

We would like to thank the three reviewers for providing valuable comments that helped in a better representation of our manuscript. Please find below our replies following the comments. Comments are listed first, followed by replies and associated changes. While revising the manuscript based on the comments, we have also corrected some errors throughout the manuscript to further improve it.

Reviewer #1 (Dr. Massey)

General Comments

This paper describes a software suite, the Regional Climate Model Evaluation System (RCMES), which can be used to evaluate the performance of regional climate models, in particular those that are contributing to the CORDEX regional climate projections project. RCMES is built on top of the Open Climate Workbench (OCW), which is an open source and community-governed software toolkit to facilitate the analysis of climate models and observations. RCMES consists of two main parts - a database of observations for which to calibrate models against, and the RCMES toolkits. The toolkits provide functionality to load datasets in multiple formats and from four data sources, without having to know the URLs, etc. of the data sources. The toolkits also provide analysis routines for model evaluation as well as plotting routines for common plots such as the Taylor diagram.

From a software point of view, RCMES provides a valuable resource to the Regional Climate Modelling community in that it provides a standardised set of analysis tools which are comprehensive and easy to use. In particular, the three tiers of user interaction (the CLI, CFiles and using OCW with Python scripting) allows for a shallow learning curve. Standardised analysis routines enable analyses to be directly comparable across research groups, i.e. you know that the bias is computed in exactly the same way for each model / observation comparison. This is especially valuable to the CORDEX consortium, to enable comparison of the model analyses.

The community driven software development and community governance ensures that the software is freely available, that individual or teams of researchers can make contributions but that code additions are approached in a peer-reviewed fashion to maintain quality.

The paper itself is comprehensive, well written and provides a valuable overview of the RCMES project, which will help users taking their first steps in using the software and provide a short reference and pointers to more in-depth resources for more experienced users.

I have few comments, mostly centred around grammar, typos, bad citations and clarifications.

I recommend publishing the paper with these minor corrections, which I have listed below.

→ We would like to thank Dr. Massey very much for providing valuable comments on the manuscript, especially those meticulous technical corrections.

Specific comments

P2L22: CMIP is currently in its sixth phase. The fifth phase has been completed, but is the one everyone is using.

→ We have corrected this.

P7L2: Capitalise "yaml" to "YAML". YAML is not part of Python, it is an independent mark up language, so remove "Python" before this

→ We appreciate this comment. YAML in the manuscript has been capitalized.

P10L9: Can the loader use the CF-compliant standard name, from the attributes metadata in a netCDF file, to load the latitude, longitude, time and level variables?

→ Yes, it can. However, as stated in the following sentence, users can provide names for the latitude, longitude, and time to load non-standard files.

P10L20: Can the loader use OpenDAP to access a remote dataset?

→ The ESGF loader in the released version downloads a NetCDF file without subsetting the data using OpenDAP. When accessing PO.DAAC data, OpenDAP is used.

P11L2: what is a "granule" in this context? Please define.

→ The following sentence has been revised.

(Before) Current functionality includes the ability to retrieve and extract full granules and/or specific variables from over 50 Level 4 blended datasets covering myriad of gridded spatial resolutions between 0.05 and 0.25 degrees...

(After, P11L6) Granules are equal size subsets of a satellite's observations along its track. Current functionality includes the ability to retrieve and extract granules meeting a search criteria and/or specific variables from over 50 Level 4 blended datasets covering myriad of gridded spatial resolutions between 0.05 and 0.25 degrees...

P12L22: Does the temporal resampling support sub 24-hour temporal resolution? e.g. 3 or 6 hourly. If not, does this mean the toolkit cannot be used to evaluate the diurnal cycle? Having a good representation of the diurnal cycle is very important for regional climate modelling, especially for heatwave extreme events.

→ We thank the reviewer for providing this thoughtful comment. However, the resampling does not support sub-daily resolution. We will try to provide this functionality in the next release.

P14L2: by "slower" do you mean slower in installation time or slower in performance once installed? Please make clear.

→ The following sentence has been revised.

(Before) The latter is slower than the package installation, but it provides Linux OS and all required python libraries that can run on any types of users' computers.

(After, P14L12) RCMES execution with the latter is slower than the package installation, but it provides Linux OS and all required python libraries that can run on any types of users' computers.

P14L20: you could also explore creating a Docker container, to make packaging the dependencies easy without the performance hit of a VM.

→ We think this suggestion is very interesting. We are testing RCMES Docker containers to test Docker Swarm to parallelize RCM evaluations for multiple CORDEX domains. We plan to publish another paper on benchmarking different parallelization methods for RCMES.

P14L25: isn't github a "collaborative platform where climate scientists can spontaneously share their software updates". What is different here? Please explain.

