Geoscientific
Model Development
Discussions

Open Access

# *Interactive comment on* "Enhancing reproducibility of numerical simulation result on the C-Coupler platform" *by* L. Liu et al.

**L. Liu et al.**

liuli-cess@tsinghua.edu.cn

Received and published: 1 September 2014

We thank the reviewer a lot for all comments and suggestions.

We started to design and develop the C-Coupler at the beginning of 2010. Before the end of 2011, we only focused on the coupling functions. Considering the importance of a model platform (a configuring, compiling and runtime framework) for model development, we began to design and develop the C-Coupler platform at the end of 2011. As we were aware of that existing works rarely concerned the bit-identical reproducibility at that time, how to support bit-identical reproducibility of simulations became a new focus in the C-Coupler platform development. When developing the C-Coupler platform, we carefully studied the CCSM3 model platform, which does not satisfactorily address bit-identical reproducibility. This is because we joined the FGOALS-g2 team

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

for CMIP5 and the FGOALS-g2 runs on the CCSM3 model platform. During developing the C-Coupler platform, we did not pay sufficient attention to other ongoing projects for model platforms. We are glad to see that there are emerging works related to this topic. We are sorry that many state-of-the-art related works have not been cited and introduced in this manuscript. We thank the reviewer a lot for listing out the related works. This problem will be corrected when revising even re-writing the manuscript.

Guided by the papers given by the reviewer, we want to download the FRE (the FMS Runtime Environment) by GFDL and the model platform in GENIE, for further careful study, in order to know how these platforms benefit from our experience. However, it seems that these two model platforms are not publicly available. For example, it seems that there is no link on the GFDL website to for downloading the FRE the model versions that can be downloaded do not include the FRE. Regarding GENIE, we failed to access its website (http://www.genie.ac.uk/). We really hope the reviewer would like to help us for downloading them.

In the following, we'd like to reply to the reviewer point by point.

1.1 Metadata

Response: the information recorded by the C-Coupler platform for bit-identical reproducibility can be viewed as one kind of metadata for model simulation. We will discuss about the metadata for earth system modelling when revising the manuscript. Regarding the library versions, the model versions on the C-Coupler platform do not use the third-party libraries that can impact the bit-identical reproducibility. The third-party libraries currently used include NETCDF and MPI. In our opinion, results should keep bit-identical when changing MPI library version, especially when bit-identical parallelization is achieved. When the third-party libraries may impact the bit-identical reproducibility, it should be recorded. We will provide this function when there are this kind of model versions on the C-Coupler platform. Regarding the pre- and post-processing, we discuss about them as a future work in Section 7.

1.2 Software framework

1) What makes the authors decide to build their own software to address the missing pieces in existing solutions rather than contributing to international efforts, e.g. the ones described by Lawrence et al. (2012) and Guilyardi et al. (2013)?

Response: As mentioned above, when we started to develop the C-Coupler platform with the enhancement of technical reproducibility, we were not aware of the new international efforts about model platform such as GENIE. We knew that FRE is a good model platform. However, it seems not open for public use. We concerned more about the CCSM3 platform because almost all model versions developed in China for CMIP5 use the CCSM3 platform. We really want to contribute to international efforts. Currently, besides publishing paper and releasing the C-Coupler platform with implementations, experiences and suggestions, we do not know other ways to contributing to international efforts.

2) Ford et al. (2012) and references therein give an overview about existing tools for configuring, building and running models. Again, in the introduction the authors miss to provide a brief overview about the state of the art, and thus fail to derive from such an analysis those aspects that make their approach unique compared to existing solutions.

Response: We will introduce and compare to existing solutions to show the unique features in the C-Coupler platform.

3) Using GIT and/or SVN in combination with log files for archiving does not sound very innovative. For model code this is state-of-the-art since many years, and I wonder whether this is very suited for the versioning of data. Clearly, a discussion is missing why the authors follow this approach rather than using some kind of data base which allows the query of information regarding a particular model run or experiment setup. It is tempting to defer a change in software solutions and strategies to forthcoming papers, but I would not accept this as an argument for the choice made here.

Response: The database will be more powerful for recording the version of the data because it can help query the information of the data. We do not use database in our current implementation because using SVN seems much simpler for us to implement the C-Coupler. China currently does not pay much support for software engineering for earth system modelling. We have a very limited funding and less than 3 persons (not full time) are involved in the development of the C-Coupler. We thought about the database before and will use it in the future version of the C-Coupler platform.

4) In two paragraphs the authors mention the use of checksum to identify data and discuss problems associated with this approach. Again, this is not very innovative. The whole concept is well documented on the web and published elsewhere. References are missing.

Response: According to the papers provided by the reviewer, the approach of checksum for input data files has been used by the FRE. We will reference this work when revising the paper.

1.3 Code development

Response: We will reference Clune and Rood (2011) and Easterbrook and Johns (2009) when revising the manuscript.

