

Interactive comment on “Spatio-temporal consistent bias pattern detection on MIROC5 and FGOALS-g2” by Bo Cao et al.

Bo Cao et al.

martincao12@gmail.com

Received and published: 28 September 2019

The manuscript presents the description of a method to cluster and classify main regions of bias in 3D grids and its application to the outputs of two climate models. The merit of the work is fair, improved by the fact that the code is shared and licensed under GPLv3. However, there are several shortcomings that need to be addressed.

[Thank you for taking your time and effort in reviewing our manuscript and we appreciate your insightful comments for improving the manuscript.](#)

Q.2-1

C1

The proposed methodology is, somehow, not new. Clustering analysis has been used to check model performance now for a long time and this is not properly put into context in the Introduction. It is surprising that the work by Yokoi et al. (2011) is not cited at all (<https://doi.org/10.1175/2011JAMC2643.1>) when moreover uses as examples MIROC5 and FGoals1. However, in the Introduction the authors cite other works that simply put the focus on regional studies or features, such as the ITCZ. I strongly recommend the authors to perform a better and more complete review of the literature on this topic and improve this section.

[Response: Thank you for pointing out the relevant references that are currently missing in the Introduction. We will improve the Introduction by reviewing more existing clustering analysis works on model performance such as Yokoi et al. \(2011\) \(<https://doi.org/10.1175/2011JAMC2643.1>\), Sierra et al. \(2018\) \(<https://doi.org/10.1007/s00382-017-4010-5>\), and Khan et al. \(2018\) \(<https://doi.org/10.3390/w10121793>\), and clarify the differences between these works and ours. That is, the existing works such as Yokoi et al. \(2011\) focused on clustering performance metrics to group correlated metrics, whereas our proposed method is applied to climate model outputs directly and tries to find local spatio-temporal consistent bias patterns in the outputs.](#)

Q.2-2

Also, the code (see also comment from the executive editor about using ZENODO) includes a README file, but this file does not contain any information about the structure of all the code in the repository. It should explain what each 'program' does and list all the files in the repository. Moreover, the different pieces of code that you share, do not have comments at all. Including commentaries in the code explaining what every set of routines or group of code does, is a basic request in software development and very important when you share it. Therefore, please, include

C2

adequate commentaries in the software with explanations. Finally, in the text of the manuscript, when you mention the analysis and the tool to generate Figure 3, include the name of the code/routines that use to do it and make clear to the reader that they can be accessed from the link available in the 'Code and data availability section'.

Response: Thank you for your valuable suggestions on code readability. We will enhance our code readability by adding the structure of the code in the README file and adding adequate commentaries in our codes. Figure 3 in our manuscript was generated using routine 'plotGlobal' defined in 'plotHeatMap.py'. We will add the names of these routines in Section 2.3 and explain how a bias family can be visualized and analyzed by using our tool.

Other minor issues:

Q.2-3

- the current title of the manuscript is misleading. Readers can think that you perform a full detection and study of bias in the two models mentioned. Indeed, your manuscript is about a clustering technique and MIROC and FGOALS are simply used as testbeds without any interpretation. The title should reflect this, maybe something similar to 'An algorithm for spatio-temporal consistent bias pattern detection: a case study using MIROC5 and FGOALS-g2'

Response: It is indeed correct that MIROC and FGOALS were used as testbeds in this manuscript. We will modify the title to reflect that the focus of this manuscript is the proposed clustering technique and MIROC5 and FGOALS-g2 are used as case studies.

C3

Q.2-4

- If you want to highlight the relevance of climate models for the study of climate change, probably the papers that you cite in the second line of the Introduction are not the most suitable. Better use Taylor et al. (2012) or IPCC reports.

Response: We will add the citation of Taylor et al. (2012) and IPCC reports to the second line of the Introduction.

Q.2-5

- When you mention the CMIP5, please, cite the paper by Taylor et al. (2012) (<https://doi.org/10.1175/BAMS-D-11-00094.1>).

Response: We will cite the reference of Taylor et al. (2012) on P1 when first mentioning the CMIP5.

Q.2-6

- The last sentence on page 2 is unnecessary, delete it.

Response: We will remove the last sentence on P2.

Q.2-7

- When you state on page 3 (first paragraph) that the computational cost would be too high, please, include extra detail at least on orders of magnitude of how more expensive it would be.

C4

Response: We will explain in details why considering spatial and temporal characteristics simultaneously is of a high computational cost on P3. Actually, the computational cost of a naive joint clustering algorithm would be the product of the costs when focusing on spatial or temporal dimensions separately. In this paper, we mainly focus on bias patterns that steadily appear in a region over a period of time and thus we can reduce the cost by tracing the bias pattern along time. The time cost of the proposed method is affected by the distribution of underlying bias patterns. In terms of data used in the case study, it can give the analysis result in minutes.

Q.2-8

- when you mention the models used, please, provide links and cite the relevant papers with the description of the models.

Response: Links to download the data have been provided in 'Code and data availability' and we will attach detailed data set Ids for data downloaded from ESGF. In addition, we will cite the papers that describe the models themselves:

1. Bao, Qing, et al. "The flexible global ocean-atmosphere-land system model, spectral version 2: FGOALS-s2." *Advances in Atmospheric Sciences* 30.3 (2013): 561-576.

2. Watanabe, Masahiro, et al. "Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity." *Journal of Climate* 23.23 (2010): 6312-6335.

The second paper has been cited in the manuscript and we will cite it the first time 'MIROC5' appears.

Q.2-9

- the citation and reference to HUANG Xin is capitalized, fix it.

C5

Response: We will fix the capitalization of the reference to Huang Xin et al. (2019).

Q.2-10

- page 7: you have already explained what CMIP5 stands for in page 1, you do not need to repeat it here.

Response: We will use the abbreviation of CMIP5 directly on P7.

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2019-107>, 2019.

C6