Interactive comment on "Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for extreme events, regional and impact evaluation and analysis of Earth system models in CMIP" *by* Katja Weigel et al.

Anonymous Referee #2

Received and published: 23 November 2020

Overall comment:

The paper gives an overview on the new implemented data analysis and diagnostic recipes in the version 2 of ESMValTool. The prerequisite for using the open source tool is that the data not only follow the CF conventions for the file structures in netCDF but also the CMOR standard of CMIP is applied. To broaden the usability of the tool, it would be good to base the work on the given standard_names of variables and using an internal dictionary to translate variable names to the CMOR standard. That would avoid reformatting of used datasets depending on the software used (in his case

ESMValTool).

The quality of the paper in terms of structure and language is high but sometimes there is a lack of correctness and attention, which leads to several minor comments. The high number of comments leads to the overall judgement 'major revisions'.

The comments are ordered by the occurrence in the text and where the comment relates not only to the text but to the figure as well the figure-comment is brought forward.

Content:

P4L96: "Observations used in the evaluation are described in the following section' – this is not the case (describing is perhaps not necessary but mentioning them)

P5L136: ". . .hydroclimatic intensity index . . .which is a measure of the overall behavior of the hydroclimatic cycle"– this is not true; it is a measure of change of hydroclimatic intensity. As the hydroclimatic intensity itself is normalized, the values are not comparable in space.

P28Figure1: The figure shows the mean change in hydroclimatic intensity and not the 'mean hydroclimatic intensity'.

The measure is an indicator for changes in time compared to the normal period, in this case 1976-2005. Therefore, it is astonishing that the authors give the mean change for the period 1976-2099. Please, explain your intension.

P8L228: Figure 7 is much more complex than Figure 6 – therefore it is not good to throw them together and explaining them a bit than coming back to Figure 6 and afterwards again explaining Figure 7.

P8L231(and Figure7) To start the discussion of Figure 7 it is misleading to use "The root-mean-square error. . ." as it is very import for understanding the colors and signs that you are talking about a normalized root-mean-square error – normalized by the

root-mean-square error of the ensemble median. At the Figure caption (L724) is written "median RMSE normalized by the spatial standard deviation of the index climatology in the reanalyses (RMSEstd)" whereas the text (L238) gives "multi-model median error on a global scale (i.e. RMSEstd)" here the wording should be exact. "Figure 7 shows that the magnitude of the multi-model median error on a global scale (i.e. RMSEstd) is generally larger for precipitation indices than for the absolute and percentile-threshold indices based on temperature with the exception of csdi and wsdi." – it is not clear why the authors praise percentile-threshold indices as the RMSEstd seem to be even worse than the precipitation values.

P33L724: "...the ensemble mean error..." – "error" this is too short for readers not familiar with (Sillmann,2013a)

P9L271: "The result is presented as annual time series of the total number of heat or cold wave days for..." – this is not the case – in Figure 8 "summer days" are shown without any restrictions regarding the sequence of days, so the 'wave' aspect is excluded.

P10L298: It would be nice to get an example for the combined index as Fig. 9f.

P10L304: (Déandreis et al., 2014) is missing in reference list. And the document references on the ESMValTool web site (https://docs.esmvaltool.org/en/latest/recipes /recipe_diurnal_temperature_index.html?highlight=D%C3%A9andreis #diurnal-temperature-range) does not exist anymore: Déandreis C. (IPSL), Braconnot P. (IPSL), Planton S. (CNRMGAME). Study performed for the DALKIA company. http://secif.ipsl.fr/images/SECIF/documents/ Communication/fiche_invulnerable/RC_indicateur_EN.pdf

P10L307: Please, check "The measure is defined as the DTR exceeding 5°C at a given location and for a given day of the year" the suspicion suggests that "...as the DTR exceeding the value of the reference period by 5 K..."

P11L328: "daily instantaneous" - Lledo is using 6 hourly instantaneous wind speed data, which is courageous enough. Taking daily means results in strong underestimation of highest wind speeds and strongly disturbs the effect of the power curve but taking one instantaneous value is a strange idea and has to be justified.

P11L337: The daily cycle of surface incoming solar radiation is not a strength of climate models – therefore it would be very interesting to see a comparison with observed data.

P14L413: The period is given with "over 1900-2005" – please, correct.

Technical comments:

General:

- Check the space between two citations, it is sometimes missing (...Teixeira et al., 2014;Waliser et al., 2020)

- Check the space between number and unit, it should be 2 m instead of 2m (e.g. L184)

- The physical unit of temperature difference is Kelvin. Please, change °C and degree to K where necessary.

- Check reference list for unintended line breaks

P3L69 (Straus et al. 2007) is missing in reference list

P3L76 (Covey et al. 2003) is missing in reference list

P7L191 (Schulzweida, 2018) is missing in reference list

P11L326 (IEC, 2005) is missing in reference list

P13L388 (Ferranti and Corti, 2011) is missing in reference list

P9L281 'American Academy of Actuaries' should be the entry name in the reference list (instead of Actuaries Climate Index)

P17L513 'Impact du changement climatique sur la gestion des réseaux de chaleur.'is

not in the text

Figures:

P29Figure2: The headlines of the subplots give 'trend 1976-2099', whereas meant is changes in 2006-2099 with respect to 1975-2005.

P31Figure5: The plot title give "of Event" – could you change to "drought Events" to make obvious what is meant?

P31L705 Please, correct "historic (1950-2100)".

P34Figure8: "a)" and "b)" should be added to the figure. "Average number of summer days" should be replaced by "Average number of yearly summer days". 80th quantile should be 80th percentile.

The colorbar should be adjusted to the 'normal' number of days - for the exceeding of the 80th percentile 18 days out of 92: green colors for less than 18 days and yellow to red colors for more than 18 days.

P35Figure9: All units are given with 'cm', why?

P36L738-739 d) and e) are mixed and "maximum" is missing: "d) drought and e) maximum precipitation".

P34-36Figure8-10: The relation of x and y dimension of 2D-Figs 8-10 should be more realistic.

P37Figure11: Please, reduce the accuracy of the colorbar and reduce its size appropriate to the plots.

P38L777: Add the colors of observations: "as ERA-Interim (yellow line) and CRU (black line) data shown".

P39Figure13: Consider a legend for the regions abbreviations beneath the annual plot.

P40Figure14: The technical quality is poor.

P41Figure15: The technical quality is poor. Axis titles are insufficient (variable + unit missing).

P42Figure16 Does it make sense to give to exact date of the example - it is not comparable to reality/observed value. A hint that it is an artificial date would help.

P43Figure17: 1900-2005

Tables:

Table1 is hard to read as the column for 'Description' is rather narrow, consider landscape format if is possible and broaden the description-column.

Language/Spelling

P11L317: Please, consider rephrasing "is the single biggest contributor".

P11L326: ".."

P13L400: "example" instead of "examples"

P15L451: Do not use small caps but write "Version 2.0"