

## ***Interactive comment on “A bias-corrected CMIP5 dataset for Africa using CDF-t method. A contribution to agricultural impact studies” by Adjoua Moise Famien et al.***

**Adjoua Moise Famien et al.**

moflod@locean-ipsl.upmc.fr

Received and published: 25 February 2018

We thank the reviewer for the constructive comments and useful suggestions. Below we answer the different comments of the reviewer. We present all reviewer comments and our answers are given in blue.

General comments:

The presented study developed the bias-corrected CMIP5 GCM daily dataset using a combination of CDF-t method and WFDEI (and EWEMBI) and then com-

C1

pare with the ISI-MIP dataset which based on different bias-correction method and reference data (WFD). Some extreme climate indices (daily-based metrics) as well as maize yield simulated by a crop model were compared across different datasets to characterize the quality of the bias-corrected GCM daily dataset provided in this study. I respect the authors' efforts conducting this comprehensive analysis. Although this study analyzes only West Africa, their findings have implications especially for agricultural impact studies across the world. I only have a few concerns (listed below) and believe that most of them are minor. I recommend the acceptance of this manuscript with minor revision.

Specific comments:

1. Section 3.1. Although the detailed description of CDF-t method may be available in earlier study (Michelangeli et al. 2009, Vrac et al., 2012, 2016), a more completed explanation of key characteristics of the method in this section is unavoidable to make this manuscript stand-alone. Otherwise readers have to scratch around for. Particularly, it would be great if you could add an brief explanation whether the method forces the maximum (or minimum) value of a climatic variable in the future projection to be the same with that in historical period or not.

A more detailed description of the method has been provided with more specific statements and equations. It is indicated that “contrary to the QQ method that projects the GCM CDF of simulated future data onto the CDF of historical data, CDF-t considers the CDF change between GCM historical and future simulations”. So the CDF-t method does not force the maximum (or minimum) value of a climatic variable in the future projection to be the same with that in historical period.

2. Section 3.3. I am not convinced whether the current design of the sensitivity analysis is appropriate to evaluate the sensitivity of correction to the length of calibration period. When the data in 1979-1996 are used as the calibration subset, those in

C2

1997-2013 are used as the validation subset; this is fine. But, there is no independent validation subset when the data in 1979-2013 are used as the calibration subset. And it is easily expected that biases in bias-corrected data become the smallest when all available data are used as the calibration data. Therefore, the conclusion that the correction with the longest calibration period leads to the smallest biases is not examined using independent data. However, I think this part is not essential in this study. Removing or reanalysis are possible for this part.

You are right. This conclusion has been removed from the text (and we keep the figures regarding this analysis in Supplementary Information for the readers).

3. Fig. 5. In GUICOAST JAS, ISI-MIP bias-corrected precipitation data distantly distributed from WFD. Why? This is unreasonable because most methods including the ISI-MIP bias-correction method forces GCM data in the historical period similar to the reference data (WFD). A plausible explanation is necessary.

We agree with the reviewer's comment. This figure has been corrected.

4. Fig. 14. Two tendencies are observed in this figure. One is that the 95th percentile precipitation values from ISI-MIP data largely varied by GCM compared to the spread across the GCMs found in CDF-t method. Why?

It is difficult for us to answer this question because we did not applied the ISI-MIP bias-correction method, but only got the simulations from a dedicated website.

Another one is that in SAHEL JAS the IPSL-CM5A-LR data corrected by CDF-t method are relatively far from other GCMs of CDF-t data. A brief explanation is required.

C3

We do not understand this remark as the IPSL-CM5A-LR data corrected by CDF-t is quite near the other bias-corrected GCMs on Fig.14.

5. P25L2. "It indicates also that WFD data and related bias-corrected simulations should not be used anymore." I think this is overstated. Please consider rephrasing or removal. I agree that a use of WFD leads to biased crop yields in crop model simulation mainly due to biases in solar radiation, as demonstrated in your analysis. However, simulated yield variability and/or projected future change in yields would not be affected in relative term when biases in solar radiation are a main issue (for instance, see Iizumi et al., 2010). Reliable projection of yield change in absolute term is challenging, and therefore projected relative change in yield is still only a main source of information for adaptation planning and other application.

