Comment on egusphere-2022-134
Anonymous Referee #1

The author shows that (according to his assessment) the results of the multi-model mean (MMM) Global Average Temperature from CMIP5 and CMIP6 models according to the different scenarios (RCPs and SSPs) is well approximated by applying a simple linear regression fitted to historical observations of temperature and CO2 emissions, since 1976 and through 2005 or 2014.

I’m sorry to say I don’t see the value of this work and I recommend rejection.

First, GCMs produce hundreds of self-consistent variables and phenomena, so to propose that this linear regression on global temperature has value in substituting for complex GCMs is preposterous.

The comparison to make is with energy balance models, that through a simple function, as simple as a linear regression but accounting for physics and, importantly, feedbacks can indeed replicate global temperature and not only multi-model means but individual models that differ at a basic level by different equilibrium climate sensitivity. Some of this simple models importantly account for possible changes in the carbon cycle strength with warming, something a linear regression fitted on a short observational period can never do. The community has also, for many years now, recognized a strong linear dependence
between global temperature and cumulative emissions. So, the idea that simple relationships can be found is not new, except here it is applied to the wrong pair.

The way the author shows agreement is extremely superficial and I question it at a fundamental level which is appreciable to the naked eye. The presentation of results is actually troubling in its hand-waviness. I don’t see good agreement in rates and even behavior for important scenarios like the highly mitigated (which are more and more the focus of the community). RCP2.6 stabilizes for CMIP5 MMM while it declines for the results of the regression, and strongly so. RCP4.5 MMM keeps increasing, even if at a slower rate, while the blue line stabilizes. The rate under RCP6.0 is completely different, with the two lines starting apart and ending close to one another. SSP1-2.6 rates of decrease are significantly different. SSP2-4.5 trajectories diverge at the end. I could go on. Why is the author not even bothering commenting on these discrepancies, let alone showing the behavior of residuals between the red and blue line, which would be obviously far from i.i.d., undermining the whole proposition of applying a linear model? The starting date of 1976 is also arbitrary and would demand some sensitivity analysis to the choice of the start date…but that is a detail that I don’t think is worth pursuing given the fundamental flaws of philosophy and performance one sees.