

Earth Syst. Dynam. Discuss., community comment CC1  
<https://doi.org/10.5194/esd-2022-26-CC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on esd-2022-26

John Fasullo

---

Community comment on "The future of the El Niño–Southern Oscillation: using large ensembles to illuminate time-varying responses and inter-model differences" by Nicola Maher et al., Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2022-26-CC1>, 2022

---

Review of The future of the El Niño-Southern Oscillation: Using large ensembles to illuminate time-varying responses and inter-model differences

by Maher et al.

The manuscript by Maher et al. seeks to diagnose changes in ENSO in 14 single model large ensembles (so-called SMILES). The manuscript builds upon a body of work that is often based on single members, individual models, or idealized models and so it represents an advance, particularly at resolving the decadal varying aspects of forced changes in variance - which the work shows can be important. The manuscript is clearly written, is explicit about its objectives, findings, and reasoning, and includes figures that are well-designed and clear. There is sufficient new material here to justify publication. Some aspects are frustrating - such as in cases where the robust take-home message seems to be that there is no robust-take home message. Though basic questions go unanswered concerning the origins of inter-model contrast and mechanisms of change, the broader community has also struggled to answer these questions and so this work is not unique in this regard. That said I do have some minor suggestions for improvement. This includes the general suggestion that multi-model means not be used for many of the metrics because of the disproportionate variance in some models that is swamping out the means. Rather I think medians make more sense since the broader goal is to make generalizations about model behavior, which implicitly seeks to screen out outliers. I have various other relatively minor suggestions listed below but otherwise view the manuscript as suitable for publication.

\*\*\*

55: There have been various studies that show the improvement of ENSO simulation across CMIP generations. These seem useful to reference in this paragraph to provide context on the numerous SMILES.

119: I appreciate that there can be useful examples of models that suggest that no simple relationship between the present day and future exists but perhaps a more thorough exploration across all models and metrics considered here would be more convincing?

126: It is worth calling out the very large differences between SSP370 and RCP85/SSP585

in regarding to early 21st sulfate aerosols and potential consequences for the evolution of ENSO.

Figure 1: There is substantial white space in the 3x5 layout of the figure. I recommend changing to 4x4 and ensuring there is little white space as the figure is somewhat inefficient and difficult to read as is. Also it looks as though the vertical extent needs to be expanded as some lines go out of range. Note also that there is substantial noise in many of the time series. I suggest applying a smoother except where the variability is not irrelevant noise. That said, various models seem to have abrupt responses to forcing, such as volcanic eruptions and perhaps even biomass effects (in MIROC-ES2L), that is too abrupt to be explained by warming alone. The authors seem not to find this worthy of discussion? I recommend addressing it.

Figure 3: The multi-ensemble mean seems to be dominated by EC-Earth3. Would a multi-ensemble median perhaps be more appropriate?

Figure 4: Fonts are too small - minimum font including axes should be on par with main text. Monthly stddev lines should be make thicker.

The authors don't provide any hypotheses for the seasonality of the change in variance? Do none exist? Again the MIROC-ES2L increase at 2000 is quite notable. Is there no explanation for why this may occur? Other models show periodic changes in variability in the future. Is this just noise? Does it suggest that again some additional smoothing is needed to deal with some small ensemble sizes using monthly data? If one looks at CESM2 there are again suggestions of periodicity? What might drive this? I suggest reducing the range of the color bar to make colors in the figure more visible.

Figure 5: Perhaps put whiskers on Fig 5 corresponding to the 2 standard error range, which seems to increase in the latter half of the year? Might multi-model median be more appropriate due to sensitivity of means to a single model?

Why are 99 members of CESM2 used for some figures and only 50 used for others? This is not discussed at all and contradicts with Table 1. I'd review all plots and ensure that the # of members used is consistent with Table 1.

Figure 6: I again question the use of multi model mean rather than median given the outsized influence of some models (CSIRO).

Figure 8: Since the abscissa is not symmetric I recommend that a vertical line at 0 be shown to avoid confusion. As is, the panels and particularly the top, are misleading.

Figure 8: Doesn't the fact that obs have become more La Niña like suggest that we should have seen a reduction in variance?

Figure 9: There should probably two more sets of arrows at the top of the plot saying La Niña-like and El Niño-like such that they are on each plot.

226: Why not also infer changes from the CESM1-SF (and now CESM2-SF if you do add CESM1 to the figure)?

Figure 10: reference to 'all models' is a bit confusing given there are only 2 models. perhaps state "both models" or better yet "both ensembles"? or one could include ALL CMIP6 DAMIP simulations to provide context for canesm5 and miroct6. Also I wonder why not show these in a similar fashion to Fig 9?

255: The discussions based on MEM should be reconsidered in the context of the median I

think, particularly when making categorical statements of models overall (since the mean is strongly weighted by only a few models with large variance and is therefore not representative).

I recommend that the paper include a discussion of possible mechanisms and paths forward for exploring.

It is unfortunate that there is so little consistency across models on many aspects. The reader is left to wonder a bit on what is the robust finding here relevant to nature, aside from little being consistent across models? Have the authors looked for connections between some of the metrics being shown across models (e.g. seasonality of mean state changes and ENSO variance)? Have the authors examined what systematic differences exist for models that do best in some metrics packages (e.g. CVDP) in the present day (if so it would be good to mention, if not it would be good to do)? or for which changes in variance in observations fall within the ensemble spread, though this may be a weak constraint but would still be good to mention.

304: Projections of nearly every climate quantity are nonlinear in time (since radiative forcing is nonlinear). Is there really an expectation of linearity? and if not is this really a significant result?

308: As stated earlier I'd avoid the MEM(ean) and use the median.

334: last sentence needs a period.

None of the panels in the paper are labeled (e.g. A, B, C, ...). I think they should be for easier reference in this and other manuscripts that may cite specific figure panels from this work.