

Geosci. Model Dev. Discuss., author comment AC1
<https://doi.org/10.5194/gmd-2021-311-AC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Reiner Jung et al.

Author comment on "CP-DSL: Supporting Configuration and Parameter Selection of Ocean Models with UVic (2.9) and MITgcm (67w)" by Reiner Jung et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-311-AC1>, 2022

Thanks for your valuable feedback to our submitted paper!

We respond to the comments that require changes.

Your comment:

- "As such, I think it is an approach that has value although I feel that the paper itself could do with making this case more strongly."

Reply:

- We will add comments from domain experts to highlight the advantages of the approach.
See also the following reply.

Your comment:

- "Related to this, although there is a discussion of the evaluation of the features and use of CP-DSL, there is no mention of what the model developers themselves think – are they keen to adopt CP-DSL or do they have reservations?"

Reply:

- We presented the CP-DSL to research software engineers and scientific modelers at GEOMAR, and they are (1) very much interested in a DSL in their domain of ocean system modeling and (2) are interested in using the DSL. However, their responses only apply to the currently adapted case studies with UVic and MITgcm (as our paper title indicates).

Your comment:

- "It seems that configuration/control of diagnostic outputs is at an early stage in CP-DSL but this is a critical and complex part of production jobs. I can see value in a common way of specifying diagnostics as this extends beyond ocean modelling: there is a relatively small number of IO systems and these tend to be common between e.g. atmosphere and ocean models."

Reply:

- The diagnostics configuration in the DSL is indeed in an early stage and primarily motivated by the diagnostics features present in MITgcm. Its inclusion in the current DSL helps to better engage in a follow-up discussion on how to specify this aspect. While it is possible to specify logging based on the parameter groups and module structure of the CP-DSL, we aim to provide specific structures for diagnostics and logging that allow

users to be concise when specifying diagnostics and instruct them safely through specific DSL features.

We are aware of XIOS and other logging and diagnostics facilities for scientific models (cf. <https://www.esiwace.eu/services/software-support/supXIOS>). However, XIOS is seen by our interviewees as complicated and not applicable to all models. Our DSL could – with a suitable template for XIOS configuration generation -- support XIOS configuration files. This is future work as mentioned in the paper.

Your comment:

- "Although the discussion of the various roles played in the development of ocean models in Section 3.1 is interesting, I don't think it adds any value to the paper (which is primarily about the new DSL) and could be removed."

Reply:

- The roles are important to understand the processes and thus determine the requirements for the DSL, which in turn leads to the design of the DSL. They are also relevant to understand which DSL addresses which role, i.e., the Declaration and Template specifications target the research software engineer, while the Configuration specification addresses the needs of the scientific modeler or model user. We will make this more explicit in the paper.

Your comment:

- "Figure 2 and its accompanying text mention that there are options that are common between ocean models. Given the subtleties that can occur when different scientists implement the same numerical scheme I think that determining such options could be problematic. I think some examples of such options would be helpful here. Is there a need for an agreed set of named quantities?"

Reply:

- The text for Figure 2 was, indeed, misleading, sorry. The Declaration and Template models are specific to one scientific model, while the Configuration model is specific to one scientific model setup. We have updated the text and illustration accordingly to clarify this.

Your comment:

- "The UK Met Office uses 'Rose' for job configuration (see <https://metomi.github.io/rose/doc/html/tutorial/rose/index.html#rose-tutorial>) which has some similarities with the approach described in this paper. It is likely that other meteorological centres have similar configuration systems."

Reply:

- Thanks for the pointer. Rose' is simpler than CP-DSL, e.g. only name-value pairs and

sections are used. Whereas CP-DSL uses higher-level concepts like groups of parameters and features with dependencies. However, we see 'Rose' as an important related work and now include it accordingly in Section 3.3.

Your comment:

- "PSyclone is developed by the UK Science and Technology Facilities Council's Hartree Centre in collaboration with the UK Met Office and the Australian Bureau of Meteorology. PSyclone has two 'modes' of operation: as an internal DSL (as used for the UK Met Office's LFRic atmosphere model) and as a code transformation tool (as used with the full NEMO ocean model)."

Reply:

- We agree that PSyclone provides both applications and mention its use of an internal language as well as the extensions to Fortran code and its ability to transform these via in-place code transformation. To further clarify this, we add the following text: "The use of code transformation makes it possible to both write code and use PSyclone for optimizations in existing code that uses appropriate code structures."

Your comment:

- "Not all models set grid sizes at compile time. This is a run-time option in NEMO for example."

Reply:

- Thanks for this additional information. It is correct that grid sizes can also be a dynamic namelist option for some models, still it can also be defined at compile time. Nevertheless, we agree with the point that there are also models that only allow to set grid sizes at runtime. Accordingly, we have replaced the example and pointed out that models may implement options differently. We added the following text: "However, not all parameters can be set at runtime as optional modules enable or disable complete parts of the earth system model, e.g. atmosphere model. Thus, they are set at compile time. For example, in some models mesh and grid sizes are set at compile time."

Your comment:

- "Deployment of models is, I believe, something that the CYLC workflow engine (<https://cylc.github.io/>) does. Please compare."

Reply:

- In our paper, we focus on configuration and parametrization. Model deployment is an important issue, but it will be addressed in our project (OceanDSL) with a separate deployment DSL (not covered by the present paper). Nevertheless, we appreciate the interesting reference and have included it in the section as follows: "These requirements are addressed by other DSLs and tools, like the CYCL workflow engine~\citep{oliver2018cylc}."

Your comment:

- "SVN is used as well as git (e.g. by NEMO)."

Reply:

- This is correct, but according to our interviewed experts SVN is outdated and mostly replaced with Git. However, we extended the text to Section 5 to emphasize this fact: "Some model projects, like NEMO~\citep{NEMO2015}, still use SVN."

Your comment:

- "What does "UVic is the reference simulation by GEOMAR of the University of Victoria model." mean?"

Reply:

- The GEOMAR Helmholtz Centre for Ocean Research Kiel uses a refined version of the UVic ESCM model developed by the University of Victoria. To make this point explicit, we moved the paragraph and added the following paragraph to Section 8: "The GEOMAR Helmholtz Centre for Ocean Research Kiel uses a refined version of the UVic ESCM model developed by the University of Victoria~\citep{keller2012new}."

Your comment:

- "A better reference for GOcean may be found at <http://nora.nerc.ac.uk/id/eprint/521162/>"

Reply:

- Thank you for the reference. We also consider this reference as appropriate and replaced the existing with the proposed one.