→ Yes, github is a collaborative platform. However, many projects on github are not sustainable beyond the period of developers' funding. This is why we transitioned the RCMES development to the Apache Software Foundation's OCW.

P15L16: "over the years": how long has OCW been active as a project.

→ The following sentence has been revised.

(Before) In addition, over the years OCW has been presented in countless conferences.

(After, P15L25) In addition, OCW has been presented in countless conferences since its first release in June 2013.

P18L16: "By combining BMA with ABC, a diagnostic based approach for averaging regional climate models becomes possible": please provide a reference for this.

→ We have added Turner and Zandt (2012), Vrugt and Sadegh (2013), and Sadegh and Vrugt (2014) to the corresponding paragraph.

Technical corrections

P2L7: comma after "Yet"

→ We have added a comma.

P2L13: Change "Because of" to "Due to"

P2L25: Change "based on GCM" to "which is based on GCM"

P2L29: Change "is" to "are" in "is now underway"

P2L32: Change "Because of" to "Due to"

→ We have changed these as suggested. Thank you again.

P3L8: Define "DOE"

→ DOE stands for the United States Department of Energy. It is defined in the revised manuscript.

P3L8: Bad citation (/citeclimatemodeling)

P3L35: Bad citation(citetpodaac)

P8L22: bad citation ("citetLee17")

P15L3: bad citation citetapache

→ We have corrected these and other bad citations.

P4L2: change to "fostering the collaboration" or "fostering collaboration between"

→ We have changed this.

P4L5: put commas around "therefore"

(Before) To promote greater collaboration and participation of the climate research community within the RCMES development process, we transitioned from a closed-source development process to an open-source software (OSS) community driven project hosted in the public forum and therefore subject to public peer review, something which has significantly improved the overall project quality and standards the community and project holds itself to.

(After, P4L2) To promote greater collaboration and participation of the climate research community within the RCMES development process, we transitioned from a closed-source development process to an open-source software (OSS) community driven project hosted **in the public forum**. **As a result**, the development process is subject to public peer review, something which has significantly improved the overall project quality and standards the community and project holds itself to.

P4L7: capitalise python to Python

P4L11: change "as well as describe" to "as well as to describe"

P5L2: change "can utilize Open Climate Workbench" to "can utilize the Open Climate Workbench"

P5L2: add "(OCW)" after "Open Climate Workbench"

P5L2: change "build up" to "write" or "produce"

P10L2: capitalise python to Python

P11L3: missing "a" between "covering" and "myriad" - i.e. should be "covering a myriad"

P13L20: change "that are deviated from" to "that deviate from"

P13L26: capitalise "yaml" to "YAML"

P14L4: capitalise "python" to "Python"

P13L26: capitalise "yaml" to "YAML"

P14L4: capitalise "python" to "Python"

P14L6: capitalise "python" to "Python"

P14L11: capitalise "python" to "Python"

P15L5: change "climate scientists" to "climate scientist"

P15L6: change "becomes of" to "would become of"

P16L6: change "last release" to "latest release"

P16L8: change "is" to "are" in "contributions from any party is reviewed"

P18L6: commas around "with associated uncertainty"

P18L17: change to "development of the OCW dataset processor" - add "the"

→ We have applied all of these to the revised manuscript.

P8L15: change "use" to "uses" in "RCMES CFiles use"

→ CFiles stand for configuration files. So, we keep 'use'.

P13L8: change "that are recently" to "that have been recently"

P13L8: change "long-term trend simulated" to "long-term trend of simulated"

(Before) Another set of metrics that are recently added enables evaluation of long-term trends simulated climate models over the contiguous United States and analysis of the associated uncertainty.

(After, P13L12) Another set of metrics that **have been** recently added enables evaluation of **long-term trends in climate models** over the contiguous United States and analysis of the associated uncertainty.

P13L24: these downscaling method are not all simple!

→ We think that these four methods are relatively simple compared to state-of-art empirical downscaling techniques.

P14L4: change "any types of" to "any type of" all "all types of"

→ "any types of" has been replaced by "any type of".

P15L6: "everyone and anyone" not necessary

→ We agree with this. "everyone and anyone" has been replaced by "everyone".

P16L3: by "committers" do you mean "code committers"?

→ Yes. "committers" in P15 has been replaced by "code committers".

P16L25: change "one which has" to "ones which have"

→ We have changed "one which has" to "those which have".

P17L34: change "interrogating" to "determining". Comma after "poorly"

(Before) Examples include interrogating why some models simulate extremes poorly as related to biases

(After, P18L10) Examples include **investigating** why some models simulate extremes poorly, as related to biases

P18L4: provide reference for "ABC"

→ We have added Turner and Zandt (2012).

References error with references: "<GotoISI>" (undefined in LaTeX maybe?) please check all references

→ We have checked all references and corrected several errors. We appreciate the comment.