

RESPONSE TO REVIEWER #1 FOR GEOSCIENTIFIC MODEL

DEVELOPMENT: MANUSCRIPT EGUSPHERE-2025-1380

BY Zhongwang Wei, Qingchen Xu, Fan Bai, Xionghui Xu, Zixin Wei, Wenzong Dong, Hongbin Liang, Nan Wei, Xingjie Lu, Lu Li, Shupeng Zhang, Hua Yuan, Laibao Liu, and Yongjiu Dai

We thank Reviewer #1 for thoughtful and constructive feedback. This Response to the Reviewer file provides a complete documentation of the changes that have been made in response to each individual comment. Reviewer's comments are shown in plain text. Authors' responses and quotations from the revised manuscript are shown after **Reply**.

1. I thank authors for providing this revised manuscript, and for incorporating previous suggestions. I have two suggestions relating to this latest revision.

Reply:

Thank you very much for your review. I will address each of your comments and propose revisions to improve our manuscript.

2. Authors introduce a discussion on the parallelisation of the software, describing its performance as demonstrating “remarkable efficiency” and “outstanding scalability.” However, the results comparing 1 to 48 cores indicate that the scalability falls significantly short of linear. I recommend moderating this language to simply state that parallelisation provides speed improvements. Additionally, the section could be made more concise, as it is quite long.

Reply:

Thanks for the valuable suggestion. We have revised our manuscript (**P8L163-P8L173**): **P8L163-P8L173**: "OpenBench achieves speed improvements through its parallel processing architecture. Benchmark tests demonstrate clear advantages over sequential processing methods. In station-based evaluations, processing a single variable across 142 stations takes 3.12 minutes using single-process execution, whereas parallel processing with 48 cores reduces this to 0.509 minutes on an Intel(R) Xeon(R) CPU E5-4640 v4 @ 2.10 GHz with 48 GB RAM. OpenBench uses Dask's lazy execution

and chunked arrays for efficient gridded data processing, balancing memory use and processing speed. Processing 0.25° resolution model outputs (2001-2010, monthly) against two reference datasets takes 2.302 minutes sequentially versus 1.301 minutes in parallel on the same hardware. These performance improvements are particularly beneficial for comprehensive model evaluations involving multiple variables, reference datasets, and spatial domains. The efficiency gains from parallel processing become more substantial with higher-resolution datasets and increasing numbers of evaluation sites, making OpenBench suitable for both rapid diagnostic evaluations on personal workstations and extensive comparative studies on high-performance computing systems."

3. Authors have linked to an updated software repository, and a thoroughly revised README here: <https://github.com/zhongwangwei/OpenBench>. I tried running the updated software following "Usage" instructions, with the command:

```
python script/openbench.py nml/main-Debug.json.
```

Unfortunately, when running on my HPC I encountered this error:

```
FileNotFoundError: [Errno 2] No such file or directory:
'./output/Debug/output/scores/Evapotranspiration_stn_GLEAM_hybrid_PLUMBER2
_station_case_evaluations.csv'
```

In my first review I was able to run this to completion without error. I also suggested in my first review that authors provide a stripped down, highly simplified example script which will allow new users to test a basic version of OpenBench on their system (the previous example test was "main.json"). I reiterate this suggestion, as "main-Debug.json" appears to be a complex configuration for debugging purposes, i.e. it runs more than a simple case.

Reply:

We appreciate the reviewer's rigorous testing and valuable feedback. We have fixed the reported error. Please find our responses below:

- 1) The reported FileNotFoundError has been fixed. We recently developed an "only_drawing" module to support secondary visualization of existing evaluation

results. The error occurred because this option was mistakenly set to true in the configuration file. We have corrected this by setting it to false in the default configuration.

- 2) We continue to recommend using main-Debug.json as the primary example script, as it enables comprehensive testing of all four evaluation types: Grid-Grid, Grid-Station, Station-Grid, and Station-Station.
- 3) Users can selectively disable specific tests by modifying the configuration file main-Debug.json according to their testing needs.

We sincerely appreciate the reviewer's suggestions that helped improve OpenBench's accessibility. Please don't hesitate to contact us if further issues arise.