Reviewer comments are in black and our reply is in blue.

This manuscript used a multiple linear regression energy balance model, EM-GC, to estimate the attributable anthropogenic warming rate (AAWR), the equilibrium climate sensitivity (ECS), and the future projections. The authors compared the results from EM-GC with those obtained from CMIP6. They found that the CMIP6 GCMs tend to exhibit a faster rate of warming, which induced larger AAWR, larger ECS, and smaller remaining budgets of carbon emissions. One highlight of this work is the use of Aerosol Weighting Method, which allowed a probabilistic estimation. This work is very interesting and the authors have done many detailed analyses. However, before I can recommend accepting this manuscript, I have several concerns that need to be addressed.

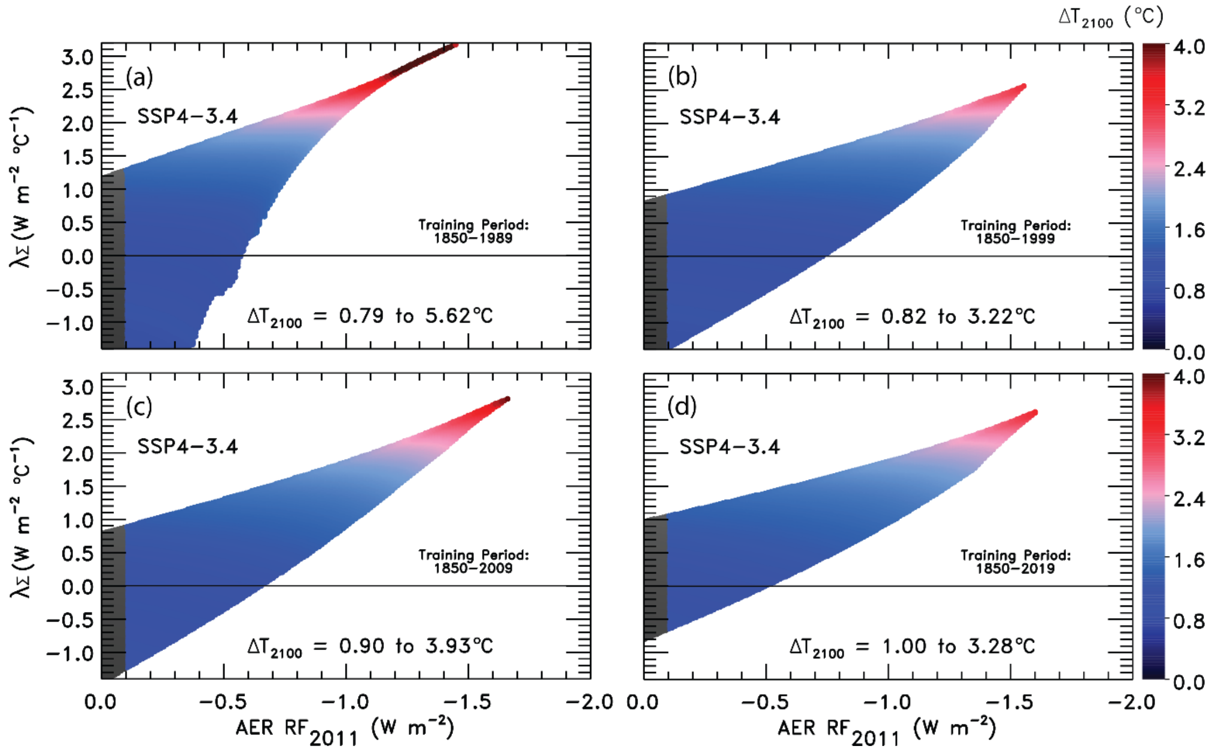
> We thank the reviewer for taking the time to carefully read our manuscript and suggest many useful changes. Upon revision, we will make changes to the paper to address all of these comments, as detailed below.

1. To run the EM-GC model, it seems that one needs to determine nine regression coefficients and parameters. Constrained by the observed GMST and the OHC, one can obtain a set of the nine coefficients/parameters to ensure a good fit to the historical observations. However, I am not sure if the selected set of coefficients/parameters is unique, or one can use a totally different set of coefficients/parameters to achieve a similar fitting skill? I also have concerns that whether the coefficients/parameters are still useful for the future projections? I would like to suggest the author to perform a test to prove the validity of the model and the stability of the coefficients/parameters. For example, the authors may consider to divide the historical period into two halves, use the first half to determine the coefficients/parameters, and use the second half to test the stability.

