in the user code but the objective is to minimize it.

Author's response.

We mean that there is no coupling through a standalone coupler in OASIS3-MCT, so the user does not have to reorganize his code according to standard interfaces (e.g., as required for the cpl7 coupler). The phrase has been changed to "Unquestionable advantage of this non-coupler design is the minimization of interference in the user code, since there is no need to adapt it to interfaces required by a coupler."

Referee comment.

- P.5, L.76: "Several ping-pong tests were carried out for interpolation system using coupled ocean-atmosphere model."

Author's response.

The phrase has been refined for explanation of ping-pong test conditions: "The performance rate of the CMF2.0 interpolation system was evaluated in several "ping-pong" tests, in which the coupler was maintaining component-component exchanges of the INMIO-SLAV ocean-atmosphere model with disabled solvers of physics equations (similarly to the ping-pong test of OASIS3 in (Valcke, 2013))".

Referee comment.

P.6, L.95-96: "work of the sequential algorithm is only possible with restriction that memory is allocated only for interpolation block, which is impossible in practice Â˙z: I am not sure what this sentence means exactly and why you write that it is impossible in practice while you do get some results on 1 core; do you mean it would be impossible with real models as the interpolation per se would require all the available memory?

Author's response.

Yes, to perform this test we had to switch off the allocation of all physical model arrays, except those particularly involved in the test. So, in real numerical experiments the node memory (at least on considered supercomputers) will be insufficient for both physics equations solving and work of the 1-core coupler. Possibly it is better to say "unlikely" instead of "impossible". The phrase has been refined.

Referee comment.

- P.13, L.63: “but we are more interested in scalability of the program on perspective sizes of computational resources Â˙z

Author's response.

The phrase has been reformulated indicating the saturation of the decomposition algorithm:

“Obviously, at high core counts the parallel efficiency curve experiences “flattering”. But assuming that the time step of the model is 5 min., the result of the experiment leads, e.g., to quite satisfactory five simulated years per wall-clock day (SYPD) rate achieved on 20000 cores of the BlueGene/Q supercomputer.”

Referee comment.

P. 13, L.71: “Time evolution of the sea-ice surface temperature is described in the same way as in prescribed ocean experiments.”:

Author's response.

The sentence has been removed.

Referee comment.

P. 13, last paragraph: please rephrase, the current sentence with the “min.” is too difficult to follow.

Author's response.

The sentence has been rephrased and split in two sentences:

“Every

72 min., nine 2D-arrays were transferred from the atmosphere to the ocean (components of wind stress, short- and long-wave

radiation, fluxes of sensible and latent heat, precipitation, evaporation, air temperature at 2 m). Conversely, every 144 min. three

2D-arrays were transferred from the ocean to the atmosphere (upper gridbox temperature, temperature and concentration of sea ice).”

Referee comment.

P.2, L25: “Coupling through shared file or sequential component is acceptable . . . Â˙z could be Â n ´ Coupling through shared file with components executing sequentially is acceptable . . . Â˙z

Author's response.

We imply slightly different meaning. Changed to “Coupling through a shared file or through a

sequential hub is acceptable...”

Referee comment.

P.2, L33: “. . . and their representation in the interfaced style understandable . . .” could be “ . . . and their adaptation to the interface understandable . . .”

Author's response.

Changed to “This approach requires some reorganization of the components' code and its adaptation to the interfaces understandable by the driver...”

(since we are talking about a set of standard interfaces through which all model procedures should be called).

Referee comment.

- P.2, L41: “GFDL FMS (Balaji, 2012) system additionally suggests fully parallel data storage with file post processing at the end of the run Â˙z could be Â nIn the GFDL FMS (Balaji, 2012) system, fully parallel data storage with file post processing at the end of the run is offered Â˙z

- P.3, L.21: “According to proposals of Earth System Modeling conference, (Valcke et al., 2012), . . .” could be “According to the analysis of coupling technologies for Earth System Modeling by Valcke et al. 2012, . . .”

- P.6, L86: “Performance is based on a standard Intel Fortran compiler” should be moved to the next paragraph and could be “On all supercomputers, the coupled system was compiled standard Intel Fortran compiler.”

- P.6, L.90: “Increasing number of coupler size”; should be “Increasing number of coupler processes” or “Increasing the size of the coupler communicator”

- P. 11, L. 12: “Optimizations regarded to ignore repeated cell requests are preserved” could be “Optimization regarding repeated cells are preserved”.

