Geoscientific

Model Development

Discussions

[GMDD]

Interactive

comment

Interactive comment on "Observations for Model Intercomparison Project (Obs4MIPs): Status for CMIP6" *by* Duane Waliser et al.

Anonymous Referee #1

Received and published: 13 January 2020

Overview:

The manuscript provides an update to the Obs4MIPs effort detailing progress in recent years and demonstrating the project's role in the most recent generation of the Coupled Model Intercomparison Project (CMIP6). The key points of this manuscript are: 1.) Obs4MIPs is undergoing expansion from $\sim$80 datasets to possibly close to 200, 2.) improved characterization of data quality, coverage, and suitability for datasets that are part of the project; and 3.) enhanced support for supplemental information and code to aid the end user. Notably, there is a fundamental shift in the type and quantity of datasets included in the effort that favors a more diverse collection of datasets that sometimes overlap each other. This more inclusive approach, however, is in potential conflict with prioritizing "ease of use" to the end user. Overall, this is an important

manuscript and project that bridges the climate observational and modeling communities. I would recommend publication pending revisions that address the concerns below.

Major comments:

1. On Page 5, Lines 16-18, one of the stated goals of the project is that Obs4MIPs is to assist a growing community of scientists "without an expert's understanding of either the observations being employed or the climate models themselves." There are several notable advances in this effort to help a user better understand the observational dataset, such as the dataset indicators and more supplemental information. How does Obs4MIPs help users understand the climate models themselves? This point wasn't clear to me. Also, the stated goal of helping the end user is in potential conflict with other aims of the project. (See Major Comment #2.) Perhaps there are some use cases or metrics that could be provided as evidence that the Obs4MIPs effort is having in the community (publication counts referencing data, etc.)

2. The discussion on data redundancy, present on Page 18 Lines 22-27, deserves much more attention. The shift from a high degree of initial oversight toward a more inclusive approach to accepting datasets marks a fundamental philosophical shift in the project. For example, there are now multiple references to different SST products (Page 7 Line 22, Page 9 Line 24, Page 12 Line 18). I personally agree that a diversity of datasets is ultimately a virtue, but it is potentially in conflict with the goal of assisting the "non-expert" user. How can a non-expert choose which product is best for given application when multiple datasets of the same product exist? Are dataset indicators enough?

3. The relationship with ana4MIPs and other similar efforts could be spelled out earlier in the manuscript. It is not until the last section that the reader discovers that the "ana4MIPs" effort is static and the introduction of CREATE-IP on Page 10 Line 11 is lacking context. In the community, Obs4MIPs and ana4MIPs are mentioned in the

same breath. Important distinctions between the two are necessary.

4. Can more be said about the different spatial coverages of the Obs4MIPs data? How are gridded and point measurements handled differently by Obs4MIPs? Are datasets derived from floats, such as ARGO, compatible with the Obs4MIPs? There are references in the last paragraph of the manuscript, but these issues need to be addressed earlier.

Minor Comments:

Page 4, lines 17-19: These are necessary capabilities, but it is not a complete list.

Page 5, line 12: Consider a sentence or two expanding about how observations help identify and correct model shortcomings. Observations allow validation of the models climatological mean state, annual cycle, and variability across timescales helping to assess model fidelity among many different Earth System Processes.

Page 7, lines 16-18: What was the criteria used to determine what datasets were "appropriated for climate model evaluations?" Was this based on demand from the modeling community, external users, assessment by the dataset creators, or some combination?

Page 8, line 15: "apprised" is an odd word choice. This implies that the Obs4MIP effort was taking place without a dialogue with the modeling community.

Page 9, lines 13-17: Rather than a standalone paragraph, consider joining these lines with the previous paragraph.

Page 12, lines 7-8: This sentence is awkward. How do the improvements in CMIP output specification impact the Obs4MIPs holdings?

Section 4b: As data standards/specifications evolve, what happens to existing datasets? Are they left as-is or converted?

Page 15, Lines 21-22: Can more be said about the indicator assignment process?

Does it rely mainly on self-reporting? How is the review conducted?

Page 19, Lines 7-10: How will the assignment of data DOIs work in practice? Would an end user be citing both a data DOI and a scientific paper describing the dataset? As noted, this problem extends to CMIP6 data and beyond, but it poses an interesting issue for the peer-review publication process.

————————————————————

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-268, 2019.