> Great suggestion, which we plan to address upon revision.
>
> The only parameters important for the future projections of GMST are the climate sensitivity parameter, γ, and the ocean heat uptake efficiency, κ. The regression coefficients C_0-C_6 in Eq. (2), which modify natural drivers of climate variability, are only used to simulate the observed change in GMST from 1850-2019. However, future temperature projections consider only the anthropogenic components governed by RF due to GHGs, aerosols, as well as climate feedback (related to γ) and ocean heat uptake (related to κ). We are able to obtain much better fits to the actual climate record upon consideration of the full range of natural drivers of climate variability; hence, the inclusion of C_0-C_6 allows for a more realistic evaluation of the range of model parameter space for γ and κ under which "good fits" to the prior climate record can be obtained.
>
> We have taken the suggestion from the reviewer to alter the training period of our model to test the stability and propose to show these results in a new Supplemental Figure. New Fig. S2 shows the projections of the change in GMST in 2100, ΔT_{2100}, as a function of the climate feedback parameter, λ_{Σ}, and the value of aerosol radiative forcing in 2011, AER RF_{2011}, for 4 different training periods: 1850-1989 (New Fig. S2a), 1850-1999 (New Fig. S2b), 1850-2009 (New Fig. S2c), and 1850-2019 (New Fig. S2d), which is the normal training period used in our analysis. Values of ΔT_{2100} are shown only for combinations of

λ_Σ and AER RF_{2011} (value of aerosol radiative forcing in 2011) that lead to good fits to the climate record, which means values of the three reduced chi-squared (χ^2) parameters are all less than or equal to 2. We project relatively similar results for end of century warming for the training periods that end in 2019, 2009, and 1999. The training period that ends in 1989 (New Fig. S2a) yields a different "shape" of model parameter space for which good fits to the climate record can be obtained, compared to the other training periods. The different shape for this shorter training period is due to the formulation of the ocean component of our model. In training to 1989, we are only considering 35 years of the observed OHC record. We are able to calculate good fits to the OHC record over this shorter time period that diverge from the OHC record after 1989. The highest values of ΔT_{2100} in New Fig. S2a are associated with the largest values of λ_Σ, which in our model corresponds to excessively high values of κ that we can rule out, based on OHC data collected during 1990 to 2020.

We propose upon revision to add a paragraph to Sect. 2.1 to the paper noting the stability of the forecasts of end-of-century warming for the training periods of 1850-1999, 1850-2009, and 1850-2019, with most of the words supporting this finding appearing in the revised Supplement along with New Fig. S2.



