

The point cloud in Fig. 7 without the CESM model looks so scattered to me, that I decided to read out the data with an online tool (plotdigitizer.com) and perform the regression myself. When excluding the upper left point in Fig. 7 I find an r-value of only 0.12, and a p-value of 0.74, contradicting the authors' claim that the relationship is statistically significant without including the outlier model. From this quick analysis, I infer that LGM TS and ECS are uncorrelated or only very weakly correlated among models when excluding the CESM model, in contrast to what the authors asserted. Only when including the CESM model does the relationship pass basic tests of significance ( $r=-0.76$ , and  $p=0.01$ ). Furthermore, the slope changes drastically when including (slope=-0.29 K/K) vs. excluding (slope=-0.05) the outlier. Finally, the data seems to be different from what is shown in Fig. 13 of Renoult et al. 2023, where the statistically significant relationship in PMIP4 is obvious to me.

I attach my code and screenshots of how I picked out the data so that the authors can reproduce what I did. I would like to ask from the authors to point out where I am wrong (e.g., by sharing code), and how they come to the conclusion that the relationship between LGM TS and ECS is significant, "not sensitive to the inclusion of CESM2", and that "the inclusion (or non-inclusion) or CESM2 and/or the CESM model family leads to the same slope and intercept for the emergent constraint relationship", beyond a reference to Renoult et al. 2023.

The remaining points of mine were addressed in a way that I am satisfied with.