Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-291-RC2, 2020

Geoscientific Model Development Discussions

Interactive comment on "ESMValTool v2.0 – Extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP" *by* Veronika Eyring et al.

Anonymous Referee #2

Received and published: 24 January 2020

General comments:

This manuscript entails a monumental effort in attempting to facilitate the development and evaluation of climate models. Examples for analysis reproducibility, particularly output figures from IPCC chapters is commendable. A pathway to expand this to output figures in the literature is also evident. Model performance metrics, diagnostics for the evaluation of processes in different realms are presented in great detail along with the corresponding recipes. Example figures as a result of integrating community metrics is also shown in the manuscript. The flow and the content could be more con-

sistent so the focus of ESMValTool goals and the impact in doing that is delivered as intended. Some level of brevity, citing references for details, providing more example figures from recipes, pointers to additional recipe documentation – should be made available through an external reference and/or supplementary material. Scalability and interoperability aspects can also be briefly touched upon, providing guidance to the community, making interoperability and practicality– a key to expanding the audience. There is scope for condensing and merging certain sections. Some key points to help improve readability is furnished below in specific comments.

Overall, thank you for the contributions. Please see more comments below.

Specific comments and technical corrections

Page1, Ln 58 Reproducibility - Specifics and explicit wording is required here, as to what aspect is targeted.

Page2, Ln 85

There may be more references that need to be cited while discussing data standardization for CMIP. E.g. https://pcmdi.llnl.gov/mips/cmip5/CMIP5_output_metadata_requirements.pdf ?

Page 3, Ln 92-83: The line about "full rewards of the effort. . .." should be reworded to provide more of a positive tone to the available observations and model output in standardized format. Expanding what is meant by "full rewards" will be very helpful in this line, rather than the subsequent paragraph.

Page2, Ln 96-97: Please cite or provide links to appropriate references w.r.t data volume estimations for CMIP. Also, what is the database that is being referred to here?

Page 2, Ln 100- I like the addition of "creativity" here.

Page 2, Ln 107- "that provides results. . .." - Substitute results with something more specific. E.g. analysis products/output?, so it better connects with Ln 108 (This is

realized through..)

Page 3, Ln 115- Does ESMValTool preserve the netCDF metadata (global attributes from input datasets) in output products? How is data provenance established? (Ln 142 may answer this, please clarify)

Page 3 Ln 118- Consider stating “Figure 1 from their paper, or from Righi et al.. rather than “their Figure 1”

Page 3, Ln 121- The flow from the introduction to companion papers and the present one can be better. Example- Precede the sentence “the use of the tool is demonstrated..” with “In the present paper,..”

Page 3, Ln 124-125, Avoid too many conjunctions (and) here. . “Diagnostics and performance metrics and the variables and observations used”.

Page 3, Ln 129: What does “partly also with CMIP3” mean here? Page 3, Ln 130: Is CF-compliance and CMOR-compliance required? Please, also cite CF and CMOR references, expand acronyms. The sentence could be changed to - tool is compatible with any CF and CMOR compliant model output? Please change this as needed so users understand what is ready to be plugged in to ESMValTool, and what requires additional work.

Page 3, Section 2. Ln 131-136. The data descriptions in this section are not satisfactory, especially where the manuscript reads “observation from other sources..” . obs4mip publications should be cited here. It will be nice if the different observation datasets used in recipes can be listed and cited thoroughly. Also, this section could be merged into the final section 6-7 on Code and Data availability.

Page 4, Ln 154- Reproducing IPCC chapter figures is impressive. Are these diagnostics-and-recipes written working directly with the IPCC authors? What is your advice to the IPCC authors to make this effort a success for CMIP6? How are the recipe names constructed- is there a recommended naming convention? How resilient

is ESMValTool to changes like the metadata conventions, DRS, etc from CMIP5 to CMIP6, or say another [inter]national assessment?

Page 4, Ln 160-161: Check and correct line,word spacing.

Page 4, Ln 164: How does one add an alternative observation dataset? One of the companion papers might be addressing this? Page 4, Ln 165: How can additional variables be added? Is it the same as the first version of the tool? Following the citation link here, I still could not get information in two hops. Page 5, Ln 174: Can there be a reference here to the regridding tools used? Why 4x5?