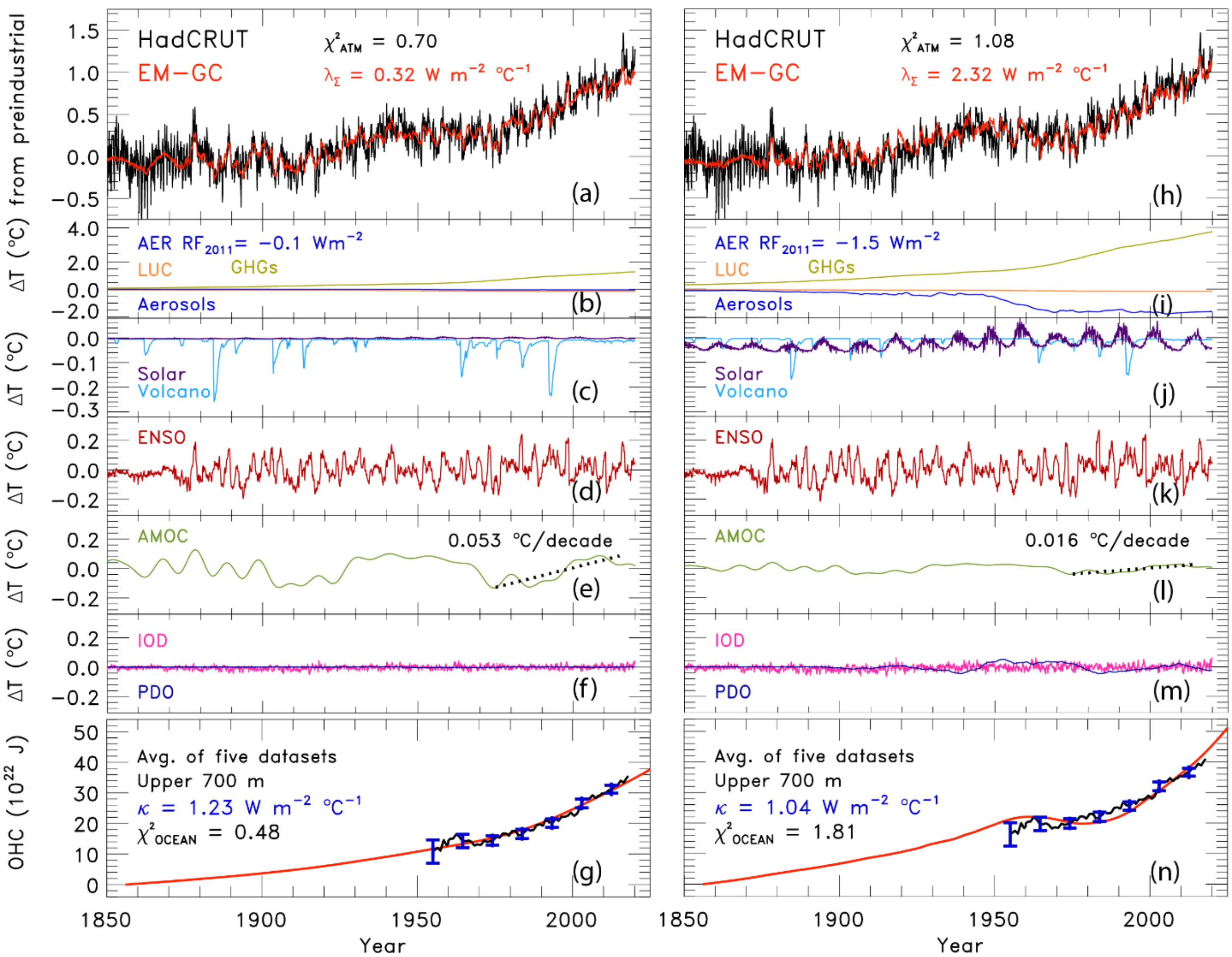
New Figure S2. ΔT_{2100} as a function of climate feedback parameter and tropospheric aerosol radiative forcing in 2011 using the EM-GC for SSP4-3.4. (a) Training period of 1850-1989. The region outside of the AER RF_{2011} range provided by IPCC 2013 is shaded (grey). Colors denote the GMST change in year 2100 relative to pre-industrial. The color bar is the same across all four panels for comparison. (b) Training period of 1850-1999. (c) Training period of 1850-2009. (d) Training period of 1850-2019, which is the normal training period used in our analysis.

2. From Fig. 1f, the authors found that the PDO has very limited contributions to the GMST. I don't understand this finding, as to my knowledge, the different phases of the PDO play an important role in modulating the GMST. For example, the recently well discussed warming hiatus in the beginning of this century has been found to be closely related to the PDO. An explanation about the findings in Fig. 1f is needed.

Another great suggestion, which we also plan to address upon revision.

In Fig. 1f of the submitted paper, we had shown the model run for the best estimate of AER RF_{2011} = −0.9 W m^{-2}. In this case, the PDO exhibits less influence on GMST than we find for AMOC. If we vary the value of AER RF_{2011}, which impacts the time series of aerosol RF of climate over the entire simulation, we find there are some model runs for which the PDO has the same or even larger influence on GMST compared to AMOC.

Upon revision, we propose to include the figure shown below as New Fig. S7, and modify the text in the paper to make clear that the expression of the PDO on GMST in our model framework is dependent on model specification of the aerosol RF of climate time series. At low values of AER RF_{2011}, the effect of PDO on GMST (New Fig. S7f) is negligible and the contribution from AMOC dominates over PDO or IOD. At high values of AER RF_{2011}, the effect of PDO on GMST (New Fig. S7m) is equal to the contribution from AMOC. Upon revision, we will add a new paragraph to Sect. 2.2.6 discussing the importance of PDO at higher values of AER RF_{2011} and include New Fig. S7 in the supplement with citations to England et al. (2014) and Trenberth and Fasullo (2015). The figure below is a robust result: the larger the scaling factor for aerosol RF, the greater the influence of PDO.