You are right. This sentence has been removed.

6. Section 4.3. The presentation of daily-based metrics is relatively not well organized in the current manuscript. The analysis and findings on the daily-based metrics themselves are useful, but not comprehensive compared to earlier study examining daily-based metrics (e.g., Iizumi et al., 2017, JGR, doi:10.1002/2017JD026613). Why did you select a limited number of metrics for this analysis? More importantly, it seems that the importance of the analysis results is not equal to that of the analysis of crop model. The simulated maize yield is used as a metric in the current manuscript, as the daily-based metrics are did so; though the maize yield has much importance in the manuscript compared to the daily-based metrics. A justification to present a limited number of daily-based metrics is necessary.

We chose to select a limited number of daily-based metrics because our main objective in this manuscript is to introduce the new bias-corrected dataset over Africa whose the CDF-t correction method has been applied for the first time. We wanted

C4

also to limit the length of the paper and therefore we focused on priority metrics defined by stakeholders in AMMA-2050 (note that some metrics have been put in Supplementary Information). We have provided now a link to a more complete metrics report produced as a deliverable for AMMA-2050. We chose also to provide a more detailed analysis on the sensitivity of crop simulations to bias-corrected data because this is a quite important issue for stakeholders and farmers, and because we think interesting to evaluate how an impact model can integrate non-linearly the diversity of forcing datasets.

Technical corrections:

7. P1L2-3. Why has CDF-t method never been applied to Africa? Is that due to low availability of daily weather observations in that region?

No, as it is said in the text, CDF-t method has been mainly applied over Europe. It is the first time that Africa has been considered, thanks to AMMA-2050 project. The different and successive reference datasets (WFD, WFDEI, EWEMBI, and others) used in bias-corrections have a global land cover. Of course, daily weather observations are less dense in some areas in Africa, and this introduces a certain level of uncertainty of the reference dataset, that must be kept in mind by users. This is another reason for comparing bias-corrections using various reference datasets.

8. P2L6. "robust" biases. Do you mean "persistent" biases?

Robust biases means biases that have not been reduced up to now, for instance between CMIP3 and CMIP5 GCM simulations. For instance warmer than normal SSTs in the equatorial Atlantic leads to a too southern location of the ITCZ in summer over West Africa. A sentence has been added.

C5

9. P2L19. "high and robust biases". Do you mean "large and persistent biases"?

"Large and robust". We indicate in the text the reference Roehrig et al. (2013).

10. P3L24. "northern summer". Is this "northern hemispheric summer" or "boreal summer"?

We modified for "boreal" summer.

11. P'L7. "anomalies". Probably, "differences" is more appropriate here.

Yes, it is been corrected.

12. Fig. 3. "observations". This should read "references".

Yes, it is been corrected.

13. P12L6. "standardized variance". This should read "standardized standard deviation" by definition of Taylor diagram.

Yes, it is been corrected.

14. P21L6. "GDHY" are a hybrid of FAO country yield data, satellite-derived crop-specific vegetation index and global crop datasets on crop calendar, harvested area and production shares achieved by different growing season. Subnational yield statistics are used to validate the grid-cell yield estimates, but not used as the input to estimate grid-cell yields.

C6

Ok, this has been corrected in the text.

15. P23L9. "Rsds" -> "rsds"?

Yes, this has been corrected.

16. Fig. 19. The label of y-axis should be replaced by "maize yield" instead of "crop yield" to be more precise.

Ok, this has been corrected.

17. Table 2. The caption of the table needs to include sufficient information to interpret the results presented in the table. Saying "see detail in the text" is not acceptable from the viewpoint of readability.

Ok, it has been completed.

All corrections will be applied in the revised manuscript.

---

Interactive comment on Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2017-111>, 2017.