

Reviewer #1 (Comments to Author):

The submitted article attempts to quantify the reliability of climate model simulations through the construction of a linear model relating global mean temperature to net climate forcing, assessing the uncertainty in current climate forcing and propagating this error to future projections.

Unfortunately, the physical arguments made in the study show a significant misunderstanding of feedback processes in the climate system, and consistent misinterpretations of the current literature. The projected uncertainties from the study are several orders of magnitude larger than the CMIP5 spread arise from the unvalidated empirical model used in the study, as described below. This paper, in the reviewer's opinion, is not suitable for publication.

The core of the study forms a linear model which directly relates global mean temperatures to net forcing, with no discussion of the effect of ocean heat uptake. The model thus effectively assumes that the earth system has no thermal inertia. The validation of the model consists of plotting global mean temperatures calculated using the A2 scenario and idealized 1 percent annual increase simulations.

The fundamental error made in this study is the confusion of base-state forcing and feedback. The author attempts to measure uncertainty in present day cloud forcing by comparing ISCCP and MODIS derived cloud fractions with GCM output. There are numerous problems with this. Firstly, the cloud fraction diagnostics in GCM output are not directly comparable with the satellite products or each other, (GCM satellite simulators are available for both products, but the author has not used them). The author correctly ascertains that GCM derived cloud distributions have systematic errors, which is not unprecedented (Bony 2011 and references therein).

The author's first catastrophic error is to assume a linear relationship between global mean cloud fraction and net global cloud forcing - an assumption which is not justified in any way, but is clearly nonsense. A less wrong approach here would have been to use Partial Radiative Perturbation, Adjusted cloud radiative forcing or even just basic cloud radiative forcing as an estimate of the cloud radiative effect.

But, even ignoring the errors thus far and assuming the  $5\text{Wm}^{-2}$  uncertainty in cloud forcing is accurate, the overwhelming error in this paper is how this uncertainty in cloud forcing is applied in the future projections made using the empirical linear model. Each GCM starts simulations in ~1850 in an equilibrium state, thus all of the errors in base state cloud forcing are already represented in the global mean temperature in 1850. The author is implying that the uncertainty in cloud feedback (the derivative of cloud forcing with respect to global mean temperature) can be equated to the multi-model spread in present day cloud forcing without any justification. Of course, the spread of cloud feedbacks in present day GCMs is our dominant source of uncertainty in future global

mean climate response, but it is critically the changes in forcing in each model as a function of temperature, and not the absolute values, which are of relevance for climate sensitivity to greenhouse gas forcing.

An entirely equivalent argument would be to say (accurately) that there is a 2K range of pre-industrial absolute temperatures in GCMs, and therefore the global mean temperature is liable to jump 2K at any time - which is clearly nonsense, but no less so than the arguments presented in this paper.

For these reasons, I consider the study to be incorrect and unpublishable. There are however, several other minor issues, for reference.

The author consistently ignores all literature in feedback theory. Water vapor feedback is well understood and remarkably consistent between present day GCMs (see Soden et al (2011), Colman (2002) etc.), and the spread in water vapor effect referred to in the model is due to the realization of other feedbacks on temperature which then influences water vapor forcing.

The 3 pages of fundamental radiative transfer in the paper dedicated to deriving the log relationship between CO<sub>2</sub> concentrations and climate forcing is completely unnecessary, and adds nothing to existing textbook literature.

The literature review is incomplete and misrepresentative. The IPCC is portrayed as making conclusions based on model results alone, and does not reflect multiple lines of observable historical data which can constrain large scale climate parameters. The repeated statement that Knutti (2008) considers only model variability is simply incorrect, as the paper discusses a wide range of systematic and parametric uncertainties in GCM projections [line 61 and others]. Stainforth et al (2005) has nothing to do with internal variability [line 62], and discusses the relationship between parameter uncertainty and climate sensitivity. The repeated statement that no prior papers have discussed propagated error in GCM projections is simply wrong (Rogelj (2013), Murphy (2007), Rowlands (2012)).

Rowlands, Daniel J., et al. "Broad range of 2050 warming from an observationally constrained large climate model ensemble." *Nature geoscience* 5.4 (2012): 256-260.

Murphy, James M., et al. "A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1857 (2007): 1993-2028.

Rogelj, J., D. L. McCollum, A. Reisinger, M. Meinshausen and K. Riahi (2013). "Probabilistic cost estimates for climate change mitigation." *Nature* 493(7430): 79-83. (HTML)

Colman, Robert. "A comparison of climate feedbacks in general circulation models." *Climate Dynamics* 20.7-8 (2003): 865-873.

Soden, Brian J., Isaac M. Held, Robert Colman, Karen M. Shell, Jeffrey T. Kiehl, Christine A. Shields, 2008: Quantifying Climate Feedbacks Using Radiative Kernels. *J. Climate*, 21, 3504-3520.

Bony, Sandrine, et al. "CFMIP: Towards a better evaluation and understanding of clouds and cloud feedbacks in CMIP5 models." *Clivar Exchanges* 16.56 (2011): 20-24.