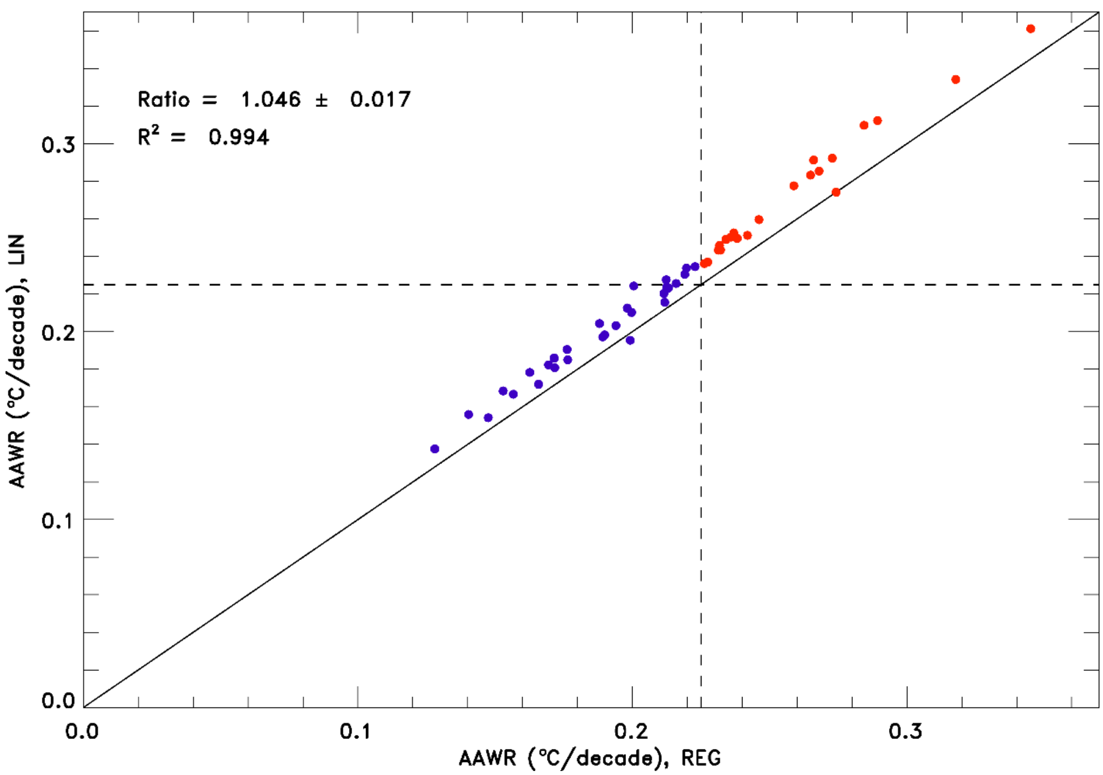


New Figure S7. Measured and modeled GMST anomaly (ΔT) relative to a pre-industrial (1850-1900) baseline for an AER RF_{2011} = −0.1 W m^{-2} and −1.5 W m^{-2}. (a) Observed (black) and modeled (red) ΔT from 1850-2019. This panel also displays the values of λ_Σ and χ^2_{ATM} (see text) for this best-fit simulation. (b) Contributions from total human activity. This panel also denotes the numerical value of the attributable anthropogenic warming rate from 1975-2014 (black dashed) as well as the 2σ uncertainty in the slope. (c) Solar irradiance (light blue) and major volcanoes (purple). (d) Influences from ENSO on ΔT. (e) Contributions from AMOC to ΔT and to observed warming from 1975-2014. (f) Influences from PDO (blue) and IOD (pink) on ΔT. (g) Measured (black) and modeled (red) ocean heat content (OHC) as a function of time for the average of five data sets (see text), the value of χ^2_{OCEAN} for this run, as well as the ocean heat uptake efficiency, κ, needed to provide the best-fit to the OHC record. The error bars (blue) denote the uncertainty in OHC used in this analysis (see Sect. 2.2.8). (h)-(n) Same as (a)-(g), except for AER RF_{2011} = −1.5 W m^{-2}.

