Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., author comment AC2
https://doi.org/10.5194/gmd-2021-372-AC2, 2022

Reply on RC1

Davide Zanchettin et al.

Author comment on "Effects of forcing differences and initial conditions on inter-model agreement in the VolMIP volc-pinatubo-full experiment" by Davide Zanchettin et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-372-AC2, 2022

We thank Referee #1 for her/his helpful comments on our manuscript. We report in italics relevant comments by the referee.

No observed values are used when evaluating model climatology and responses to volcanic forcing. Although I agree that the aim of this paper is to provide an initial assessment based on idealized experiments, not historical transient experiments which are comparable to the observations, assessing the degree of inter-model agreement in volcanic influences without any relevant comparison with observed values could be misleading given that models may have systematic biases. I strongly suggest including observed values somehow in their plots and interpreting results accordingly

We agree with the referee that a comparison with observed climate anomalies around the Pinatubo eruption may be a worthy addition to our analysis. In the revised manuscript we will therefore include observed anomalies in some of the figures (for instance, those regarding the temperature and precipitation response) and discuss how our multi-model results compare with observations.

This study aims at providing preliminary assessments but more efforts to quantify factors responsible for inter-model discrepancies would be useful. One way would be to add summary bar graphs or tables for some key variables (with observed estimates if possible, see my comment above) where readers can find actual values for individual models and how much differences exist between models and also between different ocean initial conditions (ENSO phases). Mostly, time series are displayed and it is inconvenient to identify specific model responses.

We are in favor of adding some relevant information in a more quantitative way, for instance through tables. In the revised manuscript we plan to include supplementary tables for relevant statistics of the global-mean near-surface temperature response, including ensemble-mean, ensemble spread, and maximum cooling with timing, also accounting for the effect of initial conditions and specifically the ENSO state. We would like to remind that all time series used in the manuscript are publicly available, so that quantitative estimates can be easily calculated in follow-up studies.

Some places need more explanations for better understanding. It's unclear how authors have selected samples for "equally distributed cold/neutral/warm states of ENSO and negative/neutral/positive states of NAO". Exact details of sampling methods look very

important for interpreting results as well as for planning the next VolMIP protocol. Also, authors consider radiation feedbacks in their evaluations but its association with inter-model spreads needs to be explained more clearly. Another one is why ECS is considered here, which represents equilibrium sensitivity to doubled CO2.

In the revised manuscript we will provide further metadata regarding the simulations, including the initial states sampled for the different participating models. The VolMIP protocol was somehow vague on how the "equally distributed cold/neutral/warm states of ENSO and negative/neutral/positive states of NAO" were to be sampled, so different groups proceeded differently, sometimes with a subjective approach. We will explain this better in the revised manuscript. In the revised manuscript we will also put ECS in better context.

Authors conclude that influence of ocean initial conditions is weak or even negligible but this conclusion can be dependent on how to measure ENSO-like responses. Other studies used relative SST as authors briefly mentioned, and results can be affected much by applying different metrics. Since understanding ENSO influence is one of major issues, I think that adding more discussion with appropriate sensitivity tests would be useful, e.g. comparing relative SST responses with Nino3.4 responses. In terms of NAO or AO responses, target season and region can be revised as boreal winter and high latitude areas, for better comparisons with previous findings.

We agree that relative SSTs are a better basis than absolute SSTs to calculate the Nino3.4 index and capture ENSO-like responses. In fact, we used the Nino3.4 index as defined by the VolMIP protocol. This is one of the aspects where there is room for improvement in a possible second phase of the initiative. Therefore, we will keep the original analysis based on absolute SSTs, but also add results corresponding to a Nino3.4 index calculated from relative SSTs. We will compare results and discuss the implications for the protocol in follow-up VolMIP activities.