Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-60-RC2, 2020

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for emergent constraints and future projections from Earth system models in CMIP" *by* Axel Lauer et al.

Anonymous Referee #2

Received and published: 11 May 2020

This paper describes part of the functionality of the ESMval tool that can be used to evaluate and intercompare CMIP (and other) model data. Tools that automate part of the process of collocating and analysing data are immensely useful as they increase efficiency, avoid redundancy and minimise the risk of errors. Given the ever larger flow of (model) data these tools can rightfully be considered part of our modelling toolkit. GMD is an appropriate choice of journal for this paper. The paper has a clear structure and is well written although sometimes short on detail.

The paper states that its aim is "to document and illustrate [] these newly added

Printer-friendly version

Discussion paper

ESMValTool “recipes”. However, very little information is given to the user on how to use these ’recipes’ (do we need to set certain parameters? how does the code find the data? What requirements are there for the data, both models and observations?). Rather the paper seems more an advertisement than a technical document (see https://www.geosci-model-dev.net/9/3093/2016/gmd-9-3093-2016.pdf for an example of the latter). This is not necessarily bad but the paper does not make it easy for users to find the technical documents to obtain this information. As a side note: the readme for the tutorial on github is mostly unpopulated.

Furthermore, I miss discussions of following topics:

- a discussion of the spatio-temporal resolution of the model data and observations used by ESMvaltool. P 3, l 83 states that "any arbitrary model output" can be used. Is that really true? Can I use e.g. both yearly averaged data and hourly data?

- a discussion of the tool’s expectation when it comes to the format of observational data. Presumably these should be gridded.

- a description of how ESMvaltool deals with differences in the spatio-temporal resolutions between datasets. This is alluded to in a single Figure caption but should be clearly stated in the paper as part of the tool’s functionality. As a side note, it appears the authors believe that observational errors are all stand in the way of model evaluation but there are two other issues. These are 1) differences in spatio-temporal sampling of different datasets and the representativity issues that result (see e.g. https://www.atmos-chem-phys.net/17/9761/2017/ and the references therein); 2) appropriateness of the observation operator (i.e. the model’s code that generates a diagnostic that may be compared to observations, e.g. what definition of temperature is used?).

- a discussion of the underlying assumptions in the many regressions used by the tool. My guess is that one important assumption is that individual models can be viewed as independent data points which is unlikely given that often models share (part of) their

code base or at the very least incoporate similar ideas with regards to e.g. sub-grid parametrisations.

- a mention of the graphics formats produced by the tool and whether the user has any control over them.

Finally, I think there may be substantial mistakes in Sect. 3.4.2 that need to be addressed. In addition I found it lacked sufficient explanation.

Minor comments:

Should Table 1 maybe have more information on e.g. the temporal averaging in model data that is needed or do the scripts work with high-freuqency output and perform this averaging themselves?

p. 3, l. 80-85: Can the authors provide references (even if weblinks) for CF-complaincy and CMOR?

p. 3, l. 90: Apparently users can 'import' their own favoruite datasets and use them with ESMValtool. Can the authors provide a brief description of the steps necessary for this to work?

p 5, l 139,140: I do not know whether tas and rlut etc belong to CF-compliant or CMOR definitions but can the authors clarify, also where readers may find further definitions?

p 7, l 189: 'correlation of the covariance'. Shouldn't this just be 'correlation'?

p 13, Sec. 3.4.2: I suggest there is something wrong with either the equation or the definitions here. When alpha=1, the eps_i,m would be drawn from a distribution with imaginary (!) standard deviation (unless beta=0).

There are numerous other issues with this section: - x_i is (probably) not an observation but an anomaly (y has mean 0).

- How is the 'mean correlation between a series of values (x_i,1..M) and a single value

Printer-friendly version

Discussion paper

(y_i) defined?

- What is the meaning of eps_i,m? Note: it is also called eps_i sometimes (l. 383), please correct this.

- I'm not familar with the work by Weigel but it seems odd to call alpha the predictability. Don't the random errors eps control the predictability? Probably I misunderstand something but it appears that, beyond the correction of aforementioned errors, this section needs much more explanation.

- the purpose of the toy model is not really explained. I guess it allows the user to put an error estimate on the uncertainty of observations used in emergent constraints etc? Can the user apply this toy model to every constraint or are there limitations? What underlying assumption feed into this toy model? Independent and Identically randomly distributed errors is probably a major assumption and needs to be written down explicitly!

- Toy model may be a confusing choice of word, as the ESMval tool is all about model evaluation. Maybe uncertainty simulator (or estimator) would be a better choice?

p 14, l 409: "including stippling and hatching to indicate significant changes and areas where models do not agree" I found this sentence onfusing. It suggests that stippling/hatching is used to indicate where models do not agree but the caption to Fig 18 states otherwise. Elsewhere in the paper stippling/hatching is used to indicate agreement as well.

p 14, l 412: "where the projections are still uncertain (hatching)." It appears that the use of hatching is quite inconsistent. I understand that the authors are trying to recreate figures found in a large number of papers that are unlikely to be consistent. Maybe this is something to note in the summary or elsewhere, e.g. a 'buyer be ware' clause. After all the authors provide a single tool to generate figures that will be assumed by most users to be consistent in their definitions.

p 15, l 444-446: this explanation of other papers regarding ESMval should be part of the introduction, in my opinion. I would also suggest to add more detail: as a user I want to know which paper to use to find what information.

p 15, Sect 4: I suggest removing the names of recipes. The serve no purpose in this summary.

p 15, Sect 4: The summary should contain a brief mention of data requirements and limitations of the tool. As it stands it is a brief rehashing of the the list of emergent constraints and nothing more.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-60, 2020.

Printer-friendly version

Discussion paper