3. Another concern is about the comparison of the AAWR that obtained from EM-GC and CMIP6 models. Since different methods are used to calculate the AAWR, I am not sure if the results are comparable. Especially for the CMIP6 models, the REG method seems to be too simple to calculate the AAWR. I am not sure if the AAWR values obtained from the CMIP6 models are as pure as those obtained from EM-GC.

We plan, upon revision, to add much more detail regarding how the attributable anthropogenic warming rate, AAWR, is estimated from CMIP6 GCM output.

In Sect. 2.3 of the submitted paper, we discuss two methods to determine the AAWR from the CMIP6 models, REG and LIN. REG is a regression-based approach and LIN is a linear fit method. For the GCM-based estimates of AAWR that appeared in the submitted paper, the LIN method tends to result in very slightly higher values than REG, as shown in Response Fig. 1.

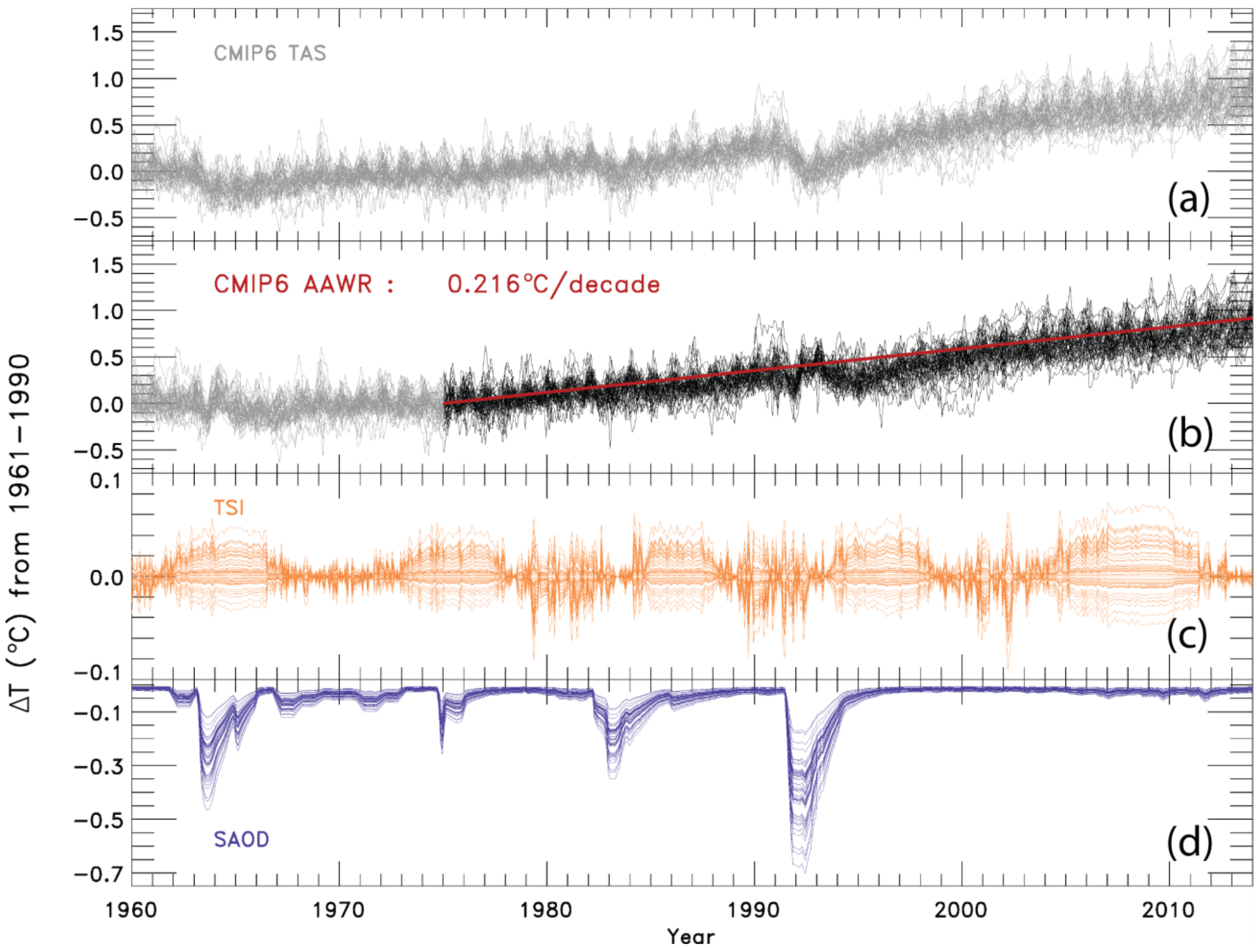


Response Figure 1. Values of AAWR for 50 CMIP6 GCMs using the LIN and REG methods. The solid black line is the 1:1 line and the vertical and horizontal dashed lines are the maximum value of AAWR determined using the