Author's response.

These issues have been corrected according to the reviewer's suggestions.

Referee comment.

- P. 13, L.56: “Latest version of INMIOWOM model was fully integrated to CMF Â˙z could be Â n ´ CMF2.0 was fully integrated in the latest version of INMIOWOM Â ż

Author's response.

Changed to “The latest version of INMIO WOM is distibuted in an integrated package together with the CMF2.0 and 3.0, all necessary libraries and a standardized folder structure facilitating the

adding of new model components (including adapter files for the CICE sea-ice model).”

Referee comment.

- P.14, 3rd line: “Ice model was built into the ocean model , land model – into the atmosphere model” could be “The ice model is integrated in the ocean model and the land model in the atmosphere model”

Author's response.

Changed to “The sea ice was simulated by the INMIO built-in ice thermodynamics model, while the land processes were incorporated into the SLAV atmosphere model.”

Referee comment.

- P.14, L.6: “structurize” could be “structure”

Referee comment.

B) Other comments:

B.1) References:

- P.2, L.6: The reference to OASIS3 should be Valcke 2013 (i.e. p18, L45)

- P.2, L.36: The reference to OASIS3-MCT should be Craig et al 2017: A. Craig, S. Valcke, L. Coquart, 2017: Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0, Geosci. Model Dev., 10, 3297-3308, https://doi.org/10.5194/gmd-10-3297-2017, 2017.

- P.3, L.7 (as annotated in the manuscript): Reference to ESMF should be the more recent one: Theurich, G., Deluca, C., Campbell, T., Liu, F., Saint, K., Verten- stein, M., Chen, J., Oehmke, R., Doyle, J., Whitcomb, T., Wall- craft, A., Iredell, M., Black, T., Da Silva, A. M., Clune, T., Fer- raro, R., Li, P., Kelley, M., Aleinov, I., Balaji, V., Zadeh, N., Ja- cob, R., Kirtman, B., Giraldo, F., McCarren, D., Sandgathe, S., Peckham, S., and Dunlap IV, R.: The Earth System Prediction Suite: Toward a Coordinated U.S. Modeling Capability, B. Am. Meteor. Soc., 97, 1229–1247, https://doi.org/10.1175/BAMS-D- 14-00164.1, 2016.

Author's response.

These issues have been corrected according to the reviewer's suggestions.

Referee comment.

B.2) P.2, L.35-36: OASIS3-MCT proposes coupled system not only as single executable, so this sentence is misleading; the important feature is that the coupling functions are not provided by a standalone coupler but by a coupling library linked to the component models. Please correct.

Author's response.

The paragraph has been rewritten as:

“The coupled system can be launched as a single or multiple executable without a separate coupler whose functions in this case are provided by by a coupling library and performed in parallel on a core subset of each model component. Such a solution was proposed in OASIS3-MCT (Craig et al., 2017). A high-level driver controlling system sequencing is not required in this case.

Referee comment.

B.3) P.2, L.44-50: This list mixes functionalities (1. and 2.) and characteristics (3. and 4.); please reorganise.

Author's response.

Sentence reworked.

“Thus, we can point out the necessary features of modern coupling frameworks:”

Referee comment.

B.4) P.3, L.11: It is not really fair to write that the computational costs of CESM coupler are quite significant 20%, as the CESM coupler does not only perform coupling and remapping but also performs the surface flux computation.

Author's response.

The phrase has been changed to “Tests showed that computational costs of the CESM coupler (including coupling, remapping and surface flux computation) are quite significant 20%...”

Referee comment.

B.5) P.3, L. 21-24: these 3 lines do not give an appropriate summary of the analysis provided by Valcke et al. 2012. Please correct.

Author's response.

The paragraph has been rewritten as:

“According to proposals of the Earth System Modeling conference (Valcke et al., 2012), today there are several common aspects in coupling software development: an ability to communicate data between components, regrid data, and manage the time evolution of the model integration. There is a lot of custom parallel coupling mechanisms, with either single or multiple executable approache. We selected the approach with single executable because it can simplify the program flow and give additional opportunities for performance optimization. Besides, we used NetCDF for parallel I/O and SCRIP (Spherical Coordinate Remapping and Interpolation Package) (Jones, 1999) for regridding, as done in OASIS3 (Valcke, 2013).”

