Authors' Response to Reviews of

Implementation and validation of a supermodelling framework into CESM version 2.1.5

W.E. Chapman et al.
Geoscientific Model Development, EGUsphere-2024-2682

RC: ***Reviewers' Comment***, AR: Authors' Response, □ Manuscript Text

We thank the reviewers for their thoughtful comments, and have included an acknowledgment to their work in the manuscript. We highlight that the manuscript has gone through a few significant changes due to this review and we believe it is in a much stronger condition. It has been refined greatly with an eye towards readability and flow. We include many more error statistics (see table 1 in the supplemental material). The github repository has been restructured for user clarity (which was a concern of reviewer 1, and this has resulted in new versions on zenodo), and we highlight the improved use-ability of the code and modeling framework. Additionally, the GPCP dataset has been remapped with a new interpolation and this updated many figures, though the broad findings are unchanged.

1. Reviewer #1

RC: ***This paper describes a supermodelling framework that combines the effects of the physics parameterization suites from two versions of the Community Atmosphere Model. The physics suites are combined by nudging each version (model component) to the averaged state on a periodic time interval, as in data assimilation.***

RC: ***Supermodelling is an interesting idea with interesting potential applications, but I think that the manuscript is not ready for publication for three reasons. First, the method is not described in enough detail to allow a reader to reproduce the results. For instance, I couldn't find a tag of the code listed in the manuscript that would allow me to recreate the figures in the manuscript. Furthermore, the scripts have hardwired paths (not variables) with little instruction given as to what lines in the scripts need to be modified. Second, the main application cited is improvement of supermodeled climatologies over the individual components, but only one or two improved fields are shown in the manuscript. The authors argue that improvement will come with tuning, but one might expect to see improvement even with simple equal weighting of the components. It might interest readers to see more discussion of which fields are improved and which are degraded, along with any insights that the authors have regarding the reasons that some fields are improved and others are not. Third, there are many typos in the manuscript, some of which impede understanding of the meaning. I list just a sample of them below. My recommendation is that the authors invest more time in the manuscript and deliver a more polished version.***

1.1. Author response:

We would like to sincerely thank the reviewer for their time and effort in evaluating our manuscript. We appreciate the constructive points raised and have made substantial revisions to improve clarity, documentation, and presentation of results. Below, we address each of the major concerns raised by the reviewer.

1. Clarity and Readability of the Manuscript

We acknowledge the reviewer's comment regarding typos and overall polish. In response, we have carefully revised the manuscript to improve clarity, correct errors, and enhance readability. We appreciate this feedback and believe that the current version is significantly more refined.

2. Reproducibility and Code Documentation

We understand the importance of reproducibility and have taken significant steps to improve documentation. The original manuscript included a GitHub repository containing all necessary code to reproduce the figures, but we recognize that the accompanying documentation may not have been sufficiently clear. We highlight that this is not the same repository that the supermodel built is in, which is stated in the **Code and data availability** section. This link to the repository has been highlighted in the supermodel REPO. Additionally, we have expanded on how a user can port their supermodel on any new machine. To address this:

- We have substantially expanded the README and associated documentation to provide explicit guidance on setting up and running the supermodel framework.
- We now explicitly link to a separate repository that reconstructs every figure from the manuscript. This was included in the original submission but may not have been sufficiently highlighted. It is now directly referenced both in the manuscript and the README.
- We have added further instructions on obtaining and installing CESM2.1.5 and the associated CONDA python environments to ensure users can set up the required environment, including explicit instructions on necessary modifications.

We believe these changes significantly enhance the reproducibility of our work.

3. Model Configuration and Hardcoded Paths

We would like to clarify that our implementation did not contain hardcoded paths except for those referencing the supported instance of CESM2.1.5. However, we recognize that clearer guidance on configuring paths for other environments would be beneficial. To this end, we have:

- Added explicit instructions for users to install their own version of CESM2.1.5.
- Clearly indicated where modifications should be made within the codebase to adapt it to different computing environments.

4. Evaluation of Model Performance

The reviewer notes that the manuscript primarily focuses on introducing the platform and does not extensively analyze the improvement of supermodeled climatologies. While we maintain that the primary goal of this work

is to introduce the framework, we agree that additional analysis can provide valuable context. Additionally, we have removed the sentence in the abstract that may mislead readers and now it is focused more on the platform itself. Therefore, we have expanded our discussion to:

- Include a more detailed comparison of improved and degraded fields of the prognostic variables (U,V,T,Q,PS) at multiple model levels.
- Included an RMSE table in the supplemental to capture the model state biases for the reader.
- Improved upon our insights into why certain fields show improvement while others do not.
- Clarify that while tuning will play a role in further optimizing performance, some improvements can already be observed with equal weighting.

These additions should help readers better understand both the potential and limitations of the approach.

Final Note

We appreciate the reviewer's feedback and believe that the revisions significantly enhance the manuscript. We hope that the changes we have made address all concerns and improve the clarity and reproducibility of our work. Thank you again for your time and consideration.