EM-GC and the HadCRUT temperature record. The CMIP6 GCMs that have values of AAWR less than the maximum value from the EM-GC are blue, and the CMIP6 GCMs that have values of AAWR greater than the maximum value from the EM-GC are red. The slope, 1σ standard deviation, and R^2 of the values of AAWR from the CMIP6 GCMs are shown.

The values of AAWR determined by the LIN method are about 4% higher than the values of AAWR determined by the REG method. The close agreement of AAWR found using both methods provides strong evidence that we have correctly extracted this important quantity from the CMIP6 archive.

We have further examined our calculation of AAWR using the REG method in response to the reviewer's comment and have a few proposed changes that lead to a more robust estimate that we will implement upon revision.

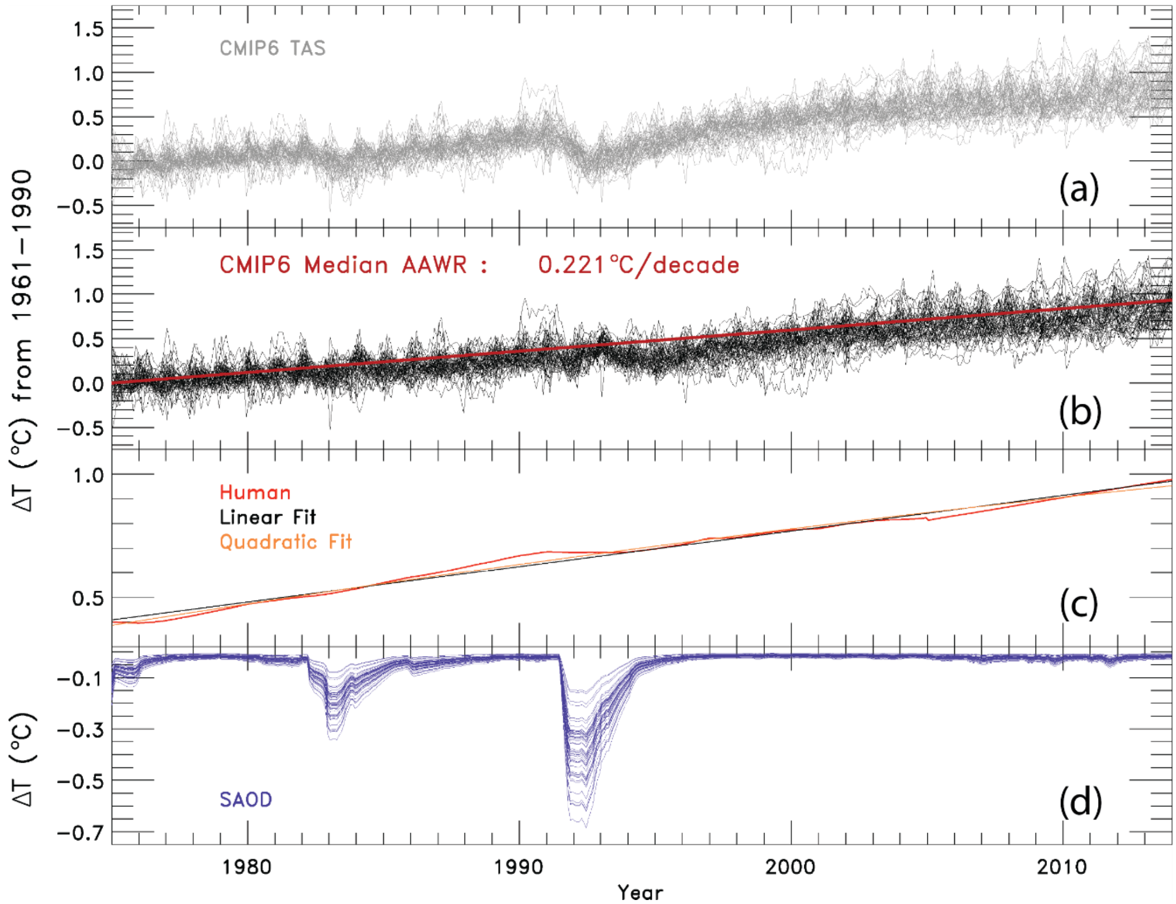
As detailed below, close examination of the CMIP6 GCM output, shows that the representation of the effect of variations in total solar irradiance, TSI, on global mean surface temperature (GMST) in the GCMs leads to a regression coefficient that seems to be randomly distributed (see Response Fig. 2).