1.4 Bit-identical results

1) While the ability to get bit-identical results during the software development process is a very useful concept when working on technical code development I have the impression that this aspect is overemphasised in this paper. Climate scientist are usually interested in guaranteeing that a particular model is capable of producing an identical climate when run at a different site or in different hard- and software environments. In other words: the climate should be independent from the parallelisation, compiler option and hardware. As single realisations are only of little help here this reasoning brings us to ensemble members. Thus, it would be interesting for the reader to know to

which extent the C-Coupler platform is capable of supporting such ensemble runs and in particular to evaluate the ensemble member results.

Response: Song et al., 2012 shows that the ensemble result with tens of members can be independent from the round-off error introduced by different parallel settings. However, we are not aware of that the simulation results from most scientific papers are from ensemble runs and whether the scientific conclusion is independent from the parallelization, compiler option and hardware has been clearly stated in the papers. That's one reason why we highlight the bit-identical reproducibility. Regarding the C-Coupler, it will support ensemble run in the near future.

2) As Li at. explicitly talk about CCSM3, I wonder why they do not mention the CCSM web site http://www.cesm.ucar.edu/models/ccsm3.0/ccsm/doc/UsersGuide/UsersGuide/node9.html where many aspects of bit-reproducibility are explicitly mentioned.

Response: We know that the CCSM3 can achieve bit-identical reproducibility in restart run, while the whole coupled model cannot achieve bit-identical parallelization and the CCSM3 model platform does not provide sufficient support for bit-identical reproducibility of simulation result. We will introduce that when revising the manuscript.

3) Furthermore, I found a presentation by Lapillone, Lardelli, and Fuhrer "Technical Test Suite for COSMO", held at the 14th COSMO General Meeting in 2012 At other sites in Europe this concept is used for almost two decades in the context of technical model development (e.g. paralellisation), and in my oppinion this is rather common practive than an innovative concept.

Response: We also think that bit-identical result is a common idea. The idea of bit-identical compiler version set and bit-identical processor version set may be new for bit-identical test. We will cite this presentation if possible.

4) In the Li et al. paper it finally turns out that bit-reproducibility can only be achieved

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

under certain very restrictive conditions. The C-Coupler platform does not perform any miracles to help achieving this beyond what we can already get with version control of our software and input data plus recording compiler versions and options. This is a disappointing result after reading in total 625 lines of text.

Response: It is true that bit-identical results can be reproduced only under certain restrictive conditions. That's why we propose bit-identical parallelization and define the bit-identical compiler version set and bit-identical processor version set.

2.1 Introduction

1) The state-of-the-art is not captured

Response: The state-of-the-art work will be briefly introduced in Introduction.

2) P4430 L18-24: The list of papers describing numerical models looks somewhat arbitrary. What is the reason for selecting those and ignoring other programming efforts?

Response: we just want to cite several models here. We selected these models because we know them. In fact, other models can also be cited.

3) P4431 L1: I would say it is still the modelling groups and people that participate in MIPs with the models but not the models alone.

Response: we will modify the manuscript accordingly.

4) P4431 L10-L12: Scientific reproducibility is not about hindcasting the climate of the past. I understand it as the ability to reproduce any climate obtained with a particular model setup (including input files and parameter settings) on any other platform independent of the compiler and hardware. In which sense does this deviate from technical reproducibility?

Response: As mentioned above, we are not aware of that the simulation results from most scientific papers are from ensemble runs and whether the scientific conclusion is independent from the parallelization, compiler option and hardware has been clearly

stated in the papers.

5) P4432 L1 - P4433 L16: This whole discussion is not new. Those aspects have been discussed at several workshops and are not new to any group involved in the development of a numerical model.

Response: We will try our best to reference the related works when revising the manuscript.

6) P4433 L4-L9: see also Easterbrook and Johns (2009) and Clune and Rood (2011)

Response: We will try our best to reference the related works when revising the manuscript.

7) P4434 L1-L22: Mentioning the C-Coupler is confusing and the relation to the C-Coupler platform is not really clear.

Response: We will clarify that when revising the manuscript.

8) What is the main message you would like to get over? The whole paper should then be oriented along this main point.

Response: We will modify the manuscript accordingly.

2.2 Brief introduction to the C-Coupler1

1) The C-Coupler is already introduced in the companion paper. A reference should be sufficient.

Response: One reference will be removed.

2.3 Necessary information for archiving the technical reproducibility

1) bullets 1 to 6: All this is not new and I claim that most professional modelling groups and companies (in particular benchmarkers from HPC vendors) already follow this strategy. It should go into a class book rather than into a publication for a peer-reviewed journal.

2) Bullet 5: this is already described by Song et al. (2012), so why is it repeated here?

3) A concise but complete list would be much more helpful.