Referee comment.

B.6) P.4, L.42-43 & P.9, L.47: Either provide details on what interceptor or Template methods are or don't mention them; one should not have to read the reference (Gamma et al., 1995) to understand the sentence.

Author's response.

The reference to interceptor methods has been removed. Section 3.1 has got some reformulations to clarify the meaning of Template methods.

Referee comment.

B.7) P.5, L77: References here are misleading as they seem to imply that "Test I condition" are explicitly defined in Valcke et al., 2012 or in Craig et al. 2012, while they are not.

Author's response.

The references have been changed to (Valcke et al., 2013) where the ping-pong test is explicitly defined.

Referee comment.

B.8) P.6, L.93: You should not use the word "failure" here as the test does not fail, it is just very slow.

Author's response.

"Failure" changed to "ineffectiveness".

Referee comment.

B.9) Using very technical coding terms along the text does not help understanding it (e.g. P.4, L.45: "Component class"; P.9, L.63:"resulted in class Communicator"; p.10, L.77-78:"since all services in CMF3.0 inherit base class Service it also allows one to easily add new Â˙z ; P.10, L.80 : receives data using Communicator Â ż ; P.12, L.25 : Â n ´ NormalEvent Â ż or Â n SyncVarEvent Â z : P.12, L.27 : Â nGenerators realize abstract class EventGenerator, so new specific generator subclasses could be easily added); I think it would be better to explain the concept that using these abstract terms.

We have extended and reformulated the explanation of these methods. Some terms have been excluded (Component class, EventGenerator). The others (class Communicator_GA, class Service, NormalEvent, SyncVarEvent) are kept, since they are names of the objects described and are needed as references in the text.

Referee comment.

B.10) P.12, L.26: Reference to Griffies et al. is not useful here.

Author's response.

Instead of this reference we pay attention to importance of supporting experiments with prescribed forcing referenced to the real calendar, e.g. the Drakkar Forcing Set (Dussin et al., 2016)

> ***Referee comment.***
>
> *B.11) P.12, section 5.1: What is the resolution of the INMIO World Ocean model in these tests?*

Author's response.

It is 0.1 degrees (added to the last paragraph of the section and to the Fig. 9 caption)

> ***Referee comment.***
>
> *B.12) P.15, L.25-27: please specify if these numbers apply to CMF2.0 or CMF3.0. On line 27, change "CPL3.0" for "CMF3.0"!*

Author's response.

They apply to CMF2.0. These issues have been corrected.

> ***Referee comment.***
>
> *B.13) In general, I think the section 6 on Conclusions and future work could be fleshed out.*

Author's response.

In our opinion, it is necessary to summarize the results. We tried to shorten and concrete this section.

> ***Referee comment.***
>
> *B.14) I think the "Code availability" section is not satisfactory regarding GMD standard but I will let the Topical Editor decide on this point.*

Author's response.

We can change this section if it is required.

> ***Referee comment.***
>
> *C) Minor comments:*
>
> *C.1) In the abstract, you write "As addition a parallel realisation of the EnOI (Ensemble Optimal Interpolation) data assimilation method as program service of CMF3.0 is presented." but this is not the only example presented in section 5.*

Author's response.

Corrected.

“As an addition, some information about the parallel realization of the EnOI (Ensemble Optimal Interpolation) data assimilation method and the nesting technology, as program services of the CMF3.0, are presented.”

Referee comment.

C.2) P.3, L.13: It is not right to write that OASIS3 is the most popular version of OASIS as most groups are using OASIS3-MCT today.

Author's response.

The sentence has been changed to “The OASIS3 system was very successful and was widely used by many research groups around the world.”

Referee comment.

C.3) P.3, L.16: In OASIS3-MCT, MCT procedures are executed on all component model cores

Author's response.

Corrected:

“The new version, OASIS3-MCT (Craig et al., 2017) resolves the issue of sequential interpolation by using MCT procedures executed on all model component cores, instead of mapping through a standalone coupler.”

Referee comment.

C.4) P.4, L.33: Define SOA the first time it appears in the text

Author's response.

Done (in the last paragraph of Section 2)

Referee comment.

C.5) P.2, L53: Define WOM the first time it appears in the text.

Author's response. Done (in the beginning of Section 2)