Response Figure 2. The change in GMST relative to 1961-1990 from the CMIP6 GCMs and the contribution from TSI and SAOD from 1960-2014. (a) The change in GMST from the 50 CMIP6 GCMs. (b) The residual in the change of GMST from the 50 CMIP6 GCMs after subtracting the contribution of TSI and SAOD determined by the REG method. The median value of AAWR is written on this panel and plotted in red. (c) The contribution of TSI in the 50 CMIP6 GCMs. (d) The contribution of SAOD in the 50 CMIP6 GCMs.

Response Fig. 2c shows the random representation of TSI in the CMIP6 GCMs. Upon the implementation of the REG method, some CMIP6 GCMs obtained negative coefficients for TSI, and others obtained positive coefficients. For some reason, many GCMs do not represent the impact of variations in solar output on GMST in a manner that mimics the actual, observed relation. There is extensive literature on possible reasons TSI affects GMST, implicating causal factors such as cosmic-ray influence on cloud nucleation, that is nicely summarized at https://skepticalscience.com/cosmic-rays-and-global-warming-advanced.htm. If the true causal factor involves something like cosmic rays, this process will likely not be present in most GCMs. Because of the varying nature of TSI in the GCMs, we propose to update our calculation of REG to not include TSI in the regression.

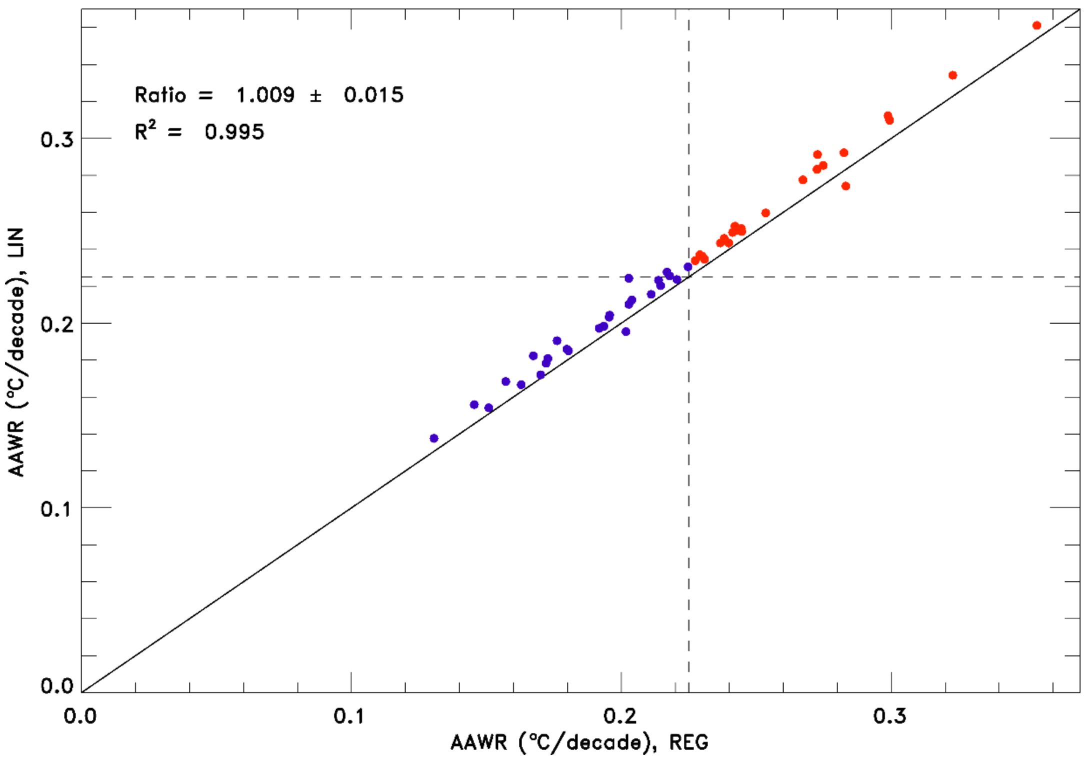
We propose to alter the REG method in the following way. We will conduct one regression from 1975-2014, instead of two regressions as had been explained in Sect. 2.3 of the submitted paper. We will exclude TSI as a regressor and only include stratospheric aerosol optical depth (SAOD) and a linear function to represent the contribution of humans to the change in GMST. For SAOD, we will determine the appropriate lag for each model that results in the largest coefficient, to accurately represent how long it takes for the effect of SAOD to have on GMST within each model. Using this new REG method results in very slightly different values of AAWR compared to those in the submitted paper, as shown in New Fig. S11 and New Fig. S12 below.