1.2. Minor Comments

RC: ***This part of the caption of Fig. 2 seems to have typos: "SUMO5 (teal); supermodel which uses CAM5-physics (SUMO5); supermodel supermodel which uses CAM6-physics (SUMO6) (purple) at location"***

... ~~SUMO5 (teal); supermodel which uses CAM5-physics (SUMO5); supermodel supermodel which uses CAM6-physics (SUMO6) (purple)~~ ... ... a supermodel which uses CAM5-physics (SUMO5; teal); a supermodel which uses CAM6-physics (SUMO6; purple)...

RC: ***Lines 32–33: "Additionally, biases which are shared across the individual models in an NIE cannot be corrected due to the linear nature of post-process averaging." But does the supermodel succeed in correcting such shared biases? Can the authors offer any insight into when and why shared biases can be corrected?***

AR: This line has been adjusted to point to literature in which the improvement of shared biases is indeed improved in supermodeling. We refer the reader to the revised text. We also encourage the reviewer to read the community comment by Gregory Duane, an expert in super modeling. Duane and Shen (2023) reveal that often improvements are manifest in representations of localized structures, rather than in reductions in RMS error. We have searched for evidence of this, but find no direct measure of improvement in our system. We have adjusted the text to include these references. Please see lines 37-40 in the new manuscript.

RC: ***Line 80: "the adaptation of the existing nudging toolbox (reference)". Typo. Please include whatever reference you're referring to here.***

AR: Fixed, Thank you.

RC: ***Lines 178–180: "We show results for four experiments: CAM5, CAM6, the supermodel which uses CAM5-physics, but is nudged to the combined state, SUMO5, and the supermodel which uses CAM6-physics, SUMO6." Does SUMO6 nudge to the combined state? If not, how is it different from CAM6? The sentence ends before the SUMO6 configuration is clearly described.***

AR: We apologize for the confusion, we were using shorthand though it is best to be explicit. SUMO6 is also nudged to the combined state. This has been fixed in the text.

> We show results for four experiments: CAM5, CAM6, the supermodel which uses CAM5-physics, but is nudged to the combined state, SUMO5, and the supermodel which uses CAM6-physics, but is nudged to the combined state, SUMO6.

RC: ***Line 195: "(Fig. ??, right column)". Typo.***

AR: This has been fixed, thank you.

RC: ***Lines 195–196: "Correlations between the SUMO5 and SUMO6 experiments are much higher (3, left column)" How can a correlation be as high as 3? Correlations must lie between -1 and 1.***

AR: The 3 refers to pointing the reader at Figure 3, not a correlation score. We have made this more clear in the text.

RC: ***Fig. 4: I'm surprised that CAM5 and CAM6 have such similar distributions of wind speed. Is there a computational or physical constraint on this distribution?***

AR: There are no computational or physical constraints, though the models have been tuned at NCAR to represent fields well in their creation, likely this is why the distributions are similar.

RC: ***Line 258: "One reason to develop supermodels lies in their potential to have smaller mean-field biases." In the example of surface precipitation shown in Fig. 7, the spatial pattern of errors looks similar (shared) in CAM5 and CAM6. The hope implied in lines 32–33 was that such error could be reduced below both components by supermodelling. But Fig. 7 shows that this doesn't happen for surface precipitation. Do the authors have an explanation for this failure?***

AR: Currently, our theory is that when we train the super model, these biases will all improve. However, for now we point to the under synchronization in the tropics as a potential cause. Generally, in the manuscript, we have de-emphasized the point that the model could contain lower biases and instead focus on the platform specifications and the system.

RC: ***Line 258: "The work presented her" Typo.***

AR: Fixed! Thankyou.

RC: ***Lines 278–279: "Positive values (grey) indicate that the SUPERense is outperforming the NIEnse and vise-versa." Are positive values grey or green?***

> Positive values (green ~~grey~~) indicate that the SUPERense is outperforming the NIEnse ~~and vise-versa~~ while negative values (blue) indicate the opposite.

RC: ***Lines 289–290: "This circumvents the in CESM costly initialization stage and introduces effectively PAUSE/RESUME capability (Fig. 1)." Typos?***

AR: Yes, we have fixed the typos.

> The exchange of information is managed through a novel Python-FORTRAN I/O interface that avoids the need to stop and start each model. ~~This circumvents the in CESM costly initialization stage and introduces effectively PAUSE/RESUME capability (Fig. 1).~~ This circumvents in CESM the costly initialization stage and introduces a PAUSE/RESUME capability (Fig. 1). This Python-FORTRAN bridge was also used to manage the timestamps in the output files, and efficiently write the pseudo-observations files.

RC: ***Lines 300-301: "To test our implementation we linked the CAM5/CAM6 atmosphere and confirmed that synchronization across various temporal scales and variables Even though the supermodels only exchange limited information . . ." What did you confirm?***

AR: We have fixed the typo that clarifies the text.

> To test our implementation we linked the CAM5/CAM6 atmosphere and confirmed that synchronization occurs across various temporal scales and variables.

RC: ***Lines 317–319: "To create all model runs and build your own supermodel, refer to Chapman et al. (2024). This second repository contains the setup for the SuperModel and its constituent models, including source modifications, model build scripts, and namelists for running the described CAM versions." Is a tag of the CESM repository listed somewhere so that readers can reproduce the figures in the paper identically?***

AR: Yes, though this was in the original manuscript, we have highlighted the text in the model README so it is easier to find and link to the figure repository in the new documentation.