Section 3: Throughout the Overview of recipes, under each sub-section, there can be more consistency. Example: For each recipe, one could ensure these are specified throughout: input (include time-frequency requirement consistently as well), output specifications, source, purpose and significance of the metrics, relevant citations to metrics calculations, summary of the recipe, a sample result. Sticking to this consistently can also condense the text. Suggest just pointing to references like how it was done for CVDP to get more information metrics. [3.3.4] Sea Ice, for instance, can be rewritten to condense text. Are the recipes part of the github repositories? Where can one find them? Though line 145 reads that the intent of the focus of the manuscript is not an assessment of CMIP5 or CMIP6 models, the construction of section 3 is not completely aligned with this. The message needs to be reiterated. If this manuscript is intended to be a documentation paper for the diagnostics and recipes used in ESMValTool, the length could be justified to an extent. Otherwise, some sections could be rewritten so focus is retained. Is this manuscript the single source for documentation for all the metrics and recipes? Page 21, Section 4.1. Automatic execution of ESMValTool at DKRZ sounds like a nice step to interface with more users. How scalable is this process? Is the idea to expand this to other nodes in ESGF? Is data replication of such huge CMIP6 volumes something that needs to be kept an eye on, leveraging distributed data access protocols or the cloud?

Page 22: Ln 843-844: Section 4 When new plots are created, is there a step that incorporates a basic automated quality assurance conducted? Is there a testing suite for each recipe?

Ln 845: The result browser looks good. Steps to reproducing figures viewed from ESMValTool result browser should be made clearer. This is probably the place where the provenance information captured by ESMValTool will come handy?

How is the performance of running ESMValTool on CMIP data in an automated fashion, and in general from a disconnected sandbox, regardless of ESGF. How is the concept of data versioning incorporated within the automated generation of plots using ESMValTool in ESGF at DKRZ ? When there is bad data retracted on ESGF, and a newer version of data becomes available, what is the current implementation like at the ESMValTool-end or the result-browser to notify its users? If there is no mechanism to notify automatically or not-show-the-corresponding-plots, what is the recommendation to the users? In general, what kind of users does the ESMValTool aim to target?

Ln 857: How does the metadata w.r.t the software version get mapped to the actual source code in GitHub? With data DOIs/data citations widely prevalent for CMIP6, does ESMValTool automatically add data citations to the output figures/files? If not, please provide a pathway to achieve this. Page 24, Ln 943: Please provide an example for "preprocessor settings". Page 24, Ln 948-949. Unable to follow this line "...and tags (i.e. what is reported) ". I think these lines are not adding much value at this point. Page 24, Ln 959. Identifying errors in the simulations early on is a key factor that is penned down as future work here. Even if there are no web-based capabilities, please address if ESMValTool can independenly be installed and run by an individual user at different stages in model running. An idea or vision here to draw more inspiration and motivation for using ESMValTool can be provided.

Ln 955, Again, enhancing quality control is a great use-case, but having ESMValTool run on published data on ESGF does not satisfy this use-case. Stand-alone, this tool

seems to work towards QC. Please clarify.

Section 5. Font size seems to be mixed up in the Summary section - lines till 950 and after 950 are different.

Page 25, Ln 965, Sections 6 and 7 should be condensed into one section. Addressing data citations briefly in Data availability will add more value to the CMIP and ESMValTool efforts.

Comments on Figures:

Adding some of the figures to a supplementary or appendix should be considered. Verify that there is not much redundancy in the text in captions (e.g. Section 3, 5, Figures text). Avoid redundancy where possible.

I find the captions in figures helpful, especially the reference to the corresponding sections. The captions are mostly like IPCC-chapter and documentation paper style. A short caption in bold followed by the description is something that will make the figures stand out.

While specifying OBS in figures, please specify names of OBS in the figures.

Name the variables/fields corresponding to the figures, example Figure 4,5,6 - zonal wind,air temperature,precipitation? respectively.

How are the color palettes picked in general and what flexibility ESMValTool allows w.r.t color palettes?

Expand the acronym QBO in Figure 4, although the section covers it.

Better labeling on the figure itself needed for Fig 15,16 especially.

What is "j" in r1i1j1 in several figure captions - e.g. Fig 22.

In Fig 20, use r1i1p1 to be consistent, not r1p1i1.

Is Figure 26 a reproduction of Fig 9.14 from AR5, Chapter 9. (Including chapter helped

me since there is some ambiguity looking up for Fig 9.14 from AR5).

In Figure 30, “typo” - “ether” vs either;

In Figure 34- typo: predictand, not predictant.

Please use long names on the figure themselves, not the short CMOR names (example Figure 35, Y axis); Units missing in some of the figures, e.g Figure 39.

In the summary section– Given the challenges of CMIP6 (and beyond) and the scientists all over the globe working on multiple research areas, this manuscript should include something along the lines of the role and future of ESMValTool in the community as a whole and how it can be interoperable with overlapping efforts. The ability to cross-function using tools like ESMValTool and making them more inter-operable is a key challenge. The experience from developing ESMValTool in the form of these manuscripts is helpful to the community, and it can also be helpful for the expansion of metrics-and-recipes used in ESMValTool.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-291, 2019.