New Figure S11. The change in GMST relative to 1961-1990 from the CMIP6 GCMs and the contribution from TSI and SAOD from 1975-2014. (a) The change in GMST from the 50 CMIP6 GCMs. (b) The residual in the change of GMST from the 50 CMIP6 GCMs after subtracting the contribution of SAOD determined by the updated REG method. The median value of AAWR is written on this panel and plotted in red. (c) The human component of global warming, $\Delta T_{ATM,HUMAN}$, from the EM-GC. A linear fit (black) and quadratic fit (red) are plotted on top to show that $\Delta T_{ATM,HUMAN}$ is almost exactly linear. (d) The contribution of SAOD in the 50 CMIP6 GCMs using a lag month calculated for each model.

A comparison of New Figure S11 to Response Figure 2 shows that AAWR found using the REG method is not much affected by removing TSI as a regressor. The values of AAWR determined from the CMIP6 GCMs are more similar to the values determined by the LIN method, under this new approach. New Figure S12 shows that there is now a 0.9% difference between the values of REG and LIN. New Figure S11c shows the human component of global warming, $\Delta T_{ATM,HUMAN}$, from the EM-GC. A linear fit and quadratic fit were taken of $\Delta T_{ATM,HUMAN.}$ The linear fit and quadratic fit are very similar, indicating that $\Delta T_{ATM,HUMAN}$ is in fact nearly linear over this period of time. This result justifies our approach of approximating a linear function to represent $\Delta T_{ATM,HUMAN}$ in the AAWR calculation.

Upon revision, these figures will be noted in Main, and detail will be added to the Supplement to document our procedure for finding AAWR from the GCMs, allowing the reader to better assess the procedure and, in our view, accurate rendering of this quantity from the CMIP6 archive.



New Figure S12. Values of AAWR for 50 CMIP6 GCMs using the LIN and REG methods. The solid black line is the 1:1 line and the vertical and horizontal dashed lines are the maximum value of AAWR determined using the EM-GC and the HadCRUT temperature record. The CMIP6 GCMs that have values of AAWR less than the maximum value from the EM-GC are blue, and the CMIP6 GCMs that have values of AAWR greater than the maximum value from the EM-GC are red. The slope, 1σ standard deviation, and R^2 of the values of AAWR from the CMIP6 GCMs are shown.

4. In line 228, "…also specified on Fig. 1f", "Fig. 1f" should be "Fig. 1e".

   Thank you, we will fix.

5. In line 975, "then" should be "than".

   Thank you, we will fix.

6. In line 1061, "…of the Paris Agreement will be achieved", "will be" should be "will not be".

   Thank you, we will fix.