Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-103-RC1, 2018

Geoscientific Model Development Discussions

Interactive comment on "BGC-val: a model and grid independent python toolkit to evaluate marine biogeochemical models" *by* Lee de Mora et al.

Anonymous Referee #1

Received and published: 19 June 2018

general comments

The manuscript from De Mora and co-authors provides a detailed description of both technical and scientific aspects of a new framework, named BGC-val, to perform routine operations and multi-model analyses of marine physical and biogeochemical quantities. Beside the wide range of functions to deeply "dissect" model data, a key aspect is the standardised approach that makes this tool model and grid independent. BGC-val represents a good step forward to support not only a single model but also the analysis of the broad ensemble of data from the CMIP exercises. The rationale supporting the development of the tool is somehow misleading (see specific comments) and can be better tailored to model evaluation purposes rather than development/application. Moreover, the potential of this tool with respect to the existing ones should be better

framed. Overall, the BGC-val workflow is clearly described and an exhaustive set of examples is provided to illustrate its usage and degree of flexibility offered to the end-user. In particular, the different functionalities of the evaluation framework are thoroughly described and the outcomes are collected in a very user-friendly interface (as from the support material).

specific comments

1. The introductory sub-section 1.1 provide a long description on different issues, spanning from the degree of complexity in marine ecosystem models, the computational effort required to analyse ocean biogeochemistry in comparison with land system, and the possible strategies to tune ESM models against observational data. Given that the purpose of the section was to describe the ideas that motivated the development of the BGC-val framework, I would have expected instead a clear review of the existing literature to support the need of a new, flexible framework. I suggest the authors to revise this section by including a description of present state of the art tools in comparison to BGC-val and by remarkably resizing less relevant paragraphs related more to model development than their evaluation. In addition, the section will benefit from a short paragraph focusing on the observational datasets that are routinely used and ingested by BGC-val within the validation of UKESM1. This could provide a good starting point to foster the discussion within the modelling community toward the definition of a common framework also for data usage in comparison exercises.

2. As far that overall considerations of computationally cheap or expensive operations are addressed, I see the need for a technical description of BGC-val usage requirements or at least a description of its computational performance/skills on JASMIN system (used CPUs, memory requirements etc...) to allow the end-user determining beforehand if the tool can be deployed on its own system. I also suggest to revise the Section 3 as single paragraph by streamlining the text on the workflow (which already have a stepwise organization) and by removing redundant comments between the existing subsections. Finally, I guess that BGC-val may include some degree of

parallelism, if so, a brief description of where parallel instances/computations are performed within the workflow could be very helpful.

3. The proposed example of BGC-val summary report contains analysis also for physical ocean quantities, e.g. Drake passage flow and AMOC. Authors report that computations for these quantities were adapted from the "Ocean Assess" tool, which is unfortunately not available to the public and methodologies are not clearly referenced. For such a reason, I think it is preferable to avoid pointing/referencing at "Ocean Assess" methodologies (see section 1.2 and 5). Authors can instead provide some details on the computation methods for these metrics (if relevant to the manuscript) or improve the description of custom functions reported in section 5 (namely cmip5DrakePassage, cmip5AMOC).

4. Authors clearly state that they "will continue to develop and apply the toolkit" in the future which is quite an interesting and promising perspective. However this seems to stride with the development of ESMValTool that is also receiving contributions and support from authors (see section 1.2). At this stage, it seems reasonable to assume that ESMValTool will include in the next period several features from BGC-val. It looks to me that BGC-val will become a mirror of ESMValTool ocean analysis at some point, so could it be possible to converge into a single tool instead of having duplicated efforts?

technical corrections

P1.L9 : the expression "marine circulate and biogeochemical parts" isn't really clear, maybe authors meant something like marine physical and biogeochemical quantities.

P2.L20 : the term "sequesterer" might be suitably replaced with "sink"

P2.L28 : remove the comma at " range, (.."

P3.L13-14: revise text as " . . . any component of an Earth system model . . ." and check if commas before citations are always in the correct place.

P7.L9 : this sentence is not totally clear.

P7.L10 : I guess is " . . . the new masks can be defined in advance, . . . "

P8.L10 : it is a detail but actually in the supplementary material there is only one index file of example.

P8.L18 : " . . . stored in one or several .." P9.L9 : please revise as " ...it is much quicker to evaluate a 2D field ..." or "...it is much quicker the evaluation a 2D field ..."

P10.L11: "In contrast, the stage ..."

P10.L22 : "observational dataset for the model to match against. ..."

P18.L4: repeated "of"

P18.L15: i think it is not necessary to repeat here and in the following text the reference to "in the HadGEM2-ES model in the historical scenario, in the ensemble member, r1i1p1" already given in the initial description of figures 3-5.

P24.L14: I think that paragraph 4.2.4 can be easily integrated somewhere else in the text above or in the general description of the package structure. Accordingly modify the reference in P25.L8 ("and the bgcvaltools package is described in sect. 4.2.4.") and other instances.

P25.L12: "the conversion is one of the standard functions "

P27.L9 : Missing capital font at " For example, the ..."

P27.L19 : Please add a reference or a http address describing the JASMIN facility

P27.L31: here it is implicitly assumed that the reader knows about the JASMIN system, maybe some detail at least on the "mirror" could be helpful.

P28.L1-2 : Maybe some more hints on this error can be profitable to describe the effective benefits of routinely use BGC-val for model development and simulations.

P29.L7 : space is missing at " . . . between 2014 and 2013 (McCarthy et al . . . "

P28.L9 & P31.L2 : here authors refers to a methodology from the “Ocean Assess” tool that has no references.

P31.L34 : Please report here the details about ESACCI product, as given in the final notes at P34.L21

P32.L13 : there is a typo in the chla value reported to be “0.39.6” mg chl m-3.

P32.L14 : the sentence on data starting with “The CCI global monthly mean .. “ was already quoted at line 3 and should be removed from here.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-103, 2018.