

This is a thorough study investigating the variability of sea surface temperature simulated in the PlioMIP1 ensemble modeling study, focusing on ENSO, and much less North Pacific Variability and the Indian Ocean Dipole. I have a few comments that the author should address to improve the paper:

1. Nino3.4: The common index for Nino3.4 is its standard deviation ($^{\circ}\text{C}$) rather than its variance ($^{\circ}\text{C}^2$). See for example, the IPCC AR5 WG1, Chapters 9 and 14. The author should justify why presenting variance is a better metric for the Pliocene. Otherwise, I would recommend using the standard Nino3.4 index to make the results more relevant to future projections.
2. Long controls: Several of the models represented in this study have long preindustrial controls in the CMIP5 database. The robustness of the conclusions should be strengthened by calculating the ranges of interannual variability of ENSO, IOD, and PDO for such models, subsetting their entire time series for the different shorter segment lengths available in the PlioMIP database. As shown in Wittenberg (2009), models can exhibit strong intercentennial modulation of ENSO behavior even with fixed climate forcings. The author should demonstrate that the Pliocene Nino3.4 variance is outside the preindustrial control variance if calculated for different 100 or 200-year periods than those in the PlioMIP1 database.
3. PDO definition: How are grid cells with sea ice present handled when calculating the first EOF of the North Pacific (20-70N) SST? Does this affect the outcomes?