



EGUsphere, referee comment RC2  
<https://doi.org/10.5194/egusphere-2022-387-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on egusphere-2022-387**

Anonymous Referee #2

---

Referee comment on "A modeling framework to understand historical and projected ocean climate change in large coupled ensembles" by Yona Silvy et al., EGU sphere, <https://doi.org/10.5194/egusphere-2022-387-RC2>, 2022

---

In this manuscript, a modeling framework is designed to investigate the transient ocean climate change and the effects of different surface flux perturbations on it. Validations of designed experiments but not all are provided based on the results from the IPSL model. Overall, the manuscript is well organized in format and the writing is clear. The idea about the application of the hist+ssp245 scenario to the ocean-only model is also interesting here. However, I think the scientific goals are still vague, as also mentioned by another reviewer, especially when I compare this modeling framework with the FAFMIP&Ocean-only FAFMIP experiments. The interactive sea ice could be an issue here, but since another reviewer has raised several questions about that, I mainly provided some comments on the treatment and interpretation of internal variability and externally forced responses. Revisions are necessary to make the manuscript publishable.

### Major Comments:

1. I think the key to making this modeling framework different from the ocean-only FAFMIP is the usage of hist+ssp245. However, the manuscript took me a while to confirm the guess when I read it. It would help readers understand this work better if the authors can mention this difference in the abstract and at the beginning of the manuscript.
2. In the abstract, you mentioned "The question of timescales ... is lacking" and may imply that these designed experiments are helpful to solve this. Is it a scientific goal here? However, I didn't find any related investigation or discussions in the main body. I believe that the validations under different timescales are needed if it is the scientific goal.
3. I agree that the usage of the large ensemble can help obtain the forced responses. However, the multi-model mean variables used in FAFMIP also do the same thing. Are

there other aspects to show the advantage of using large ensemble simulations in this study?

4. Some figures about perturbation experiments show the anomalies over the period when the CO<sub>2</sub> forcing is quite large, for example, the 2040-2059 average in Fig. 11. When the external forcing is strong, isolating forced response from internal variability is easier. Hence, I am wondering if the differences between ALL and IPSL ensemble mean are still small and insignificant when you look at the period when CO<sub>2</sub> forcing is not that strong, such as around the mid-20th century. In addition, how will it become when you only look at the difference in one year, not the average of 20 years or so. In other words, can this only one realization in ALL (HEAT, STRESS...) capture the result from the ensemble mean of 32 or 11 members at all time steps? Some validations are needed.

5. How many ensemble members are sufficient to obtain the forced surface flux perturbation? Are 11 members also fine? In particular, there are only 11 realizations after 2060 but 32 before that. Could you show a map about the difference in perturbation between the 32-member-mean and 11-member-mean in the year 2059 (and 1959 when CO<sub>2</sub> forcing is weak)? Could you also discuss the effects of ensemble size on your results in ALL, etc.?

6. The results from HEAT, STRESS, and WATER are not shown. At least, the sum of them should be compared with ALL in a figure. If possible, HEAT, STRESS, and WATER may be comparable to the corresponding runs from ocean-only FAFMIP. For the passive tracers runs, the sum of added and redistributed terms should be compared with the HEAT or WATER in a figure, at least.

Minor comments:

Title and Abstract: "large coupled ensembles" in the title is better to be also mentioned in the abstract.

Line 33: What do you mean by "decomposition of mechanisms"?

Lines 35-36: "remain unclear" to "are not fully understood"? There are some findings from previous studies.

Figure 5d,e,f: use dashed curve for CTL?

Line 291: I feel the warming in the Pacific is not "slightly".

Line 294: "...than the piControl interannual variability" but in Figure 7 you use 2 x STDDEV.

Line 366: "Differences may ...". Does this difference represent ALL-CTL minus ensemble mean anomaly or ALL-CTL?

Fig. 10a, c: Are these full-depth averages?

Line 408: "by twice the intermember standard deviation". I'm not sure whether the uncertainty range is too wide.

Fig. 11: similar to major comment #4, how about the results after 2060?

Line 438: "This paper ..." means this manuscript or the Silvy et al., in revision?