Response: we will modify this section accordingly. It may be simplified as a simple sentence and this chapter may be merged with the next chapter.

2.4 Enhancement of the technical reproducibility on the C-Coupler platform

1) P4440 L4-12: Wouldn't a directory structure be more appropriate?

Response: We will think about how to modify.

2) P4441 L11: not relevant for this topic

Response: It will be removed.

3) P4444 bullets 1 to 8: this is not really feasible. Users who do not have the appropriate hard- and software are lost. Mechanisms to guarantee that the "correct" climate is reproduce would be much more feasible. For further development on a local system and to check certain development steps for bit-reproducibility users can generate their own reference data set.

Response: We will improve this part according the suggestions.

4) P4445 L5: It is a bit weird that suggestions are made for buying a particular product to achieve intercomparison about model results. Why do the authors suggest Intel products in first place and in favour of gcc (gfortran)?

Response: We just want to use the Intel product to prove the idea of bit-identical compiler version set and bit-identical processor version set. In our team, we always use Intel compiler and rarely use the GNU compiler. In the near future, we will test about the GNU compiler.

2.5 Experiences and suggestions to the technical reproducibility

1) P4448 L10-16: Is this discussion relevant for the understanding of the main topic?

Response: We think this discussion is relevant to the main topic. We will show more details when revising the manuscript.

2) P4450 L5-L20: the topic covered here are already sufficiently discussed in previous sections

Response: We can simplify them when revising the manuscript.

3) P4451 L5-L18: I do not get the point here. Usually, model development goes beyond pure technical improvement. Changes in the dynamic kernel or new physics options will change the numerical results.

Response: We mean that bit-identical result of the same simulation in repeated runs. If the dynamic kernel or physical package has been changed, the simulation has also been changed.

4) P4452 L1-L10: This topic has been discussed earlier.

Response: We can simplify them when revising the manuscript.

2.6 Empirical evaluation

1) P4453 L19: Is it the C-Coupler platform that acheives the technical reproducibility, or isn't it rather the numerical model components if not the humans, e.g. the programmers?

Response: it should be modified as "We therefore can conclude that the C-Coupler platform can help scientists to achieve the technical reproducibility for model simulations".

2.7 Discussion and conclusion

1) P4454 L10-L11: What shall the motivation be for any scientist to use the CCoupler platform for her own development? I claim that most model development groups have implemented their own strategies for clean model development and they do it for

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

decades with quite some success (see Easterbrook and Johns, 2009 for an example). Of course there is always room for improvements, but it should be stated clearly where the added value is when using the C-Coupler platform.

Response: We will clearly state the potential new values of the C-Coupler platform when revising the manuscript.

2.8 Tables

1) What shall I do if I do not have any such system? I cannot ask my compute centre to buy one.

Response: Currently, we only have Intel processors for evaluation. In future, we will have more kinds of processors for evaluation, such as AMD processors. For the system the reviewer currently uses, we believe that the bit-identical compiler version set and bit-identical processor version set can be found. We'd like to provide the systems used in the evaluation, if the reviewer want to repeat our empirically evaluation.

2.9 Figures

1) Figure 1: What is the purpose of the colours? Simple b/w would be sufficient.

2) Figure 2: What is the purpose of using a coloured figure here. Instead of using a figure a bullet list or simple table could cary the same information?

Response: We use colors in order to make it beautiful. They can be simplified as b/w. Figure 2 can be changed into a bullet list.

In this reply, we further give a package (Package for checking reproducibility) as a supplement. This package contains the script checkout_experiment, an experimental setting package GAMIL2-sole.time_slice_from_1974.config.20140901-143902.checksum, an environmental variable file local_env and a log file of model execution GAMIL2-sole.time_slice_from_1974.gamil.log.20140901-143902 for checking bit-identical reproducibility. Please follow Section 4.2 in the paper and the user guide of

the C-Coupler platform to bit-identically reproduce the corresponding simulation. When downloading for the simulation, please modify "DATAROOT" and "PLATFORMROOT" in the file local_env, and enter the username ("Public") and password ("publicopen") to the GIT server for downloading the model code. The processor version to run the simulation is Intel(R) Xeon(R) CPU E5-2670.

We gave a simulation of atmospheric model version GAMIL2 but not a simulation of FGOALS-gc or MASNUM-POM that has been used in empirical evaluation (Section 6) because we could not publicly distribute the model code without permissions from the corresponding model groups.

Please note that is package is slightly different from the package we gave when replying to the executive editor Dr. Dan Lunt. This is because some newly developed functions have been added into the C-Coupler platform for the technical reproducibility, which will be further introduced when revising the manuscript.

Please also note the supplement to this comment:
http://www.geosci-model-dev-discuss.net/7/C1552/2014/gmdd-7-C1552-2014-supplement.zip

---

Interactive comment on Geosci. Model Dev. Discuss., 7, 4429, 2014.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper